L2 ACQUISITION OF SPANISH MIDDLE MARKER *se* BY NATIVE SPEAKERS OF ENGLISH AND KOREAN

by

JIHEE HWANG

(Under the Direction of Margaret Quesada)

ABSTRACT

This dissertation examines the acquisitional process of middle *se* constructions by English- and Korean-speaking L2 learners of Spanish. Although clitic *se* is well known for its multifunctionality, there is a general tendency in the literature to analyze any use of *se* as a reflexive marker. This tendency is especially prominent in the presentation of *se* in instructional textbooks. However, this analysis is misleading because it fails to explain the use of *se* in non-reflexive contexts. From the framework of cognitive linguistics, it is suggested that non-reflexive *se* is a middle marker whose main function is to focus on the central properties of events, rather than on the subject or object of the sentence. By determining the middle marking function of *se*, all non-reflexive uses can be generalized into one category, which should not be considered "exceptional cases" for learners to memorize.

The explanatory power of the cognitive approach on middle *se* is supported by some L1 acquisition research; however, there is still a lack of L2 studies to determine its applicability. The present dissertation aims to fill this gap in the L2 acquisition of middle *se*, by comparing two L2 learner groups with different L1 backgrounds, English and Korean, with a group of native Spanish speakers. Since both English and Korean middles are distinct from the Spanish middle

system, discrepancies between native and non-native performance were expected. Data are collected via three tasks (a picture book written narration task, an acceptability judgment task, a cloze task) designed to examine how the middle marker *se* is produced and interpreted by three different groups.

Results show that native and non-native speakers do not differ significantly in their production of *se* constructions. It was found that telicity is the main factor which affects *se*-marking. Nonetheless, differences were shown in the interpretation data: the association between middle *se* and the middleness of the context was stronger for native speakers. Also, the higher the proficiency of the non-native speakers, the more native-like their performance was. Finally, the difference between L1s of the two learner groups did not demonstrate any important significance.

INDEX WORDS: Middle voice, Middle constructions, Middle marker, Reflexives, Spanish Clitic *se*, Acquisition of Spanish clitic *se*, Spanish Second Language Acquisition, Cognitive Linguistics

L2 ACQUISITION OF SPANISH MIDDLE MARKER *se* BY NATIVE SPEAKERS OF ENGLISH AND KOREAN

by

JIHEE HWANG

B.A., Seoul National University, Republic of Korea, 2013

M.A., Seoul National University, Republic of Korea, 2015

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2021

L2 ACQUISITION OF SPANISH MIDDLE MARKER *se* BY NATIVE SPEAKERS OF ENGLISH AND KOREAN

by

JIHEE HWANG

Major Professor:	Margaret Quesada
Committee:	Timothy Gupton
	Chad Howe

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2021

ACKNOWLEDGEMENTS

This dissertation has been possible thanks to the people who have supported me at all times, especially during the past five years of my life at the University of Georgia. First of all, I would like to express my deepest appreciation to my major professor, Dr. Margaret Quesada. If there is one person who deserves all possible thanks, that is you. Without your constant guidance, thoughtful feedback, logical and reasonable suggestions, and your encouragement and support throughout the entire process, this dissertation could not have been completed. Thank you for trusting me, for motivating me, for always supporting me, and above all, for being the perfect role model and advisor that I could have ever asked for. It has been an honor being your advisee and being able to work with you. I would also like to extend my gratitude to my committee members, Dr. Timothy Gupton and Dr. Chad Howe for agreeing to be part of this dissertation and for the support that I have received from you. What I have learned from your lectures, conversations and comments has been fundamental for my academic growth in the doctorate process.

I'm deeply indebted to my *Ling Ladies*, Bethany Bateman, Camila Lívio, Madeline Critchfield and Sarah Hubbel. I would not have been able to come this far without you. Thanks for being always there, no matter what. In particular, nothing would have been possible without the support, proof reading, grammar checking, and countless hang-outs and sleepovers with Bethany. I have always wondered how you could give me everything you have and still want to give me more. I love you so much, period (there's no question mark here). I'm also extremely grateful to the families of my ladies —Dr. Jake McDonald and Nolan, Dr. Derek Beaudry and

Tessa, Ramiro Calero Granera and Leo— and my family in the US, the Batemans. You all filled my heart with love and support that always made me feel like home, and I am so lucky to be a part of your families. My thanks also go to María González, who has never wavered in her support. I cannot forget my cohorts, Shannon McDonough, Jessica Tiegs, Ingrid Abisambra, and Javier Cabezas Zapata, who welcomed me with open arms when I arrived at UGA in 2016. I also wish to thank Hyunjin Lee, my one and only "Korean friend" in Athens.

I must thank my most influential mentor, Dr. Minwook Oh; my other half of *DoD*, Dahee Ahn; and my reliable *unnies*, Dr. Min Ji Kang and Dr. Sara Choe. Your love, guidance, and advice have cheered me up and helped me persevere. I will never forget your support and, of course, the moments of laughter that we shared through endless calls and texts.

I'd also like to extend my gratitude to the professors and instructors at Seoul National University, to whom I owe so much and without whom I would possibly never have chosen this path. I want to thank all of you for your help and support in all aspects of both my career and personal life. I appreciate the encouragement and warmth that you have given me for the past 13 years starting from the freshman year of my undergraduate studies.

I cannot finish my acknowledgements without expressing my thanks to my family and friends who have been waiting for me since the first day I left home. I thank every single one of you for telling me that you love me, for showing me your love, and for letting me know what loving and caring is. You have given me so much that I was never lonely, even when I was feeling homesick. Everything has been possible thanks to you. Finally, special thanks to my mom, Minja Seo, to whom I dedicate this work with all my love.

Well, I just wanted to say that I love each of you. That is all.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iv

LIST OF TABLES viii

LIST OF FIGURES xi

CHAPTER

1 INTRODUCTION 1

1.1 Statement of problem and goals for the study 1

1.2 Outline of the dissertation 6

2 SPANISH MIDDLE MARKER *se* 8

2.1 Introduction 8

2.2 Traditional grammarians' explanations of the clitic *se* 9

2.3 Generative explanations of the clitic *se* 12

2.4 A cognitive perspective of the clitic *se* 25

2.5 Middle voice constructions in English and Korean 47

3 LANGUAGE ACQUISITION AND CLITIC *se* 57

3.1 Introduction 57

3.2 L1 acquisition and native speakers' use of Spanish clitic *se* 57

3.3 L2 acquisition of Spanish clitic *se* 68

3.4 Textbook analysis on Spanish *se* constructions 93

4 METHODOLOGY 105

4.1 Introduction 105
4.2 Research Questions 106
4.3 Participants 111
4.4 Survey design and data collection 117
5 RESULTS AND DISCUSSION 132
5.1 Introduction 132
5.2 Task 1 (Picture book written narration task) Results — identifying variables that favor *se*-marking 133
5.3 Task 2 (Acceptability Judgment Task, AJT) Results — testing learners' knowledge of absolute *vs.* dynamic verbal events 157
5.4 Task 3 (Cloze task) Results — testing learners' understanding of reflexives *vs.* middles 195
6 CONCLUSIONS 215
6.1 Summary and general conclusions 215
6.2 Contributions and future directions 226
REFERENCES 229
APPENDICES
A Language background questionnaire 241
B Spanish proficiency test 243
C Task 1 249
D Task 2 251
E Task 3 261

LIST OF TABLES

Page

Table 1: Spanish clitic pronouns....9

Table 2: Types of Spanish clitic *se* 11

Table 3: Classification of *se* constructions based on [±argument] feature 17

Table 4: Spanish basic middle voice constructions 41

Table 5: Examples of error correction by native Korean speakers....54

Table 6: Comparison of Spanish, English and Korean middle voice systems....55

Table 7: Percentage of *se* tokens as reflexive, lexical and middle forms 59

Table 8: L1 studies of the acquisition of Spanish clitic *se*....65

Table 9: L2 studies of the acquisition of Spanish clitic *se*....89

Table 10: Textbooks analyzed for the present study 94

Table 11: Examples of verbs that change meaning and verbs of *becoming* 97

Table 12: Summary of the uses of *se* 99

Table 13: Comparison of *se* introduction in four textbooks 103

Table 14: Summary of participant demographics....114

Table 15: Participant frequencies according to task 117

Table 16: Distribution of responses by all participant groups (Task 1)....134

Table 17: Frequently *se*-marked verbs by all participant groups (Task 1) 135

Table 18: Independent variables analyzed for Task 1 139

Table 19: Summary of L1 English group's Rbrul results....142

Table 20: Summary of L1 Korean group's Rbrul results 146

Table 21: Summary of L1 Spanish group's Rbrul results.....149

Table 22: Summary of the Rbrul results for all groups 152

Table 23: Distribution of *se*-marking according to *L1* and *Tense-Aspect-Mood*.....155

Table 24: Distribution of responses by L1 English speakers (Task 2) 159

Table 25: Distribution of responses according to context and proficiency (L1 English group)..161

Table 26: L1 English speakers' preference for (non-)*se*-marked option according to contexts..163

Table 27: Mean and median scores of L1 English speakers' performance on Task 2.....164

Table 28: L1 English speakers' distribution of responses by verb in dynamic contexts.....168

Table 29: Distribution of responses by L1 Korean speakers (Task 2).....172

Table 30: Distribution of responses according to context and proficiency (L1 Korean group) ..174

Table 31: L1 Korean speakers' preference for (non-)*se*-marked option according to contexts ..176

Table 32: Mean and median scores of L1 Korean speakers' performance on Task 2 177

Table 33: L1 Korean speakers' distribution of responses by verb in dynamic contexts 180

Table 34: Distribution of responses by L1 Spanish speakers (Task 2).....183

Table 35: Distribution of responses according to context (L1 Spanish group) 184

Table 36: Participants' preference for (non-)*se*-marked option according to contexts 187

Table 37: Mean and median scores of L1 Spanish speakers' performance on Task 2 188

Table 38: Distribution of responses of verbs that showed unexpected results (Task 2).....191

Table 39: Participants' preference for (non-)*se*-marked option according to contexts with 193

Table 40: Distribution of responses by L1 English speakers (Task 3) 196

Table 41: L1 English speakers' acceptance of *a sí mismo* in daily routine middle context 199

Table 42: L1 English speakers' distribution of responses by verbs (Task 3) 200

Table 43: Distribution of responses by L1 Korean speakers (Task 3).....201

Table 44: L1 Korean speakers' acceptance of *a sí mismo* with non-routine verbs.....204

Table 45: L1 Korean speakers' distribution of responses by verbs (Task 3).....205

Table 46: Distribution of responses by L1 Spanish speakers (Task 3).....207

Table 47: Participants' acceptance of *a sí mismo* in daily routine reflexive context.....210

Table 48: Distribution of responses of the verbs that showed unexpected results (Task 3).....211

LIST OF FIGURES

Page

Figure 1: Absolute constructions 43

Figure 2: Middle/Dynamic constructions 43

Figure 3: Excerpt of the survey's front page 117

Figure 4: Excerpt of Task 1 119

Figure 5: Excerpt of Task 2 124

Figure 6: Excerpt of Task 3 128

Figure 7: Association plot of *se*-marking according to *L1* and *Tense-Aspect-Mood*.....155

Figure 8: Mosaic plot of responses by L1 English speakers (Task 2) 160

Figure 9: Frequency of responses according to context by L1 English speakers (Task 2).....162

Figure 10: Box plot of scored performance of L1 English speakers (Task 2).....165

Figure 11: Significance of relevance between factors by L1 English speakers (Task 2) 166

Figure 12: Mosaic plot of responses by L1 Korean speakers (Task 2).....173

Figure 13: Frequency of responses according to context by L1 Korean speakers (Task 2) 174

Figure 14: Box plot of scored performance of L1 Korean speakers (Task 2) 178

Figure 15: Significance of relevance between factors by L1 Korean speakers (Task 2).....179

Figure 16: Mosaic plot of responses according to L1 (Task 2) 184

Figure 17: Count of responses by all participant groups (Task 2).....186

Figure 18: Box plot of scored performance of all participants (Task 2).....188

Figure 19: Significance of relevance between response, L1, and context (Task 2).....190

Figure 20: Count of responses by L1 English speakers (Task 3) 197

Figure 21: Significance of relevance between factors by L1 English speakers (Task 3) 198

Figure 22: Count of responses by L1 Korean speakers (Task 3).....202

Figure 23: Significance of relevance between factors by L1 Korean speakers (Task 3).....203

Figure 24: Count of responses by all participant groups (Task 3).....208

Figure 25: Significance of relevance between response, L1, and context (Task 3).....209

CHAPTER 1

INTRODUCTION

1.1 Statement of problem and goals for the study

This dissertation aims to determine how the middle marker *se* is used, interpreted and acquired by adult English and Korean L2 learners of Spanish. The present study adopts the definition of 'middle' following the notation of cognitive linguistics (Langacker 1987, 1991) based especially on the related analyses of Maldonado (1992; 2000; 2008; 2009, among others). The middle voice is considered to be situated in a medial position between active and passive voices, with a structure that shows different characteristics from both of the voices. In particular, Maldonado (1992:238) views the middle as "a gradual organization in which the prototype occupies an intermediate place", where the emphasis focuses[1] on the event itself rather than on the subject or on the object of the sentence. Accordingly, middle is indeed a separate category of the voice system, not a subpart of passive voice. Also, it is of importance to understand that middle voice affords relatively little attention to the participants of a verbal event and, in contrast to active and passive constructions, has a perspective centered on the event itself (cf. Calude 2017).

[1] Note that the term *focus* differs from the '(narrow) focus' under the concept of information structure. For the purpose of avoiding confusion, the present study will use "event-focus" whenever possible, in order to describe a context where the proposition or the event expressed in the sentence is more emphasized than the other components or arguments (e.g. agent, theme/patient, goal/recipient) of the sentence.

It has received attention in the literature on Spanish L2, which suggests that learners have difficulty learning and producing grammatically correct sentences with clitic *se* (Sanromán Vilas 2009; Escutia López 2010; Araya & Monteserin 2011; Guillén Solano 2015). One plausible reason why learners experience confusion with *se* constructions is the disparate grammatical definitions of such constructions. Defining what clitic *se* is or which role it plays has been a difficult task in the linguistics literature, and there still remain some specifications and (re)conceptualizations to be made. There are two major problems that traditional approaches assume, which will be discussed in more detail in the following chapter: (i) the delimitation of Spanish middle construction to a specific type of phrase, dispositional middles (e.g. *Este libro* ***se*** *lee fácilmente* 'This book reads easily'); and (ii) the analysis of middle marker *se* as a reflexive marker.

First of all, Spanish middle voice constructions were traditionally identified with dispositional middles in which the existence of an adverbial phrase such as *fácilmente* 'easily' or *con facilidad* 'with ease' is a prerequisite. According to Maldonado, this type of construction needs to be considered middle indeed. However, he emphasizes Spanish middle constructions should not be restricted exclusively to this specific type. It is argued that Spanish has a wider range of constructions that belong to the middle voice.

Moreover, the clitic *se* has traditionally been analyzed as a reflexive marker with a valence reducing or detransitivizing effect. For this analysis to be applicable to all instances of *se*, it has to be assumed that the clitic *se* only occurs in transitive constructions. However, there are various examples of intransitive *se* constructions (e.g. ***Se*** *fue* 'S/he left') which the detransitivizing approach fails to explain. Moreover, the valence reducing effect does not always

occur in transitive *se* constructions where all the arguments are present even with the employment of *se* (e.g. ***Se*** *comió toda la pizza* 'S/he ate up all the pizza').

Based on these problems, Maldonado (2000:154) claimed that not all instances of the clitic *se* are reflexive. He argues that the general tendency to analyze any use of *se* as a reflexive marker has impeded the observation of the multifunctionality and polysemy of this clitic. He argues that non-reflexive uses of *se* should not be considered 'exceptional cases' and should not be considered a hindrance to linguistic generalization. His principal thesis is that non-reflexive *se* functions as a middle marker in middle voice constructions and that the middle voice is a unified phenomenon whose main function is to focus on the central properties of events. In other words, Spanish middle voice marked by the clitic *se* pays particular attention to the event expressed in the construction.

Calude (2017) further develops this conceptualization and emphasizes that the middle voice functions in opposition to both active and passive voices. While active voice evaluates the event from the agent's point of view based on the interests of the agent, passive voice focuses on the viewpoint of the patient. The middle voice, however, affords relatively little attention to the participants and in contrast to the active and passive constructions, has a perspective centered on the event itself.

Specifically, Maldonado emphasizes the difference between *se*-marked dynamic contexts and non-*se*-marked absolute contexts. Dynamic contexts express an abrupt or unexpected change and require the use of *se* (e.g. *Juan* ***se*** *cayó* 'Juan fell'), whereas absolute contexts profile a natural or ongoing action or state and do not require the use of *se* (e.g. *Las hojas* ****se*** *caen de los árboles* 'The leaves fall from the trees'). It was shown that children who acquire Spanish as their

first language do not confuse dynamic *se* constructions with their absolute non-*se*-marked counterparts (Maldonado 1993).

Although characterizing the middle voice has been a critical issue in cognitive linguistics, its linguistic implications in L2 acquisition have yet to be examined in detail. Although some studies in L1 acquisition reported results supporting Maldonado's concept of middle domain, there is a relatively limited number of analyses in L2 acquisition which accounts for the cognitive framework proposed by Maldonado. Although the explanatory power of the cognitive approach towards the middle voice is supported by L1 acquisition research, there is still a lack of L2 studies to determine its applicability. To fill this gap, I adopt the cognitive approach of the Spanish middle domain with the following preliminary objectives: (i) to present how the Spanish middle marker *se* is acquired, interpreted and produced by English- and Korean-speaking L2 learners, whose native languages have a different conceptualization of the middle voice, (ii) to analyze the possible effect of L1 (English or Korean, depending on the learner group) or L2 (English, for most L1 Korean speakers) transfer in the acquisition process, and (iii) to make detailed in-group as well as inter-group comparisons in order to confirm if theoretical assumptions made by cognitive linguistics can be attested by empirical SLA research.

For the present study, an online linguistic survey in which English- and Korean-speaking learners of Spanish and native speakers of Spanish participated was conducted. Their performance in production and interpretation tasks was analyzed in terms of how they understand and use the Spanish middle marker *se*. The main research questions that guided the survey are as follows:

RQ 1: Are there differences between L1 English and K1 Korean speakers' knowledge and acquisition of *se*-marking in Spanish and how does this compare to native Spanish speakers' knowledge and use?

RQ 2: What are the differences among the different proficiency levels in terms of their knowledge and acquisition of *se*-marking and how does this compare to native Spanish speakers' knowledge and use?

RQ 3: What are the linguistic and non-linguistic factors that favor *se*-marking among L2 learners and native Spanish speakers?

RQ 4: Do learners and native speakers associate absolute verbal events with non-*se*-marking and dynamic events with *se*-marking?

RQ 5: Do learners and native speakers understand the difference between true *se*-marked reflexives as evidenced by the acceptance of the addition of the phrase *a sí mismo*, and *se*-marked middles as evidenced by the rejection of the phrase?

In the survey, 93 participants (33 L1 English speakers, 30 L1 Korean speakers, and 30 L1 Spanish speakers) were recruited to complete three different tasks: a picture book written narration task (Task 1), an Acceptability Judgment Task (Task 2), and a cloze task (Task 3). Task 1 aimed to examine which (extra-)linguistic factors have a significant effect on the use of *se* in native and non-native written production data. I investigate the frequencies of and constraints on verbal *se*-marking with special interest in the possible similarities and differences between the three participant groups of the current study. Task 2 examines the interpretation or comprehension of Spanish middle constructions by the participants. Following Maldonado (1993, 2008), the task was specifically designed to favor the selection of bare or *se*-marked verbal forms, depending on the nature of the contexts. It is of interest to examine to what extent

native and non-native performance on the interpretation of the middle marker *se* aligns with the expectations set up by cognitive linguistics. Finally, Task 3 investigates whether L2 learners are aware of the association between the reflexive marker *se* and its extended form (i.e. *a sí mismo* 'to oneself') and lack thereof with the middle marker *se*. Since this is a construction that is specifically taught in the foreign language classroom, it is assumed that L2 learners more often accept *a sí mismo* in non-reflexive or middle contexts to such a degree that they are even unaware of its unacceptability. In the analysis, I focus on the comparison and contrast across the three L1 groups and also across L2 learners' different proficiency levels of the target language.

The statistical analysis conducted on the survey data sheds light on the significance of the results. This study also tests the applicability of the cognitive linguistics approach to explaining in part the difficulties English- and Korean-speaking L2 learners have acquiring Spanish middle marker *se*, a topic which needs more established linguistic investigation, as mentioned above. In addition, this dissertation contributes to the broader discussion in the field of SLA by providing empirical data on an understudied L2 speaker group, Korean-speaking L2 learners of Spanish.

1.2 Outline of the dissertation

This dissertation consists of six chapters in total. The current chapter has presented a brief introduction to the problem of traditional approaches on the Spanish clitic *se* and the need for its reconceptualization as a middle marker. Also, I have described the goals for the study with a demonstration of the significance of my research objectives in the field where there is a lack of L2 acquisition studies. An overview of the design for the three tasks was briefly outlined. The rest of the dissertation is as follows:

Chapter 2 includes the first part of the literature review, in which previous studies on Spanish clitic *se*, especially as a middle marker, are discussed. The chapter provides a comparative analysis on the middle system in the three languages in question: English, Korean, and Spanish. Chapter 3 comprises the second part of the literature review, with the illustrations of how the acquisition of *se* has been studied in the field and how Spanish textbooks used in the United States and South Korea present the clitic *se*. This description is necessary since most of the L2 participants in the current study have learned Spanish in classroom settings.

Chapter 4 outlines the methodology of the study, including the research questions and hypotheses, information on participants, the procedure for data collection, and the statistical methods. The results of the study, including the statistical analyses, are presented and discussed in detail in Chapter 5. The final chapter concludes this dissertation with the summary and general conclusions about the research by revisiting the research questions. It also provides some remarks on contributions of the dissertation and directions for future research.

CHAPTER 2

SPANISH MIDDLE MARKER *se*

2.1 Introduction

In this chapter, I provide a review of the relevant literature regarding the Spanish clitic *se* as a middle marker. I first discuss traditional grammatical views on clitic *se* in general, including a summary of some previous studies which tried to define the grammatical category of *se* and other earlier descriptions which attempted to unify most, if not all, of its uses under one general function. I include studies of *se* from a generative perspective. Then, I address issues related to middle voice and how the Spanish clitic *se* has been studied as a middle marker from the perspective of cognitive linguistics. This discussion will include an overview of the ample range of Spanish middle domain and an assumption of middle voice as a basic system for some languages including Spanish. I also review middle voice constructions in English and Korean paying special attention to comparing the linguistic structures in the three languages in question. This comparison is necessary in order to predict how the similarities and differences between them affect the L2 acquisitional process of the Spanish middle marker *se* by both English and Korean speakers.

2.2 Traditional grammarians' explanations of the clitic *se*

In this section, I will examine some of the leading Spanish grammarians' descriptions of the clitic *se*. These include: the Spanish Royal Academy (*Real Academia Española*, henceforth RAE) (2009, 2010), Bello (1891), Gili y Gaya (1955), and Seco (1972).

The RAE defines a clitic as an unstressed word, especially a personal pronoun that is integrated into the accentual group of the tonic word that precedes or follows it. In the category of personal pronouns, in general, direct object pronouns, indirect object pronouns, and reflexive pronouns in Spanish are classified as clitic pronouns. Table 1 below presents the aforementioned three clitic pronoun categories with their corresponding clitic forms:

Table 1. Spanish clitic pronouns

Direct Object Pronouns	*me*	*nos*
	te	*os*
	lo/la	*los/las*
Indirect Object Pronouns	*me*	*nos*
	te	*os*
	le	*les*
Reflexive Pronouns	*me*	*nos*
	te	*os*
	se	***se***

As we can see in the table above, the clitic *se* forms part of the third category, that is the reflexive pronouns. In earlier literature, the definition of the clitic *se* appears to be limited to this reflexive category. Bello (1891), for example, provided a list of examples with clitic *se* and defined them as reflexive constructions:

(1) a. *El niño o la niña **se** levanta.* ‘The boy or the girl gets up.’
b. *Aquello **se** precipita a su ruina.* ‘That rushes to its ruin.’
c. *Él o ella **se** pone la capa.* ‘He or she puts on the cape.’
d. *Aquello **se** atraía la atención de todos.* ‘That attracted everyone’s attention.’
(Bello 1891:84, [translation mine])

However, some scholars have found problems in defining the clitic *se* as merely a reflexive pronoun due to its various and complex uses as well as its multifunctionality. While admitting that the clitic *se*’s original function derives from its reflexive value, Gili y Gaya (1955) additionally mentions its reciprocal (e.g. *El niño y la niña **se** pelearon* ‘The boy and the girl got onto a fight’ (1955:68)), impersonal (e.g. ***se** dice* ‘one says’, ***se** supone* ‘one supposes’ (1955:71)), and passive (e.g. *La pared **se** hundió con el peso de la techumbre* ‘The wall sank with the weight of the roof’ (1955:114)) uses.

In addition, Gili y Gaya also points out that there are other instances of *se* such as ethical datives or datives of interest and pseudo-reflexives, where the reflexivity is attenuated so that the pronouns are no longer direct or indirect complements of the sentence but vaguely indicate their participation or interest in the action:

(2) a. *Ella **se** tomó el café.* ‘She drank up the coffee.’ (Ethical dative)
b. ***Te** estás en casa.* ‘You are at home.’ (Pseudo-reflexive)
(Gili y Gaya 1955:68, [translation mine])

Gili y Gaya states that these expressions are commonly and widely used by Spanish speakers, however, some of them were stigmatized for being “excessively vulgar” (1955:68).

Similarly, Seco (1972:102-106) provides examples of the clitic *se* constructions in Spanish in a descriptive manner. The usages which were mentioned by both Gili y Gaya and Seco include impersonal, passive, and ethical dative or ‘expressive use’, the latter a term used by Seco. Some additional functions that only Seco (1972:104) explains consist of inchoative (e.g. *El*

niño ***se*** *ha dormido* 'The child fell asleep (in sense of the child started to sleep)', obligatorily pronominal (e.g. *suicidar****se*** 'to commit suicide', *arrepentir****se*** 'to regret'), and spontaneous (e.g. *El puente* ***se*** *ha hundido* 'The bridge has sunk') constructions.

In sum, traditional grammarians argued that Spanish clitic *se* is derived from reflexive pronouns and its usage expanded from reflexive/reciprocal to various others, including impersonal and passive as standard expressions along with some common but vulgar constructions, such as ethical datives and pseudo-reflexives. More recently, the RAE (2010) provides an extensive account of Spanish *se* constructions, also from a descriptive grammar perspective:

Table 2. Types of Spanish clitic *se* (RAE 2010:782-783, [translation mine])

Paradigmatic *se*	Reflexive	***Se*** *cuida poco.* 'S/he takes little care of her/himself.'
	Reciprocal	***Se*** *adoran.* 'They love each other.'
	Part of pronominal verb	*Yo* ***me*** *canso.* 'I get tired.'
	Non-argumental dative	***Te*** *lo comiste.* 'You ate it up.'
Nonparadigmatic *se*	Passive reflexive	*Las noticias* ***se*** *recibieron ayer.* 'The news was received yesterday.'
	Impersonal reflexive	***Se*** *vive bien siendo estudiante.* 'One lives well as a student.'

As can be seen above, the most distinctive explanation stated in the RAE is that it points out that *se* can be used as a non-argumental dative. That is, while *se* in reflexive constructions is an argumental element given that this *se* refers to an argument which is either direct or indirect object (e.g. *gente que* ***se*** *adora* ***a sí misma*** 'people who love themselves' (RAE 2009:3081),

where *se-a sí misma* is the direct object of the verb *adorar*), the non-argumental dative *se* (e.g. ***Te*** *lo comiste.* 'You ate it **up**.') lacks coreferentiality with any arguments. However, the RAE still aligns with traditional grammarians, considering that it also places impersonal and passive *se* under the category of reflexive.

2.3 Generative explanations of the clitic *se*

From the generative point of view, Kempchinsky (2004) classifies different types of constructions with the clitic *se* as follows:

(3) a. ***Se*** *trabaja duro en el mundo académico.* (Impersonal)
'One works hard in the academic world.'
b. ***Se*** *construyeron (varias) casas allí.* (Passive)
'(Various) houses were constructed there.'
c. *Las camisas de algodón* ***se*** *lavan fácilmente.* (Middle)
'Cotton shirts wash easily.'
d. *Las ventanas* ***se*** *rompieron durante la tormenta.* (Ergative/Inchoative)
'The windows broke during the storm.'
e. *La niña* ***se*** *miraba (a sí misma) en el espejo.* (Reflexive)
'The girl observed herself in the mirror.'
f. *Juan* ***se*** *afeita en la barbería (para impresionar a sus amigos).* (Causative)
'Juan gets shaved at the barbershop (in order to impress his friends).'
g. *Los estudiantes siempre* ***se*** *quejan de las clases.* (Inherent/Antipassive)
'The students always complain about their classes.'
h. *El niño* ***se*** *le comió toda la lecha a su hermano.* (Aspectual)
'The little boy drank up all the milk on his brother.' (Kempchinsky 2004:241)

Like the RAE (2010), Kempchinsky also divides these constructions into two groups: nonparadigmatic and paradigmatic. She explains that nonparadigmatic *se* is only used in third-person verbal forms while paradigmatic *se* matches the gender and number of the subject (e.g. *Yo*

me *quejo* 'I complain', *Tú* ***te*** *quejas* 'You complain'). Among all the uses listed above, impersonal, passive, and middle constructions are grouped as nonparadigmatic and the others as paradigmatic. Although it is not explicitly discussed in the study, it is plausible to infer that Kempchinsky's classification resembles that of descriptive grammar, considering that the labels she uses are common terms used in the literature.

The particularity of reflexive *se* compared to other uses which Kempchinsky highlights is its compatibility with *a sí mismo*. This is a distinctive characteristic of Spanish given that in other Romance languages such as French and Italian, the existence of *se* blocks the use of *a sí mismo*. Also, Kempchinsky provides a cartographic explanation, introducing several functional categories such as AspP, EventP, and VoiceP. For example, she suggests that *se* in ergative/inchoative constructions merges as head of Asp, while the reflexive *se* is located in the Spec of Asp.

Although this analysis might successfully describe the syntactic structure of each of the uses of *se*, its applicability in understanding the process of their acquisition by L2 learners seems questionable. Indeed, proposing and analyzing *se* with various functional categories help to examine whether L2 learners have post-critical period access to functional syntax not instantiated in the L1. However, it is important to consider that the use of *se* reflects speakers' cognitive perception on how events are construed and expressed, which is not easily accounted for in syntactic analyses. In other words, an approach based on syntactic analysis appears to be less convincing given that examining L2 learners' conceptualization of *se* needs to go beyond testing whether they have grammatical knowledge on target language specific functional categories.

On the other hand, Kempchinsky admits that the terms (i.e. Impersonal, Passive, Middle, Ergative/Inchoative, Reflexive, Causative, Inherent/Antipassive, Aspectual) used in her research

are “not necessarily contentful in and of themselves” (2004:241) because there still is disagreement in terminology and classification in the literature. The differences in determining and classifying the uses of *se* are reflected in various analyses which have focused on *se*’s syntactic and semantic characterizations. For example, in analyzing the same type of constructions termed ‘aspectual’ by Kempchinsky, Basilico (2010) and MacDonald (2017) provide disparate accounts:

(4) a. *Juan **se** leyó un libro.* ‘Juan read a book.’ (Basilico 2010:275)
b. *María **se** comió el helado.* ‘María ate up the ice cream.’ (MacDonald 2017:74)

While Basilico explains (4a) from the syntax-lexicon interface framework using the term ‘transitive *se*’, MacDonald focuses more on cartographic analysis, following Kempchinsky and dubbed the same structure in (4b) ‘aspectual *se*’.

According to Basilico, the clitic *se* in transitive constructions is “a head of an underspecified eventive light verb[2]” (2010:274). This light verb takes a bounded path or scale which entails a certain end point of the event as a complement, hence the majority of transitive *se* constructions appears to carry Accomplishment[3] type verbs. Meanwhile, MacDonald defines *se*

[2] Light verb, coined by Jespersen (1965), is a type of verb that is used in combination with other verbs or nouns due to its lack of semantic properties. For example, the verb *take* in a *take a walk* is a light verb because the actual semantic meaning of this verbal phrase is conveyed by the noun *walk*, not by the verb *take*. However, in an example such as *take medicines*, the verb *take* is not a light verb because it does convey its semantic meaning.

[3] Accomplishment is one of the four lexical aspects, also known as *Aktionsart*, classified by Vendler (1967). Vendler divided the mode of action of the predicates into four types (based on the inherent semantic notions of dynamicity, telicity and punctuality): State, Activity, Accomplishment and Achievement. The semantic characteristics of the verbs of Accomplishment are [+dynamic], [+telic], and [-punctual].

as an indirect object reflexive verb which heads Voice, arguing that aspectual *se* constructions have "the underlying configuration of a double object construction" (2017:75). Thus, his analysis is in compliance with some traditional views that considered aspectual *se* another type of reflexive. Even with this brief overview of two studies analyzing one of the many Spanish clitic *se* constructions, it is shown that there has been lack of consent among researchers on the description on the nature of the clitic *se* constructions in Spanish.

In particular, researchers who take a generativist approach have conducted their studies with a special interest in the clitic *se*'s grammatical nature in pursuit of a unified analysis in which all of the uses of *se* can be explained by a single explanation. For instance, some generativists described the grammatical phenomena of *se* as a "complex problem" (King & Suñer 2017:194) or "a matter of much controversy" (Montrul 2004:310). While Montrul's (2004) description does not differ a great deal from that of Kempchinsky (2004, 2006), given that it classified different types of *se* constructions (i.e. reflexive, reciprocal, antipassive, impersonal, passive, middle and aspectual) and provides examples of each of them, King & Suñer (2017) provided a more detailed description of various *se* constructions with examples for L2 learners of Spanish, starting from the reflexive and reciprocal function of *se*. Following the introduction of reflexive/reciprocal *se*, they compare obligatory and optional uses of *se* with intransitive verbs. The authors explain the use of *se* with intransitive verbs as "idiosyncratic" (2017:198) because the clitic *se* does not add much semantic meaning to the verb. Their remark on the acquisition process of this construction type and its variability among native speakers is worth noticing:

> The learner of Spanish as a second language has to learn the most common obligatory reflexives, just as the Spanish speaker will automatically learn them as they are heard in context. Also, not all Spanish speakers have the exact same required lexical verbs as part of their vocabulary. (2017:198, [translation mine])

For optional use of *se* in intransitive constructions, King & Suñer explain that *se* marking does not change the meaning of the sentence at all (e.g. *reír(se)* 'to laugh', *quedar(se)* 'to stay', *morir(se)* 'to die'), or the difference in meaning brought by the use of *se* is not significant (e.g. *dormir* 'to sleep' vs. *dormirse* 'to fall asleep', *ir* 'to go' vs. *irse* 'to leave'). Even though the authors provide a relatively exhaustive explanation covering various constructions (i.e. spontaneous events, unexpected events, spurious use of *se*, impersonal and passive), neglecting to analyze the semantic differences that occur with the use of *se* and dismissing them as minimal or insignificant does not seem to benefit readers' understanding of the "complex problem" (2017:194).

Meanwhile, other studies demonstrated a more theoretical analysis of the grammatical nature of the clitic *se*. For example, Cinque (1988) and Raposo & Uriagereka (1996) viewed *se* as a pronoun in the argument position, while Folli & Harley (2005), Basilico (2010), Ordóñez & Treviño (2011), and Armstrong (2013) analyzed it as a predicate or verbal morpheme. On the other hand, Torrego (1995) and Kempchinsky (2004, 2006) claim that the morpho-syntactic state of *se* is specified depending on the specific construction in which it appears, thus suggesting that *se* can function as a pronoun in certain contexts and be realized as part of the predicate in others.

Cinque (1988), for example, whose work is based on Italian impersonal constructions, assumed that *se* is a pronominal clitic in the position of [NP, IP] and that there are two types of *se*: argument *vs.* non-argument *se*. Some *se* constructions can be classified into certain categories

according to this difference in its argumental status (cf. Table 3); for the purpose of the present study, only reflexive constructions will be discussed here.

Table 3. Classification of *se* constructions based on [±argument] feature (Adapted from Cinque 1988:575)

Impersonal-arbitrary *se*	[+argument]: absorbs external theta role and nominative Case	Impersonal *se* (in transitive contexts, passive *se*)
	[-argument]: identifies arbitrary *pro* in conjunction with personal Agreement	Impersonal *se*
"Passivizer" *se*	[-argument]: suspends external theta role and accusative Case	Middle *se*
Reflexive *se*	[+argument]: absorbs external theta role and VP-internal Case	(True) Reflexive *se*
	[-argument]: suspends external theta role and VP-internal Case	Ergative/Inchoative *se*
	[-argument]: marks the absence of external theta role and VP-internal Case	Inherent *se*

Reflexive *se* can be divided into [+arg] and [-arg], as shown in the table above. Cinque explains that genuine reflexive[4] and reciprocal *se* is [+arg], whereas ergative/inchoative and inherent *se* is [-arg]. Below are examples[5] for each reflexive type:

(5) a. *Ni siquiera **se** comprenden **a sí mismos**.* (1988:538) ((True) Reflexive)
'They do not even understand themselves.'

[4] In the generative tradition, a predicate is defined as reflexive "iff (at least) two of its arguments are coindexed" (Reinhart & Reuland 1993:662).

[5] Note that these are Cinque's (1988:538, 565) Italian examples translated to Spanish.

b. *El coche* ***se*** *rompió.* (1988:565) (Ergative/Inchoative)

'The car broke.'

c. *Giovanni* ***se*** *enfermó.* (1988:565) (Inherent)

'Giovanni got ill.'

The difference in the argumental status derives from the fact that [+arg] clitics are formed in the syntactic level while [-arg] clitics are formed by a lexical process. Based on this criterion, Cinque (1988:521) argues that "this analysis is in principle compatible with the program of unifying all uses of *si*[6] (impersonal, middle, and reflexive)".

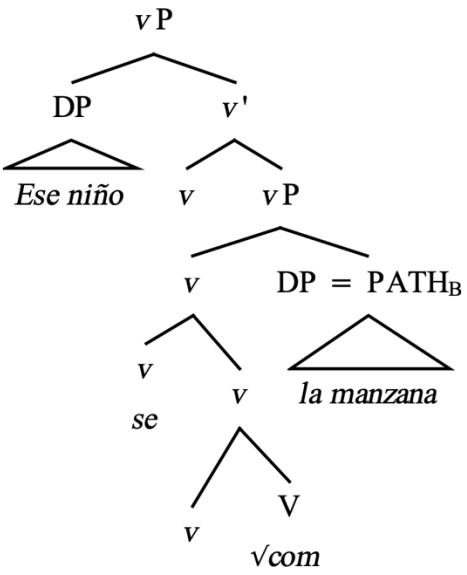
According to Folli & Harley (2005:95), on the other hand, the morpheme *se* is not a pronominal clitic but a realization of a light verb. Based on Italian verbs of consumption which use *se* with inanimate subjects, these authors argue that *se* realizes one of the different flavors of light verbs: v_{DO}, v_{CAUSE} and v_{BECOME}. For example, the Italian clitic *si* in e.g. *Il mare si é mangiato la spiaggia* (lit. 'The sea ate the beach.') (2005:112) is a realization of the light verb v_{CAUSE} which selects a complement that implies a final state of the event.

Another analysis that takes a similar approach is Basilico (2010) in which Spanish *se* constructions are examined. He assumes that the clitic *se* is a head of the projection of (a type of) a light verb. In particular, Basilico analyzes *se* as an unspecified eventive verbal head which takes a bounded path as its complement. Based on this analysis, a transitive construction with the clitic *se* in Spanish can be represented in a syntactic tree as follows:

[6] Italian equivalent of Spanish clitic *se*.

(6) a. *Ese niño* ***se*** *comió la manzana.* 'That child ate the apple.'

b.



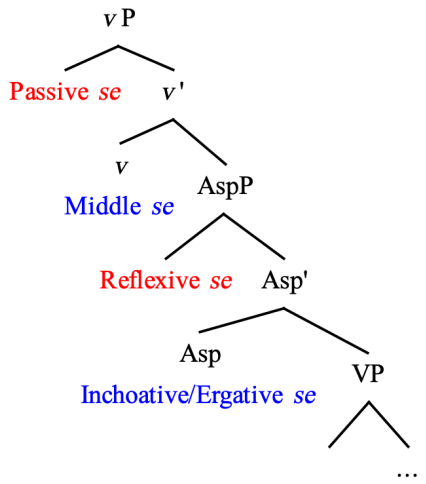
(Adapted from Basilico 2010:285)

In the above structure, the verbal root √*com* '√eat' is combined with a light verb, which later combines with *se* to create a complex verbal head. This complex verbal head requires a bounded path as a complement, *la manzana* 'the apple'. According to Basilico, the verb specifies the event introduced by *se*, specifically the manner of the activity, which is the manner of 'eating'. Here, it can be understood that "once the object has been used up and its volume or material extent is zero, the event is complete" (2010:282). As is shown, both Folli & Harley (2005) and Basilico (2010) align with the assumption that clitic *se* is related to verbal rather than pronominal properties.

Although it is not explicitly explained by the author, it is possible to consider that Kempchinsky's (2004, 2006) analyses combine the aforementioned two approaches: *se* as a pronominal argument *vs. se* as part of a predicate. Kempchinsky argues that the clitic *se* is an

aspectual element and can function either as an argument or as part of a predicate, depending on its location in the syntactic structure. In other words, as is shown in (7) below, if it is located in the position of specifier, it functions as an argument and results in passive *se* or reflexive *se* constructions. On the other hand, in middle and ergative/inchoative constructions, *se* appears as the head of an aspectual projection and has a role of a predicate.

(7) *se* in the syntactic structure (Adapted from Kempchinsky 2006:533)

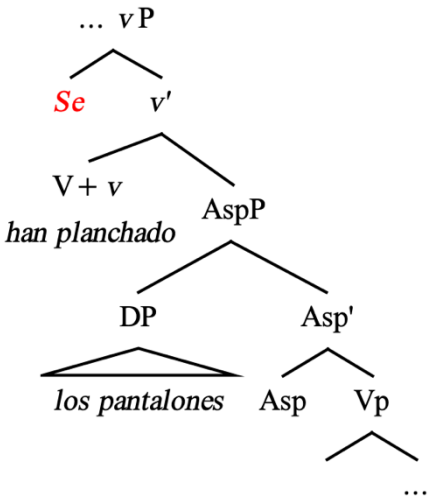


To contrast and analyze these passive, middle, reflexive, and inchoative/ergative *se* constructions in detail, examples with their corresponding syntactic trees[7] are provided as follows:

[7] These syntactic trees (8b-12b) are partial derivations and I do not assume that these are base-generated positions.

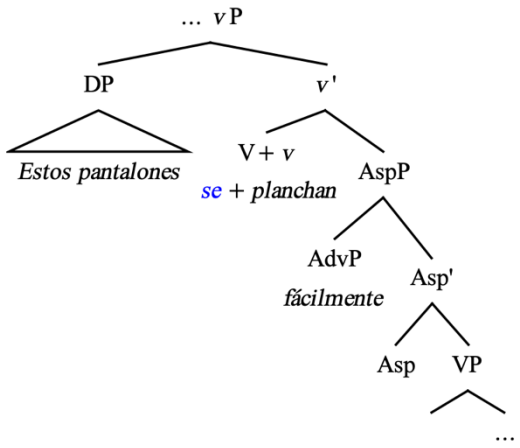
(8) a. ***Se*** *han planchado los pantalones.* 'The pants have been ironed.' (Passive)

b.



(9) a. *Estos pantalones* ***se*** *planchan fácilmente.* 'These pants iron easily.' (Middle)

b.



The main difference between the passive construction in (8b) and the middle construction in (9b) is what is located in [Spec, *v*]. In the former, this is where the clitic *se* is merged, while in the

latter, the subject DP is moved here from its verb internal position. Kempchinsky (2006:546) argues that it is the position where *se* merges in the derivation in which the semantic and syntactic distinction between these two constructions arise. The clitic *se* in middles merges as the head of *v*, suspending an agentive external argument. However, in passives, the clitic *se* itself is an external argument, implying that there is a certain degree of agentivity. This difference is demonstrated by the incompatibility with purpose clauses and agent-oriented adverbs in middle constructions:

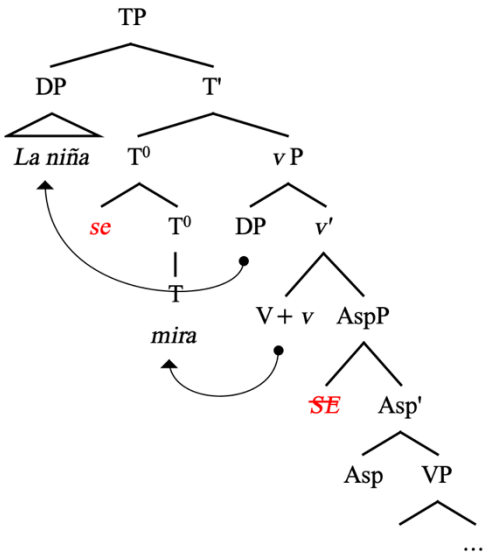
(10) a. *Este libro **se** lee fácilmente (*para ayudar a los estudiantes).* (Middle)
'This book reads easily (*in order to help the students).'
b. *Esta raíz **se** come para adelgazar.*[8] (Passive)
'This root is eaten in order to lose weight.'
(Kempchinsky 2006:536, 542, respectively)

Meanwhile, the reflexive construction undergoes cliticization, as shown below in (11). Once merged as an argument in [Spec, Asp], *se* adjoins to T^0 for cliticization. Kempchinsky explains that this process of cliticization allows the clitic *se* "to inherit specific person and number features from its antecedent, at the end of the derivation, when *se* is cliticized to T and the subject DP is in [Spec, T]" (2004:246).

[8] This example is translated from Kempchinsky's French example: *Cette racine se mange pour maigrir.* (2006:542)

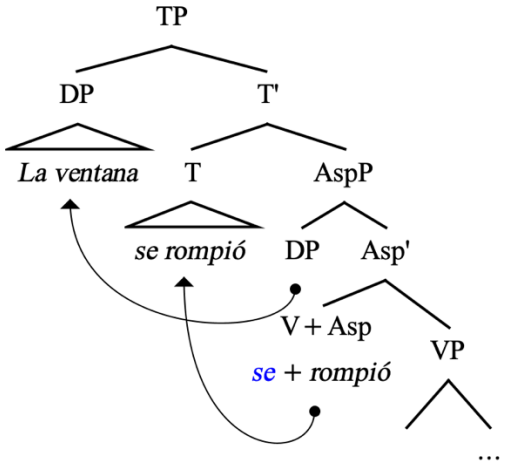
(11) a. *La niña* ***se*** *mira.* 'The girl looks at herself.' (Reflexive)

b.



(12) a. *La ventana* ***se*** *rompió.* 'The window broke.' (Inchoative/Ergative)

b.



Finally, the distinctive characteristic of inchoative/ergative *se* in (12) above is that it lacks the *v*P layer. This absence of *v*P means that there is no place for the external argument, thus neither purpose clauses nor agent-related adverbs are allowed. Though it may seem similar to the middle *se* because both lack an agentive external argument, middle *se* merges as head of *v* while in inchoative/ergative structure, *v* does not exist at all. Also, Kempchinsky proposes that in inchoative/ergative constructions, "what happens is that *se*, as a predicate, introduces the temporal role (subevent) CHANGE OF STATE" (2004:250), which describes their particular semantic property.

In summary, while there is no unanimously accepted analysis, one thing that both traditional grammarians and generativists share is their pursuit of a unified approach based on the grammatical nature of clitic *se*, such as the position where it is generated and the type of complement it selects. Although they succeeded in coming to a conclusion with a reasonable system which explains the underlying grammatical processes of the use of *se*, it still seems problematic to name some disparate constructions as genuine and non-genuine usages of 'reflexive' *se*. In other words, we still lack a unifying approach to Spanish *se*. Furthermore, the debate regarding using the term 'reflexive' as a category that contains all these constructions is derived from the fact that the detransitivization analysis based on the traditional view of the transitive system is misleading. In the following section, I discuss some of the arguments set forth by different authors that reject the detransitivization approach and propose an alternative explanation for the polysemous nature of Spanish *se*.

Finally, there remain questions related to the acquisition of *se*. Specifically, how do we explain that a child comes to dominate this complex and (partially) disparate system of clitic *se* in her/his native language at an early age (usually by 2 to 3 years), when L2 adult learners

struggle to comprehend and productively use *se* after years of study and experience with Spanish? I explore some of these issues in Chapter 3.

2.4 A cognitive perspective of the clitic *se*

2.4.1 Rejecting the *detransitivization* analysis

Maldonado (1992:17) observes that, while the arguments made in generative approaches sound reasonable, both pronominal and non-pronominal analyses need to be revised. His main argument is that they are falsely based on a sharp distinction between transitive and intransitive constructions. Maldonado rejects the detransitivization approach because under a detransitivization analysis, constructions with *se* have no other option than being intransitive, which fails to explain numerous transitive *se* constructions. In other words, "the intransitive is much too wide a category where no distinction is established between a wide variety of not always closely related constructions" (1992:16).

On the other hand, the aforementioned alternative analyses which assume that the clitic *se* is a predicate and needs a certain complement to function as a (light) verbal head are also based on the traditional dichotomous concept of transitivity, and they fail to capture the concept of transitivity increase or decrease which will be explained later in 2.4.2. In other words, in generative literature, where it is commonly accepted that a predicate can only be either transitive or intransitive, a conceptualization that some predicates can be more or less transitive than others might not be well accounted for. However, as will be reviewed with more detail in 2.4.2, Maldonado points out that this gradualness of transitivity deserves more attention than has been formerly given to explore the system behind the various clitic *se* constructions.

Furthermore, under the seemingly challenging objective of discovering and establishing a unified analysis which covers the linguistic phenomena with all of the clitic *se* constructions, Maldonado (1992) states that the clitic *se* has traditionally been analyzed based on two assumptions: the coreferentiality hypothesis and the argument-reduction hypothesis. The coreferentiality hypothesis comes from a tradition that generalizes any instances of *se* as coreferential reflexive constructions. This approach is not applicable in numerous cases, such as the example below:

(13) *Juan Carlos y Pilar* ***se*** *tomaron un tequila.* 'Juan Carlos and Pilar drank a tequila.'
(Maldonado 1992:11)

In (13), the subject and the object of the sentence are not coreferential, i.e. Juan Carlos and Pilar are not drinking themselves.

Later, the argument-reduction hypothesis emerged as an alternative to the coreferentiality hypothesis, according to Maldonado. This new approach claims that the main function of the clitic *se* is to mark the valency reduction of the structure. This way, more non-coreferential constructions can be explained without getting into the problem of the impossibility of describing instances in which the arguments are not coreferential. However, the argument-reduction hypothesis still fails to explain (13), based on the fact that there isn't any valency reduction in the construction. Here, the subject *Juan Carlos* and *Pilar* and the direct object *un tequila* 'a tequila' are both present, occupying the two necessary argument positions. It is expected that the example (13) results in a one-argument construction or an intransitive sentence, i.e. one of the two arguments being absent, if the clitic *se* reduced any valency of the structure.

As such, between these two approaches, Maldonado (1992:13) affirms that the argument-reduction hypothesis has more explanatory power than the former. However, they do not differ remarkably from each other given that both analyses define the main function of the clitic *se* as a

reflexive marker which shows a detransitivization effect. Based on this, I will use the term 'detransitivization analysis/approach' to refer to the coreferentiality and argument-reduction hypotheses together.

Detransitivization is a linguistic process in which a transitive construction loses its argument, resulting in an intransitive construction (cf. Levin & Rappaport Hovav 1995). In other words, when the clitic *se* replaces one of the arguments of a transitive verb, it becomes detransitivized and approximates an intransitive one. By losing one argument, the verb's semantic and syntactic valence is reduced. This valence-reducing process has been considered to be a detransitivizing effect of the clitic *se*, especially in pre-Minimalist approaches.

Under the traditional analyses, most of the clitic *se* constructions were regarded as having "the same grammar" (Grimshaw 1982, cited in Maldonado 1992:12). Reflexive and ergative/inchoative constructions show a good example of this approach. According to the detransitivization analysis, both constructions are marked by the clitic *se* and have only one explicit argument at the surface structure —the subject of the sentence— like intransitive constructions because *se* replaced one of its arguments, causing it to be 'invisible'. Consider the following examples:

(14) a. *Juan miró a María.* 'Juan looked at María.'
b. *Juan* ***se*** *miró (a sí mismo*). 'Juan looked at himself.'
(15) a. *Juan rompió la copa.* 'Juan broke the glass.'
b. *La copa* ***se*** *rompió.* 'The glass broke.'

The clitic *se* in (14b) is a reflexive marker in the sense that it shows that the subject and object of the sentence are coreferential. Here, the clitic *se* absorbs the direct object, *Juan* himself, and makes the subject seemingly the only argument in the structure, thus it 'detransitivized' the sentence. This reflexive construction can be compared to the transitive construction in (14a)

where two separate arguments, the subject *Juan* and direct object *María,* are found. According to the traditional view, the clitic *se* in (15b) also plays the same role as a reflexive marker because it deletes the subject *Juan* from the transitive construction (15a) and makes the object, *la copa* 'the glass', become the grammatical subject. Essentially, the detransitivization approach claims that the main function of the clitic *se* is to reduce a transitive construction to an intransitive one.

Although this analysis seems plausible at first glance, Maldonado argues that it is "insufficient" (1992:13) because there are several clitic *se* constructions that the detransitivization approach fails to explain. First, a reflexive, by definition, is transitive. In spite of the fact that the direct object in (14b) is not explicit in the surface structure, it still does exist and is a two-argument construction with a subject and an object. Thus, for true reflexives where *a sí mismo* is compatible, the reflexive marker *se* does not detransitivize the structure nor reduce the valency. Also, as mentioned earlier, there are some examples in which the clitic *se* is used in constructions where both subject and object are explicitly existent (Example (13) repeated below for convenience):

(16) *Juan Carlos y Pilar* ***se*** *tomaron un tequila.* 'Juan Carlos and Pilar drank a tequila.'
(Maldonado 1992:11)

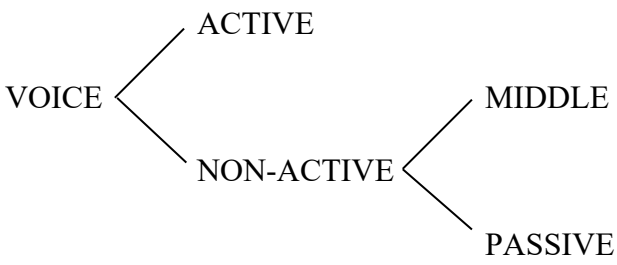
For *se* to have a detransitivizing function in this sentence, the verb *tomar* 'to drink' has to be a predicate which requires three arguments. However, it only requires two arguments in this context, and both are present in the surface structure, which makes it impossible for the detransitivization analysis to describe.

2.4.2 An alternative analysis: Spanish clitic *se* as a middle marker

In general, most theoretical syntactic studies show a dichotomous analysis when it comes to voice: active voice and non-active voice, the latter of which is frequently dubbed as passive

voice. Instead of being classified as a sub-category of voice, middle has been used to denote a restricted type of construction, which is called dispositional middles, e.g. *The bread cuts easily* (Alexiadou & Doron 2012:1). Similarly, in the case of Spanish, dispositional middle sentences (e.g. *Este coche* ***se*** *conduce con facilidad* 'This car drives easily' (Kempchinsky 2006:536)) were traditionally considered middle voice constructions.

According to Kemmer (1994:179), the term *middle voice* was originally used by grammarians to designate a category of verbal inflection of classical Greek. She mentions that the middle voice has been viewed as "being situated halfway between the active and the passive voice" (1993, cited in Balcom 2003:169). According to Calude (2017:599), although it has received extensive attention in the literature, the precise boundary between the middle voice and other related constructions, especially passive structures, remains difficult to pinpoint. However, passive and middle are indeed different voices. The passive is only derived when the corresponding active structure is derived, while the middle is derived independently of the corresponding active structure (Doron 2003:55). As such, it is more common in typological studies to recognize passive and middle voice as two separate components of non-active voice:



Voice changes the grammatical function of arguments (Doron 2003:63). For example, when there's a change from active to passive voice in the structure, the theme argument in the active construction alters its grammatical function from object to subject in the passive sentence. Consequently, it is closely related to a way of approaching the relationship between participants

in a situation or event (Bartra 2014:3), which results in the verb's argument structure alternations. Some argument structure alternations that were analyzed in various literatures include active-passive alternation, causative-anticausative alternation, dispositional middle alternation and reflexive alternation (Alexiadou 2014:21-22). The active-passive alternation is realized by the analytic form, i.e. *ser* + past participle, in Spanish. For example, an active sentence *Juan leyó el libro* 'Juan read the book' has a passive counterpart, *El libro fue leído por Juan* 'The book was read by Juan'. The other three alternations mentioned above are marked with *se*:

(17) The causative-anticausative alternation:

a. *Juan rompió el jarrón.* 'Juan broke the vase.'

b. *El jarrón* ***se*** *rompió.* 'The vase broke.'

(18) The dispositional middle alternation:

a. *El carnicero corta la carne.* 'The butcher cuts the meat.'

b. *La carne* ***se*** *corta fácilmente.* 'The meat cuts easily.'

(19) The reflexive alternation:

a. *Juan lavó a María.* 'Juan washed María.'

b. *Juan* ***se*** *lavaba todas las mañanas.* 'Juan washed every morning.'

(Alexiadou 2014:20-21, [translation mine])

Alexiadou calls the clitic *se* in above examples (17-19) a 'reflexive' weak pronoun/clitic. It is worth noting that although the clitic *se* is not used to mark reflexivity, but to mark argument structure alternations, it is still defined as a reflexive clitic.

This shows that there has been a long tradition in linguistics to analyze any instances of clitic *se* as a reflexive marker (Maldonado 1992:2). For example, with the objective of describing Spanish dispositional middle structures, Bayona (2005:125) states "middle constructions are a notional category that employs the reflexive marker *se*, but whose syntactic and semantic characteristics differ from the reflexive, inchoative, impersonal, and passive structures that make

use of the reflexive clitic as well". It is not difficult to find similar explanations of non-reflexive uses of the clitic *se* applying the term 'reflexive', which makes it even more confusing.

The miscorrelation of clitic *se* and reflexivity has received attention in the field of cognitive linguistics, especially by Kemmer (1994) and Maldonado (1992; 2000; 2008; 2009, among others). The non-reflexive use of *se* did not receive sufficient attention in previous analyses prior to cognitive linguistic studies. Initial studies were conducted under the generative approach and mainly centered on the grammatical properties of *se*. Cognitive linguists later stressed the importance of the semantic notions of *se* that extend beyond strictly reflexive or detransitivizing uses. From the cognitive perspective, the grammar of a language can be characterized as a structured inventory of conventional linguistic units whose degree of automation depends on how they are rooted in the cognitive organization of a specific language (Maldonado 2000:155-156). Emphasizing the point of view of the speaker/conceptualizer of a given construction is one of the fundamental notions of cognitive linguistics:

> "[…] In choosing a particular expression or construction, a speaker construes the conceived situation in a certain way […] (Langacker 1988:7)". […] The meaning of an expression includes both the knowledge system it evokes when the expression is activated (the various cognitive domains), as well as the particular CONSTRUAL the conceptualizer imposes on a scene. (Maldonado 2000:156)

On the basis of cognitive linguistics, Maldonado highlights that not all instances of the clitic *se* are reflexive; therefore, he comes to a conclusion that referring to *se* as a reflexive pronoun is misleading. Moreover, he argues that the general tendency to analyze any use of *se* as a reflexive marker has impeded the observation in more detail of the multifunctionality and polysemy of this clitic (Maldonado 2000:154) and of the underlying general semantic property that unifies the different meanings and functions. Thus, it is important to acknowledge that not all instances of *se*

are related to reflexivity, and thus, clitic *se* cannot always be considered a reflexive marker. That is, although the relationship between the reflexive and middle markers is evident, since in both types of constructions the clitic *se* refers to the subject of the clause, there are still important differences between them. In order to analyze these differences, Maldonado bases his studies on Langacker's (1991) concept of 'domain' and Kemmer's (1993, 1994) 'degree of distinguishability'.

Firstly, based on Kemmer's work, Maldonado analyzes the contrast between reflexive and middle markers in terms of degree of distinguishability. When the agent and the patient or the goal/recipient of a given sentence, i.e. a direct or indirect object, can be fully distinguished and is coreferential, *se* functions as a reflexive marker:

(20) *De tanto decir**se** que era insoportable, Ceci acabó deprimiéndo**se** a sí misma.*
'After so much telling herself that she was unbearable, Ceci ended up depressing herself.' (Maldonado 2008:162)

Example (20) shows both uses of the reflexive marker *se* as a direct and indirect object. The first *se* in *decirse* 'to tell oneself' is the representation of the coreference between the subject-agent and the indirect object-goal/recipient of the sentence, while the second *se* in *deprimirse* 'to depress oneself' demonstrates the coreference between the subject-agent and the direct object-theme of the given construction. Also, the use of the *a sí mismo* phrase affirms that there is a split representation between the subject and object.

On the other hand, when a divided representation of the agent and patient is not present, the subject is represented more as an experiencer rather than an agent, and *se* functions as a middle marker:

(21) a. ***Me*** *enfermé al salir de la fiesta.* 'I got sick as I left the party.'

b. ****Me*** *enfermé a mí mismo al salir de la fiesta.*

Intended reading: 'I sickened myself as I left the party.' (Maldonado 2008:160)

As can be seen above, the use of the phrase *a sí mismo*, results in ungrammaticality. As *a sí mismo* illustrates a split representation between the subject and the object, its incompatibility with the middle marker *se* reflects the fact that middles are different from reflexives where *se* can be expanded by *a sí mismo* phrase. Also, as Maldonado explains, "middle constructions do not profile the way an agent acts on himself, instead they focus on the change of state undergone by the experiencer" (2008:61). Here, example (21a) describes a situation in which the experiencer 'I' went through a change of state from being fine to getting sick. As such, middle constructions illustrate the ways in which the complexity of the event and the representation of its participants are related to the specific use of *se*.

Kemmer (1994) points out that there is a crucial semantic property for the nature of the middle voice that has not been previously observed, which incorporates the notions of the subject's affectivity. She terms this general property 'elaboration of events', which is "the parameter along which the reflexive and the middle can be situated as semantic categories intermediate in transitivity between one-participant and two-participant events, and which in addition differentiates reflexive and middle from one another" (1994:181). For example, in the case of the middle, since the subject's action cannot be distinguished from the affectivity of the object, the event is simplified and determines a low degree of elaboration of the event. This is closely related to what Maldonado (2008) calls 'split representation'. As we saw in the example (21) above, Spanish middle marker *se* indicates that there is not a split representation and that the subject is more of an experiencer, hence the subject's low degree of elaboration of the event.

Recall that the subject in (21a) is merely an experiencer who undergoes the change of state. Compare this example with a transitive construction in (22) below:

(22) *Corté el pan.* 'I cut the bread.'

This is a two-participant transitive construction with a subject-agent 'I' and an object-theme, *el pan* 'the bread'. Here, the subject-agent's action (i.e. cutting) directly affects the object-theme, hence a high degree of elaboration of the event.

On the other hand, Maldonado (2000:156) defines 'domain' by Langacker as the virtual area in which a participant has mental or physical access to manipulate, control, or have mental contact with the object(s). For example, a prototypical transitive construction is characterized as a chain of action where the subject transfers some energy to the object. Langacker describes the energy transfer as follows:

> We think of our world as being populated by discrete physical objects. These objects are capable of moving about through space and making contact with one another. Motion is driven by energy, which some objects draw from internal resources and others receive from the exterior. When motion results in forceful physical contact, energy is transmitted from the mover to the impacted object, which may thereby be set in motion to participate in further interactions.
>
> (Langacker 1990:209)

As such, in transitive constructions, energy is transferred from subject-agent's domain to that of object-theme, and the transfer of energy results in certain change of the patient-object from one state to another. Following example illustrates this:

(23) *Juan comió una manzana.* 'Juan ate an apple.'

Example (23) is a transitive construction in which the subject's action transfers some energy to the object, *una manzana* 'an apple'. The action of eating conducted by *Juan* brings out some change of state in the apple, probably a certain decrease of its volume and content. In other

words, this example shows that the energy is transmitted from the subject's (*Juan*) domain to the object's (*una manzana* 'an apple') domain.

By extending this concept of domain to reflexive and middle constructions, it is possible to obtain a clear contrast between them: the reflexive form designates that the transferred energy involves two participants with the same reference, while in the middle constructions there are no differentiated participants, thus transfer of energy cannot be fully illustrated. Middle constructions represent actions, events or states that belong to the domain of the subject; therefore, the function of *se* cannot be attributed to the reduction of valence or transitivity; rather, it modifies the conceptualization of an event in several ways.

Moreover, as is described in 2.4.1, the main function of the clitic *se* has been analyzed as detransitivization, i.e. to reduce transitivity of the given structure. However, this is only the case for ergative verbs, where the grammatical object (patient) of the transitive structure, as seen in (24a), becomes the grammatical subject of the intransitive version, as in (24b):

(24) a. *Juan rompió la ventana.* 'Juan broke the window.'
b. *La ventana* ***se*** *rompió.* 'The window broke.'

This approach fails to account for some constructions where there is no valence that can be reduced, as in (25a) and (25b):

(25) a. ***Se*** *cayeron tres vasos.* 'Three glasses fell (down).' (Cuervo 2014:49)
b. *Tachita* ***se*** *sabe la canción.* 'Tachita knows the song by heart.' (Maldonado 1992:18)

The verb *caer* in (25a) is an intransitive verb belonging to the class of unaccusative verbs whose only argument, the grammatical subject, is an experiencer (or patient), never an agent. Since the intransitive structure itself is already 'detransitivized', there is no possibility for *se* to play a role as a detransitivizer, therefore it does not make sense to apply a detransitivization analysis in this case. In other words, there is no way the use of clitic *se* in (25a) can be an instance of valency

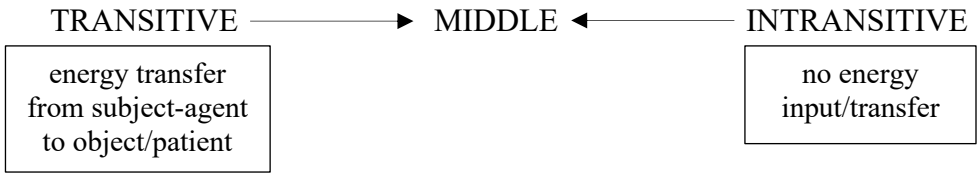
reduction because the verb *caer* itself is fundamentally intransitive (Alexiadou & Doron 2012:22).

Furthermore, it's also not possible to analyze (25b) from the detransitivization approach, which would argue that *saberse* should be considered an intransitive verb since *saber* is transitive. Although the clitic *se* used with ergative transitive verbs reduces the valence or argument of the sentence, making it an intransitive construction, we can see that *saberse* in (25b) still maintains its status as a transitive construction with a subject *Tachita* and an object *la canción* 'the song' even with the clitic *se*. From this example, it is plausible to infer that analyzing *se* as a detransitivizer is not always compatible with every construction where *se* is used.

Since the existence of non-detransitivizing *se* constructions is witnessed not only in Spanish —similar constructions are also found in various languages (Kemmer 1988)—, Maldonado points out that it is not pertinent to categorize these types of examples as exceptional cases and that they should not be considered a hindrance to linguistic generalizations. In Spanish, especially, the range of exceptions in which the use of *se* does not account for transitivity or valence reduction is too wide to disregard them as unusual or extraordinary, given that the use of non-detransitivizing *se* is exceedingly productive.

Based on this limitation of the previous analysis, Maldonado questions the traditional clear-cut distinction between transitive and intransitive verbs. It is also acknowledged by Doron (2003:6) who points out that "the direction of derivation which takes the transitive verb as basic and derives the middle verb faces a serious semantic problem". Following Kemmer (1993, 1994), Maldonado proposes an alternative view regarding transitivity in which there is a cognitive area called 'middle domain' and which contains a wide range of constructions

somewhere between two extreme poles: transitive and intransitive. The middle constitutes the unmarked choice for constructions which are neither transitive nor intransitive. These constructions are, in fact, in the middle along a gradation of maximum to zero energy input (Maldonado 1992:382-383):



According to Maldonado, the clitic *se* is used in these intermediary constructions functioning as a middle marker. He suggests that Spanish middle *se* constructions account for both a decrease in transitivity of transitive constructions as well as its increase in intransitive constructions. By moving away from the traditional binary framework of transitivity, Maldonado's alternative approach shows explanatory power with a more unified analysis of Spanish *se* constructions, which might seem like disparate and even idiosyncratic systems.

In summary, given that the traditional detransitivization approach on clitic *se* and the strong association of reflexivity with it do not seem to be relevant in numerous constructions and that they are too limited to cover important generalizations with regard to the structure of non-reflexive constructions marked with *se*, a different analysis needs to be considered. An alternative approach proposed by Maldonado is, to put it broadly, that the clitic *se* must be separated into reflexive and non-reflexive functions. His principal thesis is that non-reflexive *se* functions as a middle marker in middle voice constructions and that the middle voice is a unified phenomenon whose main function is to focus on the central properties of events. In other words, Spanish middle voice marked by the clitic *se* pays particular attention to the event expressed in the construction. As such, the conceptualization of the event can be reduced to the change of

state designated by the verb, and the function of the middle marker is to focus on (the pivotal moment of change of) the event itself (Maldonado 2000:159). Moreover, Maldonado emphasizes that Spanish middle constructions should not be restricted exclusively to English dispositional middle types of structures. He argues that all instances of *se* where it is used as a middle marker, i.e. non-reflexive *se* which cannot be used with an *a sí mismo* phrase, with its event-focusing function, belong to the middle domain in Spanish. These instances include grooming or body care, self-benefit actions, change in body posture, change in location, emotional reaction, emotive speech actions, mental change, and spontaneous event, which later on will be listed in Table 4 with some example verbs.

Instead of approaching the multiple functions of constructions marked with *se* to propose the existence of a variety of markers, Maldonado analyzed the semantic and/or syntactic composition of the clitic *se* from the perspective of cognitive linguistics. By analyzing a variety of constructions with *se*, focusing on their motivations and interconnections, he affirms that all non-reflexive uses of the clitic *se* are significant and well-motivated, not complex and/or controversial as some scholars (cf. King & Suñer 2017, Montrul 2004) have argued. In other words, instead of differentiating specific characteristics of each *se* construction, Maldonado focuses on the *middleness* that non-reflexive *se* constructions have in common. This middleness includes the aforementioned concepts such as lack of split representation of subject and object, low degree of elaboration of subject, lack of agentivity of the subject, and energy transmission remaining in the subject's domain. These result in the focus of the construction on the event itself, not on the subject or the object. Compare these two examples:

(26) a. *Juan **se** miró.* 'Juan looked at himself.'
b. *Juan **se** fue.* 'Juan left.'

(26a) is an example of reflexive *se*, in which the subject *Juan*, who is looking at himself and the object, the same *Juan,* who is being looked at by himself have split representation even though they refer to the same entity. Here, *Juan* carries out the action of looking at himself with his own will, hence the agentivity of the subject. On the other hand, a non-reflexive *se* construction in (26b) shows a situation where *Juan* is no longer present. Not only is it unknown if *Juan* left intentionally or not, it also doesn't matter. In this example, what really matters is that *Juan* has already left the place, thus he is not present at the scene anymore. According to Maldonado, this event-focusing reading is derived from the employment of the middle marker *se*. With the use of *se*, the speaker/conceptualizer gives the spotlight to the pivotal moment of the change of state of the event. As such, the middleness motivates the use of the middle marker, resulting in various non-reflexive *se* constructions in Spanish.

Calude (2017) further develops this concept of middleness and emphasizes how middle and passive constructions are different from each other, although they could be considered similar under the fact that they both belong to the category of non-active voice. In fact, middle voice is in opposition to both active and passive voice:

(27) a. *Juan abrió la puerta.* 'Juan opened the door.' (Active)
b. *La puerta fue abierta (por Juan).* 'The door was opened (by Juan).' (Passive)
c. *La puerta **se** abrió (*por Juan; *a sí misma).* (Middle)
'The door opened (*by Juan; *itself).'

By placing the agent in the subject position, as in example (27a), active voice evaluates the event from the agent's point of view based on the interests of the agent. On the contrary, passive voice in example (27b) focuses on the viewpoint of the patient, placing the patient in the position of the subject. The middle voice, however, affords relatively little attention to the participants and in contrast to the active and passive constructions, has a perspective centered on the event itself.

The distinction from passive is even clearer in that the agentive phrase, *por Juan* 'by Juan', is not compatible with the middle construction. Example (27c) also shows that the clitic *se* cannot be interpreted as a reflexive marker as shown in the ungrammaticality of adding the extended form, *a sí misma.*

Unlike the active and passive voices which focus on the participants —agents and patients, respectively—, the middle voice focuses exclusively on the event (Calude 2017:600), and the middle marker *se* highlights the central semantic properties of the verb with which it is combined to derive a variety of intensified readings (Maldonado 2000:155). This notion is also discussed by Elliott (1995), who summarizes three general types of participant roles as active 'participant', passive 'participant', and middle 'experiencer'. Upon describing the middle experiencer role in sentences with verbs that denote emotional changes (e.g. *volverse loco* 'go crazy', *quejarse* 'complain', *reírse* 'laugh'), he explains that "the role of the clitic pronoun is that of a middle experiencer because it generally describes a participant who undergoes (i.e. experiences) a mental change of state" (1995:208).

Based on these characteristics of the middle domain and by considering the clitic *se* as a middle marker, Maldonado expands the types of constructions which belong to the middle domain in Spanish and summarizes them as follows. What these constructions have in common is (i) low degree of distinguishability between subject and object, (ii) representation of the subject more as an experiencer than an agent, (iii) event remaining in the domain of the subject, and (iv) focus on the pivotal moment of change of state.

Table 4. Spanish basic middle voice constructions (Adapted from Maldonado 2008:164)

Interaction limited to body part or inalienable possession ~ grooming or body care	*lavarse, peinarse*
Self-benefit actions ~ benefactive middle	*conseguirse, allegarse*
Non-translational motion ~ change in body posture	*pararse, sentarse, voltearse, estirarse*
Change in location ~ translational motion	*irse, subirse, meterse*
Internal change (emotional) ~ emotional reaction middle	*alegrarse, entristecerse, enojarse*
Verbal actions manifesting emotions ~ emotive speech actions	*quejarse, lamentarse*
Internal change (mental) ~ cognition middle	*acordarse, imaginarse*
Changes of state whose energetic source is not identified ~ spontaneous event	*romperse, quebrarse, cerrarse, abrirse*

Given the ample range of middle domain, Maldonado further develops this notion to make an assumption that in some languages including Spanish, middle functions as a basic system. He argues that Yucatec Maya, Tarascan, Toba, Otomí, Amharic and Balinese, for example, show surprising similarities in terms of middle marking. Although this point of view contrasts with the general tendency to consider middle markers as evolved from reflexive markers, a cross-linguistic analysis shows that middle voice constitutes a separate system by itself and, in fact, can be considered as a basic form for other voices (2009:69).

In order to assure his argument that middle voice is the basic system, Maldonado takes a slightly different point of view from Kemmer, who perceives middle as evolving from reflexive constructions: Transitive > Reflexive > Middle > Intransitive. Maldonado points out:

> While reflexives may be the source for middles in languages whose prototypical event is the transitive construction, for other languages the middle may evolve directly from the transitive without depending on the reflexive and there may even be other languages like ergative ones for which events involving only one participant may be the starting point to derive other constructions.
>
> (Maldonado 2009:72)

In other words, Maldonado assumes that the middle voice can be the basic source with a less prototypical concept of transitivity in some languages of the world.

Such a distinctive nature of the middle voice, neither transitive nor intransitive, naturally explains the use of *se* to designate meanings that imply benefit, emphatic participants, uncontrolled acts, unplanned events, sudden events, counter-to-expectation events, and other meanings that depart from the prototypical transitive conceptualization of world events (Maldonado 1992). By rejecting the chain of evolution of verbs put forth by Kemmer (i.e. Transitive > Reflexive > Middle > Intransitive), Maldonado (1992:382) states that the middle involves an intermediate level of activity in which actions that are not totally agent-like and not totally patient-like tend to be expressed. Moreover, it reflects well the fact that the middle voice also involves "different manifestations of energy input either by the subject or by the conceptualizer" (1992:384) when the middle's own particular system is taken into consideration.

In sum, under the cognitive perspective, the middle voice is defined as a system of constructions whose main property is to represent the events that remain in the domain of the subject (Maldonado 2008:155). In most cases, the subject is considered an experiencer that somehow lacks agentivity, and it is proposed that the crucial function of the middle marker is to highlight the affectedness undergone by the experiencer. Therefore, internal emotions and emotional reaction middles (e.g. *alegrarse*, *entristecerse*, *enojarse*, etc.) can be represented as

the prototype for the middle voice because they designate events happening within the domain of the experiencer subject (2008:165). This fact also explains why spontaneous events (e.g. *romperse*, *quebrarse*, *cerrarse*, *abrirse*, etc.) belong to Spanish basic middle voice structures:

> In the same way that middles focus on the emotional change of the subject, they also focus on the crucial point in which the physical change takes place. The energy used to produce a change is not evident either because it is applied internally or because it is not profiled, as in spontaneous events.
>
> (Maldonado 2008:165)

In other words, Spanish clitic *se* plays an important role as a middle marker in various sentences that "share the property of focusing on the change of state undergone by an experiencer subject" (2008:165). The systematicity of this event-focusing property, which is the basic characteristic of middles, is determined by the contrast between the 'absoluteness' and the 'energeticness' of a given context. Maldonado proposes that, in its most schematic representation, Spanish middle *se* constructions have two basic functions: (i) they focus on the end point of the scene in which a thematic element undergoes a change of state; (ii) they allow the initiative force to be present in schematic terms (1993:534). The following figures adapted from Maldonado (2008:181) represent the contrast between absolute constructions which are not marked with *se* and energetic constructions which are marked by the clitic *se*:

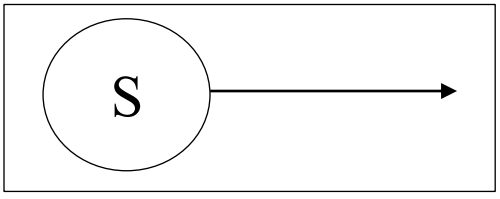


Figure 1. Absolute constructions

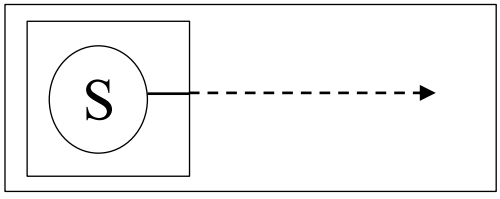


Figure 2. Middle/Dynamic constructions

Unlike Figure 1, which shows just one rectangle, the small rectangle inside the bigger one in Figure 2 represents the context focused on the pivotal moment of the change of state. This focus on the change of state is a characteristic that differs from absolute constructions and indicates the energetic feature of middle *se* constructions. That is, in absolute constructions, the event can be stative or dynamic, but no energy is put into profile (Langacker 1991, cited in Maldonado 2008:180). Although there is some kind of energy that is obviously involved in the context, no special attention is given to the energy/dynamicity involved in the event in these constructions. On the contrary, middle *se* constructions tend to focus on the energy involved in the pivotal moment of the change of state.

Furthermore, on top of this energetic nature, middle *se* constructions have some extended properties such as abruptness and unexpectedness. In the middle event, not enough information about the process or development of the action is given because the spotlight is on the result itself. Therefore, the event is depicted as rapid or abrupt since the energy's trajectory is not illustrated in the context. A further extension of this feature is the unexpectedness of the event. Here, there is an extra layer of energy which constitutes the speaker's expectations of the event. "The event is energetic not only in that it happens suddenly but the force-dynamic construal (Talmy 1985) makes it even more dynamic" (Maldonado 2008:183), as the following examples illustrate:

(28) a. *Cuando la lluvia cae sobre el lago hay una paz infinita.*
b. **Cuando la lluvia **se** cae sobre el lago hay una paz infinita.*
'When the rain falls on the lake, there is infinite peace.'

(29) *La lámpara **se** cayó de la mesa.* 'The lamp fell off the table.' (Maldonado 1993:536)

The use of *se* in example (29) implies that the conceptualizer's point of view is incorporated such that the event is sudden and accidental, i.e. out of the ordinary. Lamps are supposed to be on

tables (or other flat surfaces) and are not intended for the action of falling. On the other hand, a common event such as rainfall in (28) is not compatible with the use of the middle marker *se*. Rain is supposed to fall; it's the natural order of things. Thus, the grammaticality of the following examples can be explained in a similar way:

(30) a. ***Desperté** a media noche y ya no pude dormir.*

b. ***Me** desperté a media noche y ya no **me** pude dormir.*

'I woke up at midnight and I couldn't sleep anymore.' (Maldonado 1993:537)

The above examples show that the clitic *se* in (30b) marks an event in which a sudden or unexpected change of state from the viewpoint of the subject is outlined. In (30a) without the clitic *se*, the reading of the event described must be interpreted as one in which the change of state was not pivotal or in some way was the natural order of things. Such an interpretation is illogical and therefore ungrammatical since the act of waking up in the middle of the night is considered an unexpected change of state, that is, the speaker did not intend to wake up at that time.

When it comes to spatial domain expressed with movement verbs, concentrating our attention on the change of state with *se* reduces the profile of the process to the point of transition:

(31) a. *Juan **(*se)** va al cine todos los días.* 'Juan goes to movies every day.'

b. *No pude ver a Juan porque ya **se** había ido cuando yo llegué.*

'I couldn't see Juan because he was already gone when I arrived.'

(Maldonado 1993:538)

As Maldonado explains, in (31a), a complete path of movement is followed through space and time until the end point is reached. On the contrary, (31b) expresses a situation in which the spatial domain of the event is reduced, focusing only on the final state of the predicate, i.e. *Juan*'s having left already. Thus, *Juan* cannot be in the same spatial location as the subject, I. In

other words, the event in which *Juan* moves along the path from the source to goal is illustrated in (31a), while in (31b) the progression of *Juan*'s spatial movement is not visible: *Juan* is in the location of the source at one point, but in the next, he is no longer there.

Verbs which indicate the direction of motion, e.g. *subir* 'to go up', *bajar* 'to go down', also show the contrast between absolute and dynamic/energetic contexts. For example, *subir* 'to go up' outlines the entire path of the movement from a lower plane to a higher one, in the same way as *ir* 'to go' does in (31a). On the other hand, *subirse* 'to jump on' highlights the specific point in space and the crucial moment in which the change of location takes place:

(32) a. *Ximena subió la escalera.* 'Ximena climbed the stairs.'

b. *Ximena* ***se*** *subió a la mesa.* 'Ximena jumped on the table.' (Maldonado 1993:520)

As seen in the above examples, the verb *subir* 'to go up' expresses the movement that covers the entire path while *subirse* 'to jump on' only shows the end result of the movement that focuses on the moment *Ximena* escaped from the location of the source to the goal, *la mesa* 'the table', by jumping.

As described so far, a number of middle *se* constructions have a natural tendency to be interpreted as expressing dynamicity of the event. Also, it is possible to summarize that the two main functions of the middle marker *se* are to indicate "abruptness and accidentality" (Maldonado 1993:543). In other words, the Spanish middle marker *se* expresses contexts where abrupt and unexpected events or situations take place. As this interpretation of middle marked constructions directly relates to the conceptualization of the event, Maldonado concludes that "taking into account the point of view of the conceptualizer not only allows us to explain constructions that are not totally transparent, but also helps us to have a more enriched perception of language" (1993:562).

In sum, cognitive perspectives of the authors —especially Maldonado (1992, 1993, among others)— discussed in this section provide an alternative explanation on the complex Spanish *se* constructions. While detransitivization approaches fail to describe the polysemous nature of Spanish *se*, especially the *se* constructions in which the clitic *se* does not actually detransitivize the structure, the analyses provided by the cognitive approach can be applied to the ample middle domain (cf. Table 4) because they do not limit their explanation to a single grammatical process.

From the cognitive perspective, the clitic *se* in non-reflexive constructions marks the middleness, which bears the characteristics in between the transitive and intransitive constructions. These characteristics include inseparable representation of the subject and object, lack of agentivity of the experiencer subject, focus on the event itself and indication of abruptness and unexpectedness. This discussion is relevant to the present study because it is the main objective of the current study to examine how L2 learners of Spanish understand and utilize the clitic *se*, i.e. if they are aware that *se* has middle marking function along with the reflexive marking one. In the following section, a contrastive analysis on middle voice constructions in English and Korean will be provided in order to predict which properties of the two L1s would affect L2 Spanish learners' acquisition on middle *se* constructions.

2.5 Middle voice constructions in English and Korean

The Spanish middle system is quite distinct from both English and Korean middles. Contrasting the middle voice of each language with that of Spanish is necessary to predict how native English and Korean speakers will interpret and produce Spanish middle constructions. For

the purpose of the present study, the existence of the middle marker and its relationship with reflexivity will be the main focus of discussion.

According to Alexiadou & Doron (2012:4), in line with Kemmer (1993, 1994), there are four types of constructions which are marked as middle voice in various languages: (i) anticausatives, denoting spontaneous events (*break, open*); (ii) naturally reflexive verbs, e.g. verbs of body care (*wash, comb*) and naturally reciprocal verbs (*meet, kiss*); (iii) dispositional middles (*This book sells well*); and (iv) medio-passives, typically underdetermined for passive/anticausative constructions. The last type is not found in English. English examples of the other three types are as follows:

(33) a. The window broke from the pressure/by itself. (Anticausative)
b. The children kissed. (Reciprocal)
c. This book sells well. (Dispositional Middle)
(Alexiadou & Doron 2012:19)

As seen in the examples, all three types of middle constructions in English are formed by the active morphology of the verb. In other words, there is no explicit morpheme which marks middle, although it shows other characteristics such as the presence of adverbs (e.g. *easily, well*) and transitivity transformation in dispositional middles. Frequently, researchers investigating middle constructions exclusively analyze dispositional middles with no mention of other types of middle constructions. However, it is important not to limit the discussion to dispositional middle sentences given that the present study aims to recognize the middle domain in much broader terms.

The most distinctive difference between Spanish and English middle voice constructions is that Spanish employs a middle marker whereas English lacks one. In Spanish, the clitic *se* is used in both reflexive and non-reflexive, or middle, constructions. However, as Elliott (1995:23)

points out, the English reflexive pronoun *oneself* cannot be employed in non-reflexive cases when the subject is not logically capable of acting on itself. Although Elliott uses the term 'reflexive' to refer to the middle marker, his main argument is in agreement with the fact that Spanish makes use of *se* in non-reflexive contexts, unlike how the reflexive pronoun is used in English. It is clear that the English reflexive pronoun is only employed for constructions in which the subject-agent and object-patient/goal are coreferential. Additionally, active morphology without any specific middle marker is used in English middle voice constructions.

Notwithstanding the fact that English lacks a middle marker, some Spanish middle constructions marked by the clitic *se* are expressed by one grammatical structure in English, namely *get*-passive. As Alexiadou (2012:1089) points out, *get*-passives are discussed in connection with anticausatives and dispositional middles in the literature (cf. Hatcher 1949; Givón and Yang 1994; Arrese 1999; McIntyre 2012). She attributes this comparability to the status of the verb *get*. In comparison with the verb *be* in *be*-passive constructions (e.g. The door **was** closed), *get* behaves as a lexical verb rather than an auxiliary.

Alexiadou explains that there are two sentence types of middle voice that can be also expressed by the *get*-passive in English. First of all, *get*-passives and dispositional middles share intransitive-like structures with typically transitive verbs, e.g. *wash*:

(34) a. The clothes got washed. (*get*-passive)
b. These clothes wash well. (Active morphology)
c. *The clothes washed. (Alexiadou 2012:1089)

Second, *get*-passives are similar to anticausatives in the sense that both refer to eventualities that come into being without duration. In other words, *get*-passives may alternate with anticausatives with verbs of Achievement[9]:

(35) a. The mould got broken. (*get*-passive)
b. The accelerator cable broke. (Active morphology) (Alexiadou 2012:1089)

Also, there is another type of *get* construction that is closely related to the middle marking function of Spanish *se*. It is not explicitly described by Alexiadou (2012) because this type does not belong to the category of *get*-passives. Consider the example below:

(36) Sometimes students get anxious about this, and believe they are not making audience contact early enough. (Alexiadou 2012:1081)

Alexiadou explains that the example (36) shows the inchoative use of the verb *get*. Based on the semantic similarities, this would be compared to the emotional reaction type in Spanish (e.g. *ponerse ansioso* 'to get anxious', *alegrarse* 'to get happy', *entristecerse* 'to get sad', *enojarse* 'to get angry').

From the above examples, it is plausible to relate Spanish middle *se* constructions with English *get*-passives and the inchoative use of *get,* based on the fact that they both have the event-focusing function on the pivotal moment of change of state. However, there still remain some differences in that English *get*-passives are not as productive as the Spanish middle marker *se*:

(37) a. Grooming or body care: ***Me*** *peino antes de salir.* 'I comb my hair before going out.'
b. Self-benefit anctions: *¿Dónde* ***me*** *consigo esa camisa?* 'Where do I get that shirt?'

[9] Achievement forms part of the *Aktionsarten*, along with State, Activity, and Accomplishment (cf. Footnote 2). The distinctive characteristics of the Achievement verbs include the semantic notions: [+dynamic], [+telic], and [+punctual].

c. Change in body posture: *Juan siempre **se** para a saludar a Pedro.*
'Juan always stops to say hello to Pedro.'

d. Change in location: *María **se** fue.* 'María left.'

e. Emotional reaction: *Juan **se** enoja muy fácil.* 'Juan **gets mad** really easily.'

f. Emotive speech actions: *María **se** queja de dolor en la cabeza.*
'María complains about headache.'

g. Mental change: ***Me** acuerdo que la última vez que fui de compras, no tuve suficiente dinero.* 'I remember that the last time I went shopping, I didn't have enough money.'

h. Spontaneous event: *El molde **se** rompió.* 'The mould **got broken**.'[10]

Above examples represent each Spanish basic middle voice situation classified by Maldonado (2008:164) in Table 4. It is shown that only two examples in Spanish are translated in English with *get* constructions: emotional reaction (e.g. *enojarse* 'to get mad') in (37e) and spontaneous event (e.g. *romperse* 'get broken') in (37h). This is not surprising if we take into account that Spanish *se* is a clitic which can be utilized with a wide variety of verbs whereas its English counterpart discussed here is limited to certain examples in which the verb *get* is used. In other words, while Spanish *se* plays a role as a middle marker in middle constructions, English *get* is still a verb and it does not necessarily mark the middleness of the events.

In summary, the English middle voice system is different from that of Spanish in that (i) the English reflexive pronoun cannot be employed in middle as well as in any type of non-reflexive construction; (ii) a specific morphology to mark middle voice does not exist; and (iii) a regular active voice morphology is used in middle constructions. However, it is possible to draw some similarities between English *get*-passive and inchoative *get* constructions and Spanish

[10] This example (37h) is taken from Alexiadou (2012:1089).

middle *se* constructions based on the lexical property of the verb *get* which is compared to the 'true' auxiliary *be*, used in prototypical *be*-passive constructions.

On the other hand, Korean is like Spanish in that there is a distinct middle marker. However, Spanish and Korean are different in that the latter uses one morpheme (*-i*) to mark middles and a different morpheme (*casin*) to mark reflexivity:

(38) a. *Ella **se** vio en el espejo (a sí misma).* (Reflexive marker *se*)
'She looked at herself in the mirror.'

b. *Este libro **se** vende bien.* (Middle marker *se*)
'This book sells well.'

(39) a. *kunye-nun kewul-eyse **casin**-ul po-ass-ta.* (Reflexive marker *casin*)
she-NOM mirror-in REFL-ACC see-PST-DECL
'She looked at herself in the mirror.'

b. *i chayk-un cal phal-**li**[11]-n-ta.* (Middle marker *-i*)
this book-NOM well sell-MM-PRES-DECL
'This book sells well.'

In fact, determining the morpheme *-i* as a middle marker in Korean is controversial among linguists. Some consider it a passive marker, denying the existence of middle voice in Korean, while others define it as a multifunctional morpheme, with uses as a causative, middle and passive marker. The present study follows the latter approach, based on Kim (2013) who takes the theoretical framework of Alexiadou & Doron (2012). Kim argues that middle, passive and causative constructions in Korean should be categorized under non-active voices given the fact that they (i) share the multifunctional marker *-i* (and its phonological variants *-hi, -li, -ki*) and (ii) prevent the existence of an external argument, which is generally the subject of the

[11] The morpheme *-li* is a phonological variant of *-i*.

sentence. More specifically, middle constructions differ from the other two constructions in that they are not compatible with agentive phrases:

(40) *kam-ta* > *kam-**ki**-ta*
wash[12]-DECL wash-*i*[13]-DECL

a. *aki* *meli-nun* *cal* *kam-**ki**-n-ta*. (Middle)
baby hair-NOM well wash-*i*-PRES-DECL
'Baby hair washes well.'

b. *aki-uy* *meli-ka* *emma-ey-uy-hay* *kam-**ki**-n-ta.* (Passive)
baby-POSS hair-NOM mom-A wash-*i*-PRES-DECL
'The baby's hair is washed by her/his mom.'

c. *emma-ka* *aki-uy* *meli-lul* *kam-**ki**-n-ta.* (Causative)
mom-NOM baby-POSS hair-ACC wash-*i*-PRES-DECL
'Mom makes the baby's hair get washed.'

As can be seen in the above examples, a transitive verb *kam-ta* is marked with the multifunctional morpheme *-(k)i* in dispositional middle, passive and causative constructions in Korean. However, previous studies on the voice marker *-i* focused mainly on determining its causative and passive uses, paying little attention to its middle-marking function. Even in recent studies which examine *-i* as a middle marker, they approach it as a passive marker employing terminologies such as 'potential passive', 'middle or medio passive' and 'generic passive' (Yap & Ahn 2019:3).

---

[12] *kam-ta* is polysemous: to close or shut (one's eyes); wind (up), coil (up/around), reel in; twine (one's arms around somebody); wind, fast-forward, rewind; wash, bathe (*kam-ta*. 2020. In *en.dict.naver.com.* Retrieved February 10, 2020, from https://en.dict.naver.com/#/search?query=%EA%B0%90%EB%8B%A4). When it refers to 'wash, bathe', it generally means to wash one's hair.

[13] The multifunctional morpheme *-i* is not glossed as MM (Middle Marker) here in order to ascertain that it has other functions (e.g. passive marking and causative marking) as well.

Based on the fact that the middle-marking function of *-i* is not sufficiently acknowledged in the literature, it is not surprising that native Korean speakers overuse passive constructions in their L2 English in cases where middle constructions are necessary. For instance, Suh (2013), who reports her findings from a Grammaticality Judgment Test (GJT) with 59 Korean university students who learned English as a second language, concludes that native Korean speakers recognize English middle constructions as having a similar, if not the same, pattern with passive constructions. Specifically, almost all of her participants judged English middle sentences with verbs in their intransitive structures as ungrammatical and, when asked to correct them, they changed them into passive sentences[14]:

Table 5. Examples of error correction by native Korean speakers (Suh 2013:181, 184)

Middle sentence	Corrected sentence by the participants
The potatoes **bake** well.	The potatoes **are baked** well.
Whole wheat bread **cuts** easily.	Whole wheat bread **is cut** easily.
This piano **plays** easily.	This piano **is played** easily.

Above examples show native Korean speakers who participated in Suh's (2013) study, although they are considered to be highly proficient in English, incorrectly judged middle sentences ungrammatical. By using '*be* + past participle' phrases, they rewrote the 'ungrammatical' middle examples into passive sentences.

[14] It is also possible to assume that the same judgments could be made with Korean examples as well, since native Korean speakers are explicitly taught, usually in secondary school, that the morpheme *-i* functions as a passive marker, and it is not very common for them to perceive it as a middle marker.

To summarize, Spanish and Korean middle voice systems are relatively similar since both employ the multifunctional marker *se* and *-i*, respectively. Although they show some overlapping usages, they are not completely equivalent. Korean has a distinct reflexive marker (*casin*) which marks coreference between the subject-agent and object-patient/goal of a given construction. Also, the Korean middle marker *-i* cannot be used in anticausative constructions (Kim 2013:48) while Spanish *se* is compatible with them, which results in *-i* lacking an event-focusing function that Spanish *se* possesses. Instead, the auxiliary construction *-a/e pelita* marks accidentality and dynamicity of an event.[15]

Based on the above descriptions, Spanish, English and Korean middle voice systems can be summarized according to the following table:

Table 6. Comparison of Spanish, English and Korean middle voice systems

	Spanish	**English**	**Korean**
Reflexive marker	*se* (*a sí mismo*)	oneself	*casin*
Reflexive:	*Ella **se** vio **(a sí misma)** en el espejo.*	She looked at **herself** in the mirror.	*kunye-nun kewul-eyse **casin**-ul po-ass-ta.*
Middle marker	*se* (**a sí mismo*)	N/A	*-i,* in certain contexts
Dispositional middle:	*Este libro **se** vende bien.*	This book sells well.	*i chayk-un cal phal-**li**-n-ta.*
Grooming/body care:	*Ella **se** lavó la cara.*	She washed her face.	*kunye-nun elkwul-ul ssis-ess-ta.*
Dynamic (energetic/ spontaneous) events:	*La puerta **se** cerró.*	The door (got) closed.	*mwun-i tat-hi[16]-e peli-ess-ta.*

[15] For example, *kata* means 'to go' whereas *ka-pelita* means 'to go away'. In Spanish, a similar comparison can be made between *ir* 'to go' and *irse* 'to go away' (cf. Strauss 2003).

[16] The morpheme *-hi*, which is a phonological variant of the marker *-i*, is considered to be used as a passive marker (cf. Lee 1993). Whether it can be analyzed as a middle marker needs further research.

The most prominent difference found between the three languages is that the Spanish clitic *se* has the widest range of usage, that is, in both reflexive and non-reflexive contexts.[17] Reflexive markers in English and Korean cannot be used in non-reflexive constructions, and English generally employs active morphology in most contexts in question. Korean seems to appear in a medial position between Spanish and English in that it has a multifunctional marker like Spanish, but its presence is not as easily observed as in Spanish. Given the mismatch in the three languages, it is plausible to predict that both English and Korean speakers will have to develop a target language-specific system in order to fully acquire the Spanish middle domain. Detailed discussion on possible positive and negative L1 (and L2, in case of native Korean speakers) transfer related to this issue can be found in Chapter 5.

[17] There are some analyses that explain such multifunctionality of clitic *se* in Spanish as a result of diachronic changes. Latin had the pronoun *se* with a reflexive use and to transform active voice to middle voice. Spanish inherited both the reflexive *se* form and these functions, but also expanded its uses during the Middle Ages and the Renaissance. For example, when the Latin passive markers were lost, passive voice was expressed with *se* (Whitley 1998). Since the Renaissance, the uses of *se* continued to expand to include more middle constructions (Ranson & Quesada 2018).

CHAPTER 3

LANGUAGE ACQUISITION AND CLITIC *se*

3.1 Introduction

In this chapter, I discuss how the Spanish clitic *se* has been studied in First and Second Language Acquisition. The reason for reviewing first language acquisition studies on *se* is twofold: (i) there is a lack of previous literature on Spanish L2 acquisition of *se* as a middle marker, and (ii) the results of first language acquisition studies may offer some meaningful insight in analyzing L2 data. Also, it should be noted that the scope of this review is not limited to L2 acquisition of clitic *se* as a middle marker but also includes the L2 acquisition of *se* in general, including various 'types' of *se* (e.g. impersonal, inchoative, passive), which reflects to a large degree how Spanish clitic *se* is presented in textbooks. Thus, it is also necessary to include a review and analysis of how the clitic *se* is presented and taught in some textbooks that are used in Spanish L2 classroom for English and Korean speakers. Finally, the lack of attention towards Spanish *se* as a middle marker in L2 acquisition and its importance in learners' learning processes will be discussed.

3.2 L1 acquisition and native speakers' use of Spanish clitic *se*

Characterizing and determining the middle voice domain is a critical issue in cognitive linguistics and its linguistic implications in acquisition are yet to be examined in detail. Nevertheless, there have been some studies in L1 acquisition whose results have supported

Maldonado's concept of middle domain and have revealed that *se*'s middle function is acquired before its reflexive property (Jackson-Maldonado, Maldonado & Thal 1998; Soto & Muñoz 2000). Other studies such as Anderson (1998) suggest *se* is acquired early on but its contrastive use as a middle marker is developed later in children's speech.

First of all, Jackson-Maldonado et al. (1998) examined L1 acquisition of the clitic *se* by Mexican children aged 2;4 to 3 years. The purpose of this research was to apply a cognitive grammar model to analyze children's first uses of *se*. They worked under the assumption that existing studies of the acquisition of Spanish clitics used inadequate models, and that "cognitive models are needed for an accurate explanation of clitic acquisition" (1998:404). Specifically, they pointed out the basic problem in most of the previous acquisition studies is that "they have classified all clitics as reflexives, although most of them are really object clitics, inchoatives, impersonals, datives or emphatic datives" (1998:414), and some studies "tell us that clitics appear at an early age, but not how they are used" (1998:415).

A total of 37 children from monolingual Spanish-speaking families living in Mexico or California participated in the study with their parents and/or research assistants. They were asked to play with a variety of toys placed on the floor as they would at home. Researchers extracted three separate 10-minute samples from each child and transcribed them using the Child Language Analysis System (MacWhinney 1995) to identify every occurrences of *se*. They tabulated the number of tokens and classified them into 11 different categories and reported them in percentages (presented below in Table 7); the small size of the data set made statistical comparisons impossible.

Table 7. Percentage of *se* tokens as reflexive, lexical and middle forms (Adapted from Jackson-Maldonado et al. 1998:420)

Category	Example	# of tokens	Percentage
Reflexive forms		**15**	**9%**
Reflexive *per se*	*Valeria **se** vio en la foto.* 'Valeria saw herself in the picture.'	0	0
Dative reflexive	***Se** puso el sombrero.* 'S/he put on her/his hat.'	11	7%
Reciprocal	*(Los leones) **se** pelearon.* '(The lions) had a fight.'	4	2%
Lexicalized forms	*¿Cómo **te** llamas?* 'What is your name?'	**6**	**4%**
Middle forms		**140**	**87%**
Motion	***Se** sentó en un árbol.* 'S/he sat on a tree.'	52	32%
Unexpected changes	***Se** cayó.* 'S/he fell down.'	49	30%
Change of state	***Se** durmió.* 'S/he fell asleep.'	16	10%
Impersonal	***Se** pone aquí.* 'One must put it there.'	15	9%
Uncontrolled actions	***Se** pegó así.* 'S/he hit her/himself like this.'	3	2%
Irregular analogies	-	2	2%
Grooming	***Se** va a peinar con esto.* 'S/he is going to comb (her/his hair) with this.'	3	2%

The results showed that the Spanish-speaking children's most frequent uses of *se* were observed in middle forms instead of reflexive ones, a fact that affirms that the middle voice can be a basic system or a starting point for other constructions. According to the authors, use of the middle marker is motivated especially by the need to focus on a critical moment of change and is an aspectual phenomenon based on cognition. Therefore, the early appearance of *se* in native Spanish speakers is better explained from the cognitive perspective, in that it occurs due to the cognitive relevance of the event described by the verb with the clitic *se* as a middle marker.

Similar results were obtained from Soto & Muñoz (2000), who confirm that children from very early ages prefer the use of *se* with the event-focusing function. Using a corpus collected in Argentina and Chile, they conducted ontogenetic research that explored the use of the middle marker *se* by Spanish-speaking children between 3 to 11 years of age. The data were collected from children's narrations of a picture book[18] story. Unfortunately, this study does not provide any information about how the data were collected and measured. Only several examples extracted from the corpus are given.

Soto & Muñoz's (2000) results demonstrated that in the development of narrative discourse, children focus and use *se* where there is great energy transmission, among other uses, from the earliest age:

(41) *Ahí metió la cabeza, estaba mirando en la ventana y ¡**se** cayó! el perro y **se** rompió.*
'(The dog) stuck its head in it (lit. there), looking out the window and the dog fell down and (it) got broken.' (Soto & Muñoz 2000:204, [translation mine])

[18] The same picture book was used in one of the tasks of the present study, which will be described in detail in Chapter 4. By using the same material in the current experiment, my results are expected to display similarities and differences between the acquisitional processes of L1 and L2/3.

According to this study, the function of *se* that marks dynamic and energetic events has an ontogenetically basic conceptual status. When there is an instance of greater energy transfer which does not easily conform to the prototypical transitive constructions, children tend to employ the construction with a middle marker *se* as a means of conceptualizing the event. Therefore, this research reaffirms that the middle marker *se* in Spanish is highly productive and is mainly employed to focus on the event, especially on the pivotal moment of change of state.

Anderson (1998), on the other hand, suggests that while children may employ non-reflexive or middle *se* constructions early on, that does not necessarily mean that they are able to use them contrastively. The author classifies the various functions of Spanish clitic *se* in five different categories:

> **(a) regular reflexive meaning**, where the clitic is used with a transitive verb to indicate equivalence between the subject and the object of the clause (e.g. *María **se** vió en el espejo* 'María saw herself in the mirror'); **(b) reciprocal meaning** (e.g. *María y Juan **se** vieron en el espejo* 'María and Juan saw each other in the mirror'); **(c) impersonal meaning**, where an inanimate entity is the focus and the verb is transitive (e.g. ***Se** venden libros* 'Books are sold'); **(d) spurious dative**, where *se* substitutes for the dative third person clitic *le* (e.g. *Teresa **se** lo quitó* 'Teresa took it away from him/her'); and **(e) Romance reflexive meaning**, where the reflexive clitic is used with verbs that have been traditionally described as intransitive (e.g. *La niña **se** durmió* 'The girl fell asleep')
>
> (Anderson 1998:489, [emphasis mine])

Anderson emphasizes that the lexical meaning of the verb changes when the last category, a Romance reflexive, is used. According to the author, the clitic *se* used with Romance reflexive meaning functions as transitivity marker by expressing punctuality and telicity of the event. Although the use of *se* results in a semantic difference in this type of construction, Anderson describes that the difference is "faint" (1998:492) and "minimal" (1998:492), which explains the

reason why the contrastive use of the Romance reflexive form is seldom observed in children's speech.

With the purpose of assessing the differential development of regular and Romance reflexive constructions, Anderson conducted an experiment with 40 monolingual Spanish-speaking children between the ages of 2;0 and 3;11 from the San Juan metropolitan area of Puerto Rico. The study consisted of two tasks and the children's speech data were collected to assess the contrastive use of *se*. Additionally, 20 adult native speakers of Puerto Rican Spanish participated in the study as a control group.

The two tasks conducted in this study are defined as Romance reflexive task and Regular reflexive task, respectively. The first one was a storytelling task[19], in which the participant was asked to complete a story that the researcher started. The story narrated by the researcher implied a situation which favored either the non-*se*-marked form (e.g. *dormir* 'to sleep') or the *se*-marked form (e.g. *dormirse* 'to fall asleep'). For example, if the researcher tells a story such as "*Luisito estaba muy cansado de tanto jugar. Por eso, se sentó en el sillón, cerró los ojos, y ¿qué le pasó?* 'Luisito was very tired from playing. He sat on his rocking chair. He closed his eyes, and what happened?'" (1998:510), it is expected that the participant would answer using the *se*-marked form, e.g. "*Se durmió* 'He fell asleep.'" (1998:510).

The second task, Regular reflexive task, was only provided to the children group. Three people were included in this task: the participant, the researcher, and her assistant. First, the researcher would comb her hair and say "*Me estoy peinando* 'I am combing my hair'"

[19] This is similar to the Acceptability Judgment Task (AJT) conducted in the present study, in a sense that a situation targeting a specific verbal form was provided to the participant. More detailed description of the AJT will be given later in Chapter 4.

(1998:498). Then, she asked the participant what she was doing, expecting the child to respond "*Se está peinando* 'She is combing herself'" (1998:498). Next, the researcher would comb the assistant's hair and say "*Te estoy peinando* 'I am combing you'" (1998:498). The last procedure was to ask the child what the researcher was doing, and the expected answer was "*La estás peinando* 'You are combing her'" (1998:498).

In sum, the first task was designed to examine participants' contrastive use of the Romance reflexive, while the second task was to compare a prototypical two-participant transitive construction to a 'regular' reflexive one. Since the latter is not directly related to the focus of the present dissertation, further discussion will be centered on the results of the Romance reflexive task.

For each of the prompts and their corresponding responses provided by the participants in the first task, Anderson (1998) analyzed the presence of contrastive use of the Romance reflexive *se*. Both children and adults revealed contrastive use, although the frequency rate was higher for adults. Particularly, over 70% of the total number of noncontrastive responses consisted of the overuse of *se*.

The results demonstrated that children use *se* in Romance reflexive constructions as early as 2;1 years of age. However, it takes more time for them to understand that using *se* in Romance reflexive context expresses a different layer of meaning such as abruptness, accidentality or unexpectedness. Meanwhile, adult control groups' data showed that even adults mostly failed to produce the Romance reflexive *se* contrastively. For this unexpected result, Anderson (1998:497) explains that "the fact that adults varied in their responses suggests that individuals vary with respect to how they interpret the discourse context, and that their choice of patient focus

(Romance reflexive) or action focus (no Romance reflexive)[20] will depend on each individual's perspective of the situation".

Furthermore, children mostly used *se* with traditionally intransitive verbs (e.g. *caer* 'to fall', *ir* 'to go', *dormir* 'to sleep'). This pattern suggests that children are biased toward a transitive interpretation of events and employ the clitic *se* as a transitivity marker. Anderson (1998:509) attributes this pattern to the argument that "children do evidence a bias toward marking transitivity, but that the perceived causative effect varies across verb meanings and affects the children's (and adults') choice of using the reflexive pronoun *se*".

The question of the order of acquisition between two clitic *se* constructions, reflexive *se* or middle *se* more closely reveals the L1 acquisition process and needs further research with more data to be collected. Because there is a great deal of confusion regarding the domain of middle *se*, the answer may well lie within how researchers define the functions of *se*. Furthermore, it will be of interest to compare the existing L1 data to L2 data because any similarities and differences between them can give us some important insights into the cognition of this particular linguistic form and into the acquisitional process of human language in general. It is not an easy task to compare them, unfortunately, given that L2 studies so far have not conducted enough experiments with spontaneous or open-ended production tasks. The present study intends to partially fill this gap.

A brief summary of L1 studies discussed in this section is as follows:

[20] The term "no Romance reflexive" (Anderson 1998:508) refers to the non-*se*-marked bare verb form.

Table 8. L1 studies of the acquisition of Spanish clitic *se*

Researcher (Year)	**Age of participants**	**Number of participants**	**Location**	**Type of data**	**Result(s)**
Anderson (1998)	2;0-3;11	40	Puerto Rico	Spontaneous oral speech	Middle (Romance reflexive) *se* is acquired early, without contrastive understanding.
Jackson-Maldonado et al. (1998)	2;4-3;0	37	Mexico and California	Spontaneous oral speech	Middle *se* is acquired before reflexive *se.*
Soto & Muñoz (2000)	3;0-11;0	N/A (Corpus analysis)	Argentina and Chile	Guided open-ended oral speech	Clitic *se* is employed mainly to mark dynamic and energetic events.

Aaron & Torres Cacoullos (2006) is a variationist analysis on the use of the middle marker *se* by native speakers of Mexican Spanish. Although this study does not belong in the category of acquisition studies, it is worth addressing here, because, like the present study, it treats Spanish *se* as part of Maldonado's (1992 and so forth) Spanish middle domain. It is well known that the use of pronominal verbs and constructions with *se* varies from dialect to dialect in the Spanish-speaking world (Liceras 1999; Ibarretxe-Antuñano, Cadierno & Hijazo-Gascón 2016). Aaron & Torres Cacoullos' (2006) work contributes to the literature recognizing this linguistic variation of the use of *se*.

In particular, the authors examined the form *salir(se)* 'to go out; to leave' using the Variationist method of Labov (1972) to investigate factors that determine the choice between the two alternating forms. They specifically analyzed *salir(se)* because of Maldonado's (1999) work, which characterized intransitive verb constructions with *se* as 'energetic'or 'dynamic' constructions that contrast with the 'absolute' constructions which lack the use of *se*. These energetic constructions have evolved along the path of 'subjectivization' (a term used by Aaron & Torres Cacoullos (2006); also known as 'subjectification' or 'pragmaticalization' (a term also used by Maldonado (1999)): Absolute (without *se*) > Focus (with *se*) > Unexpected (with *se*).

In contexts where the form *salirse* occurs, it could be considered that the speaker expresses, in addition to the propositional meaning, her/his point of view and attitude toward the situation. Specifically, *salirse* belongs to the third stage of subjectivization/pragmaticalization, since it expresses an event that is unwanted or unexpected, e.g. leaving against obstacles; leaving suddenly; permanent abandonment; leaving without practical purpose (Aaron & Torres Cacoullos 2006:46-49). Yet, despite what at first seems to be a clear differentiation between the functions of the *se*-marked and non-*se*-marked forms, some examples are found —even in the speech of the same person, separated by some intermediate lines— in which *salir* and *salirse* appear in almost identical contexts, apparently with the same semantic and pragmatic value.

Faced with this situation, Aaron & Torres Cacoullos (2006) confirm that "in cases of variation like this, the most powerful tool is the Variationist method, which allows us to take a glimpse into the general patterns of use and thus measure the greater or lesser subjectivity in one way or another" (2006:51, [translation mine]) because "given this polyvalence of the relations between form and function, a qualitative examination of the uses of these two forms alone is not enough to verify that the marked form is really more subjective than the unmarked one"

(2006:51, [translation mine]). The quantitative analysis of seemingly random alternations under the Variationist method allows for analyzing patterns of use of certain linguistic forms or elements.

The dataset that Aaron & Torres Cacoullos (2006) analyzed consists of a total of 557 cases of *salir(se)* extracted from a corpus. Only 12% of cases showed the use of *salirse*, while the majority (88%) was the non-*se*-marked form, *salir*. After excluding all irrelevant contexts, 274 tokens (49% of the total) were examined statistically with the GoldVarb (Rand & Sankoff 1990) program. The independent factors that were measured included: grammatical person (first person singular *vs*. others); relationship with the speaker (close *vs.* distant); co-presence of dative pronouns; Time-Aspect-Mood (preterite *vs.* others); co-presence of the preposition *de* 'of'; and polarity (affirmative *vs.* negative). They assumed that the first person singular grammatical person, close relationship, presence of dative pronouns, preterite, presence of *de* and affirmative polarity should favor the use of *salirse*.

Their results showed that the first person singular, close relationship with the speaker, preterite and affirmative sentence condition the choice of *salirse*. In addition, although they excluded the factor 'co-presence of dative pronouns' in the statistical analysis due to their few cases of appearance in the data ($n = 15$), their hypothesis on this factor was confirmed as true given that 66% of the tokens show the pattern of use of the datives with the form *salirse*. The last factor that remains to be discussed, co-presence of the preposition *de*, was not considered statistically significant, with a probability of .48. Even though this probability was below .50, Aaron & Torres Cacoullos (2006) claimed that it still shows the direction towards the expected pattern.

What this research elucidates for acquisition studies is that the Variationist approach with its quantitative analysis is more suitable for production data when the linguistic phenomenon in question demonstrates a certain degree of variability, as is the case for the Spanish middle marker *se*. Also, that the use of *se* is far from being paradigmatic in native speaker speech leads us to predict that non-native speakers will encounter a great deal of difficulties, even fail to grasp the pattern of usage, when interpreting and producing constructions with *se*. For instance, over- and/or under-use of *se* can be shown in L2 learners' speech due to their lack of understanding of clitic *se*'s various functions. Also, they are less likely to recognize the contrastive reading of the context relevant to the presence and absence of *se* (i.e. dynamic reading of *se*-marked constructions *vs.* absolute reading of non-*se*-marked constructions).

3.3 L2 acquisition of Spanish clitic *se*

As described above, although there are handful of studies in L1 Acquisition whose research is based on the cognitive perspective and confirm that the clitic *se* functions as a middle marker, L2 acquisition studies following this line of research are nearly nonexistent. In general, existing analyses of the uses of the clitic *se* in Spanish by L2 learners are carried out especially with English-speaking learners with different levels of proficiency of the target language, and mostly from a generativist perspective and relying on grammaticality judgments. There are a handful of studies that examine the acquisition of Spanish *se* by speakers of other L1s (Finnish, French) and from other theoretical perspectives (usage-based and cognitive grammar) using corpus and other production data.

Despite the differences in methodologies and theoretical framework, previous studies agree on one thing: in contrast to L1 acquisition, the L2 acquisition of *se* is marked by great

difficulty and occurs at a later stage of second language development. It is assumed that this learnability problem is derived from inefficient and inappropriate traditional pedagogical approaches of the clitic *se* in Spanish (Maldonado 2008:156). In addition, the difference in form-function mapping between L1 and L2 can also affect the acquisition process.

3.3.1 Generative studies

Generativist research on L2 acquisition of Romance 'reflexive' constructions with *se* has gained attention from the late 1990s (Tremblay 2005). Clitic *se* constructions in Spanish are considered ideal for generative and second language acquisition analyses because it is predicted that English-speaking L2 learners cannot make any direct transfer of surface structure between their L1 and L2. This makes it possible to test learners' access to Universal Grammar (henceforth, UG) in the acquisition process of L2.

According to the generativist approach, there is an innate linguistic faculty which consists of UG that is responsible for the acquisition of language(s). UG imposes restrictions on the form of languages in terms of Principles and Parameters. Principles are "constraints on grammatical wellformedness that apply universally" and Parameters are "limited enumerated options on how a given property can be obtained in particular grammars" (Rothman & Pascual y Cabo 2014:47). In other words, there is a set of principles that all languages share, and the differences between languages are due to the way the parameters are set.

The task of L2 learners is to learn the target language's idiosyncratic properties from their L1 and to reestablish target-specific settings for parameters, along with UG principles (Sánchez & Toribio 2003:221). How different the systems are between L1 and L2, thus to what extent learners have to (re-)set parameters in the L2, results in learnability problems. The generativist

framework assumes that in this process of overcoming learnability problems and (re-)setting parameters in L2, UG plays a significant role. In simple terms, if there is access to UG, L2 learners can overcome the learnability problems and succeed in acquiring target-specific features that are non-existent or different in their L1. However, if access to UG is not available, non-native speakers' language performance will be impaired even in their later stages of L2 development.

Based on the fact that various Spanish clitic *se* constructions greatly differ from their similar counterparts in English, there has been relatively more research done from the generativist perspective. However, conclusions regarding UG access vary, depending on each study's specific experimental results.

In an attempt to analyze adult French- and English-speaking learners' L2 acquisition of Spanish, Tremblay (2005), conducted an experiment which partially replicated the work of Bruhn de Garavito (1999a, 1999b). Bruhn de Garavito examined the acquisition of three *se* constructions (i.e. passive, impersonal and inchoative)[21] with three groups of L2 learners: L1 English Advanced, L1 English Near native, and L1 French Near native speakers of L2 Spanish. There was also a group of L1 Spanish speakers as a control group. In the study, the participants were asked to judge the grammaticality of various Spanish *se* constructions based on a Likert scale ranging from -2 to 2. Test items provided in the survey included both grammatical and ungrammatical sentences with different word order, subject-verb agreement, subject position, *pro*-drop, secondary predication, and object marking with the preposition *a*. Results showed that the L1 English and L1 French near natives performed in a native-like way, without significant

[21] In Bruhn de Garavito's (1999a, 1999b) terms, impersonal passive, impersonal reflexive and inchoative.

differences between their L1s. L1 English advanced learners also revealed native-like judgments, although to a lesser degree. The author suggests that this study gives support to the assumption that L2 learners have access to UG when it comes to the acquisition of these three *se* constructions.

Similarly, Tremblay (2005) also studied the L2 acquisition of Spanish *se* by English- and French-speaking learners, focusing on passive and impersonal *se* constructions since they pose a potential learnability problem. Neither construction exists in English, and the impersonal *se* does not exist in French. As a result, passive and impersonal *se* constructions are not present in the learners' L1, nor are they salient in the input, and thus are particularly difficult for both French and English speakers to acquire in Spanish.

In addition, passive and impersonal *se* constructions are highly similar in their superficial structure, but each of them shows idiosyncratic morphological and syntactic characteristics. Passive *se* checks the accusative case, and the internal argument, *los mejores profesores* 'the best teachers', has the nominative case that agrees with the verb, as in (42a). On the other hand, the impersonal *se* checks the nominative case, and the internal argument has the accusative case and hence provokes the default agreement (i.e. third person singular) with the verb as in (42b):

(42) a. ***Se*** *contrataron los mejores profesores del país.* (Passive)
b. ***Se*** *contrató a los mejores profesores del país.* (Impersonal)
'The best teachers in the country were hired.' (Tremblay 2005:251)

Moreover, the internal 'animated' argument in the impersonal *se* construction, *los mejores profesores*, must be preceded by the preposition *a* which indicates that it is not the nominative subject and that as an object, it has a [+human] feature as in (42b). In the passive *se* construction in (42a), this argument is not preceded by the preposition *a*.

Using the structures in (42) as items of the experiment, Tremblay (2005:258) addressed two research questions with two corresponding hypotheses in her study. The first question asked if advanced L1 French and English L2 learners of Spanish have grammatical knowledge to differentiate passive and impersonal *se* constructions. If so, they would correctly relate the agreement on the verb with passives and lack thereof with impersonals, and they would employ the preposition *a* with impersonals and would not use it with passives. The second question was concerned with L2 learners' interlanguage (Selinker 1972) grammar. The underlying hypothesis of this question assumed French L2 learners of Spanish would face fewer obstacles than English L2 learners if the L1 influences the formation of interlanguage grammar. This is because French and Spanish are typologically more similar than English and Spanish.

A Grammaticality Judgment Task (henceforth, GJT), whose format was based on that of Bruhn de Garavito (1999a, 1999b), was conducted for the study. Two groups of university-level third-year Spanish students at the University of Ottawa participated in the experiment. The number of participants was 29 in total: 13 native speakers of English and 16 native speakers of French. Most of the participants had knowledge of French or English as a second language, hence Spanish was technically their third language although it was referred to as their L2 in the study. Proficiency in the L2 of each experimental group corresponded to the course level in which the participants were enrolled at the time when the experiment was conducted: all of them were classified as advanced learners. There was also a control group, which consisted of 27 native Spanish speakers from Valladolid, Spain, aged 19 to 30 years.

The results showed that the L2 learners had not (yet) acquired the structural and thematic properties underlying passive and impersonal *se* constructions, nor their internal arguments. These results contrast with the findings of studies that confirmed that adult learners of an L2 can

acquire subtle grammatical properties of the target language that are not present in their L1 nor are salient in the input. Such is the position of the Full Transfer Full Access Hypothesis (Schwartz & Sprouse 1994, 1996). One possible reason for this result mentioned by Tremblay (2005) is that the participants' L2 proficiency may not have been sufficiently advanced to acquire the structural and semantic properties involved in the use of passive and impersonal *se*, since their input was limited to classroom instruction. Particularly, it may be that both groups of learners do not associate the preposition *a* with non-nominative arguments as native speakers do. In one study on the acquisition of reverse psych-verbs (e.g. *gustar* 'to be pleasing/liked'), it was found that even the advanced L2 learner participants in the study accepted sentences with the missing preposition *a* to mark the non-nominative arguments, whereas the native speaker participants almost unanimously rejected them (Kanwit & Quesada 2018). The findings of the Tremblay study confirm that L2 acquisition by French- and English-speaking adults of passive and impersonal *se* constructions in Spanish reflects a learnability problem, not yet overcome even at an 'advanced' level of proficiency.

On the other hand, Bailey (2013) conducted a series of experiments, also from a generative point of view, to investigate the acquisition of the clitic *se* by L1 English speakers. Tremblay (2005) and Bailey (2013) both tested accessibility to UG. However, Bailey (2013) also took dispositional middle[22] *se* constructions into account along with passive and impersonal ones, and her findings suggest that L2 learners' access to UG is impaired in a way that the

[22] Bailey employs the term 'middle' to refer to dispositional middle constructions (e.g. "*Los pantalones **se** lavan con un detergente suave* 'The pants are washed with a mild detergent' (Bailey 2013:24)". In order to minimize confusion, I will henceforth specify them as 'dispositional middles'.

learners' interlanguage is partially similar to native speakers' grammar in some respects but failed to be fully native-like. In addition, Bailey (2013) based her analysis on the Processing Instruction (PI) model of VanPatten (2004), a pedagogical model which focuses on improving learners' processing strategies and analyzes their pedagogical influence in the classroom. There were four research questions that guided the study: (i) Does PI improve learners' accuracy of interpretation of the functions of the clitic *se*?; (ii) Does PI have an effect on learners' accuracy in their production of *se*?; (iii) Does PI have an effect on the internal grammars of the L2 learners as evidenced by performance on GJT?; and (iv) Assuming that L2 acquisition is UG constrained, what are the L2 learners' grammar representations of the clitic *se* at this early stage of their L2 Spanish acquisition as evidenced in the adverb placement task?

The author set forth hypotheses for each function of *se* that she investigated. There were two hypotheses for dispositional middle *se* constructions, one for impersonal *se*, and one for passive *se*: (i) Learners will reset their L1 parameter from null to overt morphology in dispositional middle constructions; (ii) Learners will not rely solely on the presence of the modifying adverb (e.g. *fácilmente* 'easily', *con facilidad* 'with ease') in order to arrive at a dispositional middle voice interpretation; (iii) Learners will reset their L1 parameter from overt subject to non-overt subject in impersonal constructions; and (iv) Learners will move the grammatical subject outside VP in passive constructions (i.e. the notional object rises to the specifier of TP).

A total of 63 participants were randomly divided into two experimental groups: 31 participants were assigned to the PI treatment group and 32 participants to the control group, that is, the group without PI treatment. The PI treatment group received explicit grammatical explanation with metalinguistic information of structured input that focused on processing and

interpreting the clitic *se* (e.g. *se* should not be interpreted as the subject of the sentence, Spanish subject nouns can be in preverbal and postverbal position).

To test the effectiveness of PI, both groups completed a pretest and posttest which consisted of interpretation tasks (identifying and differentiating between the agent and the sentence superficial subject), production tasks (a sentence completion task that asked the participants to insert the clitic *se* in the given sentences and a translation task that asked them to translate English sentences into Spanish), a GJT and an adverb placement task. Some examples of each type of task from the study are provided below:

(43) Interpretation task:

Read the following sentences carefully and identify the function of pronoun "*se*" in each of the sentences below: ***impersonal, middle voice or passive***. If you think the "*se*" has none of these three functions write ***other*** and if you are not sure of its function write ***don't know***. In addition, identify the subject of the sentence: If it is an implied subject write *implied*, if you cannot identify the subject write *don't know*. Please do not leave any spaces blank.

Sentences	**"se" function**	**Subject**
Se puede mirar por los huecos de la pared		
Las casas prefabricadas se construyen fácilmente		

(44) Production task:

a. Sentence completion task: Read the following paragraph and fill in the blanks with the pronoun *se* wherever you think it is needed. Leave the space blank if you think no *se* is necessary.

Mi vecino y su casa embrujada

Marcos es un chico extraño. Marcos (1) ___ lava a su perro con cera. Pero él también hace cosas que no son muy comunes. Él solo (2) ___ lee libros hechos en hoja de banano porque él dice que estos libros (3) ___ leen mejor, tiene camisas de algodón porque no (4) ___ arrugan y no tiene familia porque la familia no (5) ___ escoge. [...]

b. Translation task: Translate the following sentences into Spanish. *Translate as many as you can by APPROPRIATELY using the pronoun "se", keeping in mind that "se" is not appropriate in all sentences.* If you do use the pronoun "*se*", please identify its function: ***passive, middle voice or impersonal*** or if you are not sure of its function write ***don't know***.

a) The walls were painted yesterday (paint 'pintar')

_____

b) John walks to school (walk 'caminar')

_____

(45) GJT:

Please read the sentences and rate them on the following scale:

0 = ungrammatical, 1 = somewhat grammatical, 2 = grammatical and 3 = do not know

Sentences	**0**	**1**	**2**	**3**
Las manchas de tomate limpian difícilmente				
Los bosques se quemaron con gasolina				

(46) Adverb placement task:

Look at the adverb in parentheses at the right of each sentence. Mark with an X all the possible places where you could insert the adverb in the sentence.

d) ___ las blusas de seda ___ se ___ lavan ___ con cuidado ___ (no)

e) ___ se ___ necesita ___ a esa profesora ___ (supuestamente)

(Bailey 2013:240-246)

The quantitative analysis of the data showed that the PI helped participants interpret the dispositional middle constructions but not the passive and impersonal ones. Although the author mentions that the participants may have used other strategies to interpret the dispositional middle function correctly, the participants with PI treatment improved in identifying the subject of dispositional middle constructions correctly in their posttests. With regard to production, the group with the PI treatment correctly used the clitic *se* in dispositional middle and passive

constructions but not in impersonal ones. However, there was a difference according to the type of tasks: sentence completion task results showed that having all three functions together in one task can be too cognitively demanding for the participants.

A hierarchy of difficulty was found from the experimental tasks: dispositional middle < passive < impersonal. This hierarchy can be explained based on the parameters of the L1, English. Participants preferred preverbal to postverbal subjects and used an explicit NP in impersonal constructions to fill the subject position in the sentence, as can be seen in the following example:

(47) **Alguien* ***se*** *asustó a los niños.* 'Someone scared the children.' (Bailey 2013:182)

Upon diagnosing that the participants' failure may be due to insufficient exposure to L2 input, Bailey (2013) states that it seems necessary for the learners to be exposed to more input and presented with one function of the clitic *se* at a time because the syntactic similarities between the functions make it difficult to differentiate one from another. That is, if a form has several functions, it is essential to target only one of them at a time to help learners interpret and process that specific function better. This, of course, is in contrast to the cognitive model adopted in this dissertation and as set forth by Maldonado (2008, among others), who suggests that learners should be made aware of the general event-focusing nature of *se* and then shown how it plays out across several functions of the structure. Bailey particularly emphasizes that explicit instruction is necessary for learners to reset their L1 parameter successfully. In conclusion, unlike Tremblay (2005), Bailey (2013) claims that the learners' interlanguage is the result of input properties of the L2. In other words, when given appropriate input, learners would successfully acquire native-like properties of the target language. This supports the assumption that L2 learners have access to UG, although this access could be impaired due to a lack of sufficient L2 input.

A hierarchy of difficulty different from that of Bailey (2013) was provided by Escobar & Teomiro (2016). As mentioned above, Bailey (2013) examined L2 learners' acquisitional pattern of dispositional middle, passive and impersonal constructions. In contrast, Escobar & Teomiro (2016) tested the following constructions:

(48) a. Anticausative/Inchoative: *La ventana **se** cerró.* 'The window closed.'
b. Inherent Reflexive: *La niña **se** peinó.* 'The girl combed (herself).'
c. Inalienable Possession: *Juan **se** lavó las manos.* 'Juan washed his hands.'
d. Consumption: *El niño **(se)** comió el sándwich.* 'The boy ate the sandwich.'
e. Non-anticausative Inchoative: *El anciano **(se)** cayó.* 'The elder fell.'

(Escobar & Teomiro 2016:18)

The authors group clitic *se* constructions based on the grammatical nature of the clitic *se*. In anticausative/inchoative and inherent reflexive types, the clitic is an expletive that does not introduce any argument or any further syntactic complexity. On the other hand, in the other three constructions, the clitic is an argument in the specifier of the phrase projected from a low applicative head (2016:18).

Since the expletive *se* in anticausative/inchoative and inherent reflexive types shows a similar syntactic configuration with that of English, the authors assume that the first two constructions will be acquired easily and early on. However, although the remaining three functions share the status of *se* as an argument in [Spec, Appl], the inalienable possession use of *se* will be easier to acquire than the other two because the low applicative head of English is very similar to this construction. Due to a somewhat different syntactic structure and semantic interpretation, consumption verb types will be somewhat difficult to acquire. Finally, non-anticausative intransitive constructions will be acquired last because this type of syntactic configuration does not exist in English.

The hypothetical hierarchy of difficulty proposed by Escobar & Teomiro (2016) is the following: Anticausative/Inchoative & Inherent reflexive < Inalienable possession < Consumption < Non-anticausative intransitive. Based on this hierarchy, an experimental study with 55 English-speaking university students of Spanish was conducted. A GJT with 41 items revealed that the presumed hierarchy did not reflect the L2 learners' acquisitional process. Although there was not a definite pattern on the performance of the learners, the authors concluded that "L1 transfer errors were obtained only at the early stages whereas non L1 transfer errors were found at all levels, indicating that the learners are making use of their implicit knowledge, which is gradually developing over time in both groups" (2016:27). This conclusion is based on the result that the overall performance did improve when low proficiency learners and intermediate level students were compared. Also, it was claimed that L2 learners have access to UG and can gradually acquire complex grammar structures not present in their L1.

In two related investigations, Gómez Soler (2015a, 2015b) also attempted to account for acquisitional patterns of Spanish argument structure and inchoative *se* based on the generative framework. There were two research questions raised in her first study, but only the second question, which is related to the antipassive *se*, will be discussed here. In particular, she asked how well L2 learners distinguish two separate types of psych-verbs that differ in compatibility with clitic *se*. In (49a), the transitive verb *asustar* 'to scare' can be 'intransitivized' with *se*, whereas in (49b), the intransitive verb *gustar* 'to be pleasing/liked' cannot:

(49) a. *Carolina **se** asustó.* 'Carolina got scared.'
b. **Carolina **se** gustó.* 'Carolina got liked.' (Gómez Soler 2015a:632)

A total of 65 L1 English-L2 Spanish speakers participated in two GJTs, the second of which tested acceptability of *se* marked psych-verbs. The results of the second experiment showed that, although the L2 learners' judgments were in accordance with that of the native

speakers, they also contrast with native performance when the ratings for acceptance with ungrammatical constructions with *se* are taken into account. Gómez Soler (2015a) considers this non-target-like behavior is derived from overgeneralization because reflexive/reciprocal *se* is compatible with both psych-verb types (50), whereas antipassive *se* is grammatical with only one of them (51):

(50) a. *María **se** enfadó (consigo misma)*. 'María got angry at herself.'
b. *María **se** encanta (a sí misma)*. 'María loves herself.' (Gómez Soler 2015a:640)

(51) a. *María **se** enfadó.* 'María got angry.'
b. **María **se** encantó.* 'María got pleased.'

Thus, it is possible to assume that L2 learners, especially those with lower proficiency, may encounter problems with the multiplicity of meanings and functions of *se* which "might have blurred" (2015a:640) their judgments. However, the results of the high-proficiency groups showed that these difficulties seem to diminish as L2 competence develops, suggesting that non-native speakers also have access to UG.

Gómez Soler's (2015b) subsequent research measured the use of causative/inchoative *se* among a group of near native speakers of Spanish. In this study, she specifically tested whether highly advanced L2 learners are aware of the ungrammaticality of the presence of *se* with certain types of intransitive verbs:

(52) a. **Mi hermana **se** llegó.* 'My sister arrived.'
b. **María **se** caminó.* 'María walked.' (Gómez Soler 2015b:351)

Using *se* correctly (in other words, not using *se*) with these types of verbs poses a learnability problem since *se* can only co-occur with change-of-state unaccusatives, according to the author. Also, it can be predicted that due to L1 transfer, non-native speakers might underuse *se* in inchoative constructions where its use is obligatory (e.g. "**El jarrón rompió* 'The vase broke'"

(2015b:363)). In addition, the complexity of the multiple functions of *se* complicates the learning process even further.

A GJT was conducted with 65 English-speaking college students whose L2 is Spanish. The results revealed that the learners with lower proficiency overgeneralized unaccusatives over unergatives, which means that they employed *se* as a marker of intransitivity or, more specifically, non-agentivity. Additionally, as was predicted, omission of *se* was highly accepted by less competent learners in inchoative sentences, which are ungrammatical without *se*. Based on the finding that non-natives' judgments comply with those of native speakers in general, and that they seem to recover from the over- and under-generalization gradually, Gómez Soler (2015b) again supports full access to UG on the part of L2 learners.

In sum, the aforementioned generative studies aimed to explore the availability of non-native speakers' access to UG mostly by testing their grammaticality judgments on various Spanish clitic *se* constructions. Researchers conclude that L2 learners have access to UG, although the degree of access differed from study to study depending on their individual task results. Some studies support the assumption that there is full access to UG based on their findings that L2 learners gradually overcome learnability problems and perform native-like in their ultimate attainment stage of acquisitional process. Meanwhile, other studies report that even the most proficient learners still have not acquired some grammatical properties, especially those that are different from their L1s, suggesting that L2 learners' access to UG is available but impaired.

3.3.2 Usage-based studies based on corpora and other production tasks

Guillén Solano (2015) summarized three different L2 studies on Spanish *se* acquisition in the literature, which employed an error analysis methodology by using data collected from learner corpora. I provide a brief summary of these three analyses. First, Sanromán Vilas (2009) studied a corpus that consisted of written academic texts by 45 native speakers of Finnish. It was found that only 11% of the texts were free of errors in constructions using *se*. There were two fundamental types of errors committed by the learners: the absence of *se* with verbs that need the use of it (e.g. omission of *se* in *quejándo**se*** 'complaining' in "**las mujeres lloraban amargamente quejando por la ingratitud del muchacho* 'the women cried bitterly complaining about the boy's ingratitude'" (Sanromán Vilas 2009:963, [translation mine]), and the misuse of *se* when forming impersonal constructions (e.g. "***Se** busca camarero* 'Waiter wanted'" (Sanromán Vilas 2009:956, [translation mine]) to express a situation in which someone is looking for a waiter). Based on the observation that these types of errors persist over the years, Sanromán Vilas claims that it is necessary to use or adapt materials that might help Finnish speakers to learn Spanish *se* constructions more efficiently.

Secondly, a corpus of 49 written texts by seven English-speaking adult L2 learners of advanced level of Spanish who studied Spanish in classroom settings was analyzed by Escutia López (2010), focusing on unaccusative constructions. The most frequent error was that the non-native speakers inserted a semantically empty pronoun *se* before the verb in sentences where it would not be acceptable in Spanish (e.g. "*... *porque **se** va a haber huelgas* '… because there will be strikes'" (Escutia López 2010:133, [translation mine]). The author attributes the reason for this error to the transfer from the learners' L1, English. In English unaccusative constructions, either the internal argument rises to the subject position or the expletive pronoun

such as *there* is placed in that position. The advanced L2 learners of this study seem to have acquired the presence of null subjects in Spanish and do not tend to overgenerate pronominal subjects unnecessarily in unaccusative constructions. However, they seem to need a lexical item to fill in the subject position and resort to the clitic *se*. This result not only reveals some patterns of L1 transfer, but also those of interlanguage, which represents certain grammatical structures that are absent in both the learners' L1 and L2. Based on the findings, Escutia López (2010) reinforces that L2 learning does not simply consist of relexification or reestablishment of L1 structures with L2 words.

Finally, Araya & Monteserin (2010) analyzed 200 written texts by students of different levels of a Spanish proficiency exam in Argentina, CELU (*Certificado de Español: Lengua y Uso* 'Certificate of Spanish: Language and Use'), and demonstrated that the most frequently produced type of *se* constructions was ergative (e.g. "*preocuparse* 'to worry', *enojarse* 'to get angry'" (2010:6)), followed by impersonal (e.g. "*tratarse* 'to be about', *basarse* 'to be based on'" (2010:6)), and reflexive (e.g. "*defenderse* 'to defend oneself', *inscribirse* 'to register'" (2010:6)). Although L2 learners committed some errors including omission of *se* in obligatory contexts, incorrect use of *se,* erroneous use of *se* where it is not required and wrong word order, the authors focused on the findings that 72 out of 200 written texts used *se* in native-like manner 100% of the time. In addition to the L2 learners' data, an analysis of various *se* constructions obtained in the CHILDES corpus (MacWhinney 2000; MacWhinney & Snow 1985, cited in Araya & Monteserin 2010) for Spanish L1 was conducted in order to compare the acquisitional process between L1 and L2. Based on the results that the use of *se* is productive from the beginning, at a very young age for native speakers and from the basic level for non-native

speakers, the authors claimed that the L2 acquisition of the clitic *se* presents the same characteristics as its acquisitional process in L1.

Guillén Solano (2015) investigated L2 learners' and native speakers' use of the clitic *se* as a strategy for expressing anticausativity in sentences such as *Se me quebró el plato* 'The plate broke on me' (vs. *Quebré el plato* 'I broke the plate'). The participants of the experiment were six English-speaking learners of L2 Spanish. Three of the participants had a formal education in Spanish at the university level and three had no formal education but had lived in Costa Rica for at least three years at the time of the study. Therefore, any difference in the results between the two groups in this study not only shows how L2 learners use the clitic *se* to express anticausativity, but also how the learning environment (i.e. formal *vs.* informal) can affect acquisition.

The design of the study was a guided oral interview. Based on a series of images, the participants answered the question "*¿Qué le pasó a…?* 'What happened to…?'" (2015:93). The results showed that the participants were aware of the passive function of *se*, whereas its anticausative property was used less by the learners than by the native speakers. In terms of the contrast between the two groups of learners, there was no significant difference between those who learned Spanish in classroom settings and those who learned it in a natural context. L2 learners, in general, demonstrated a pattern in which the use of *se* was acquired in a later stage. Also, they tended to use other forms that were more familiar to them, such as transitive constructions, instead of those with *se* (e.g. "*No sabia dónde iba* 'S/he didn't know where s/he was going.'" (2015:95, [translation mine]) rather than "***Se** perdió* 'S/he got lost'" (2015:97, [translation mine]), which was preferred by native speakers of Costa Rican Spanish). Based on these results, Guillén Solano (2015) claims that classroom materials must represent authentic

language used by native speakers and make decisions about which structures to emphasize and include in Spanish L2 instruction.

Escutia (2016) also looked at L1 English-L2 Spanish learner oral production data. The peculiarity of this research is twofold: (i) it qualitatively analyzed data from one 74-year-old advanced learner of L2 Spanish who had been exposed to Spanish at home and work and had lived in Spain for over a decade and even taught in Spanish; (ii) in addition to analyzing oral data, an acceptability judgment task was conducted with the participant's own utterances in order to observe L2 interpretation. The grammatical phenomenon in question was how unaccusativity is expressed in Spanish and how different it is from the way English marks unaccusative structures. In English, expletive pronouns such as *there* are employed to fill in the subject position in certain unaccusative examples (e.g. "**There** exist many possibilities" (2016:35)). Unlike English, using *se* as an expletive pronoun in Spanish results in an ungrammatical sentence (e.g. ***Se existen muchas posibilidades***).

Upon examining the data extracted from the participant's Religious Studies classes, which he taught in Spanish, non-target-like patterns, which could be due to the properties of his L1, were found. First of all, clitic *se* was overgeneralized and used to fill the subject position, like the English expletive *there*. Also, in some cases such as "*A María no le gustaría que* ****se*** *naciese con tanta pobreza* 'María would not like him to be born in the midst of so much poverty'" (2016:43), it is likely that *se* functions as a 3rd person agreement marker in this participant's interlanguage. Some overextensions of the pronominal verb status to verbs which are not pronominal are also reported:

(53) ****Nos*** *tenemos que estar haciendo lo previsto.*
'We've got to be doing what we're supposed to.' (Escutia 2016:44)

To summarize, Escutia (2016) generalized his findings by mentioning that "the clitic *se* seems to be particularly favored by the presence of unaccusative verbs, apparently acting as an expletive of sorts" (2016:46). The interpretation data in which the participant corrected his own utterances also demonstrated that his interlanguage related to clitic *se* was still inconsistent, albeit near native-like in the target language.

3.3.3 A usage-based study from a cognitive grammar perspective

Elliott's (1995) study is also based on production data, but differs from the aforementioned studies in that, like the present study, it adopts the cognitive grammar theoretical framework, arguing that generative research and SLA research are not mutually beneficial. According to the author, since generative analyses are motivated by theory-internal concerns and are too self-absorbed with concerns of the generative paradigm, they often fail to fully understand second language data. Based on previous literature from the cognitive principle (cf. 2.4.2) such as work by Langacker (1991), Maldonado (1988, 1992) and Smith (1993), Elliott (1995) claims that his approach is more desirable than a generative one because:

> It attempts to unify and relate the various functions of *se*, which finds support in the works of other researchers. […] It suggests that a unifying approach to *se* also makes more sense pedagogically-speaking. (Elliott 1995:64)

Moreover, since cognitive grammar is a usage-based theory and attempts to relate language use and language learning with cognitive bases, he supports the idea that it is beneficial to use a cognitive approach in interpreting L2 *se* data.

With the objective of presenting how and why the learning of *se* has been problematic, Elliott (1995) introduces a set of rules that L2 learners of Spanish, especially L1 English speakers, should be aware of:

(54) a. Use the reflexive marker *se* with animate subjects like you would use the reflexive in English (e.g. ***Me*** *veo* 'I see myself').

b. Use the reflexive marker *se* with inanimate subjects, unlike how you would use the reflexive in English (e.g. *No abrí la puerta;* ***se*** *abrió* 'I didn't open the door; the door just opened').

c. Use *se* to indicate that no outside agent is involved (e.g. ***Se*** *me olvidaron las llaves* 'The keys were forgotten on me').

d. Use *se* with impersonal constructions like you would use the indefinite pronoun 'one' in English (e.g. ***Se*** *habla español aquí* 'One speaks Spanish here').

e. Use *se* to indicate changes in body (e.g. ***Me*** *envejezco* 'I age/get older'), body position (e.g. ***Me*** *levanto* 'I get up'), position in space (e.g. ***Me*** *voy* 'I'm leaving/moving'), and emotional state (e.g. ***Me*** *vuelvo loco* 'I'm going crazy'), where it is frequently implied that the subject takes responsibility for the changes.

f. Use the pragmatic context to interpret the specific meaning of *se* (e.g. ***Se*** *lava la ropa* 'Clothes are being washed' or 'S/he is washing her/himself (some) clothes', depending on context). (Elliot 1995:16-27)

Two research questions guided the study: "How do university learners of Spanish at the intermediate level of proficiency use the polyfunctional morpheme *se*?; What factors appear to constrain their use of *se*?" (Elliot 1995:93). Written conversational texts of 86 second-year university students of Spanish[23] were collected and analyzed for the research. The data examined in the study, computer conferencing messages, were expected to reflect characteristics of naturalistic and authentic language data with a conversational tone.

A VARBRUL analysis of the data revealed that L2 learners clearly concentrated on the first person and third person uses of *se* and that they performed better with constructions where the subject's position was postverbal than preverbal. However, what is of greater importance is

[23] Most of the participants had 2-3 years of previous high school Spanish instruction.

that the learners showed a pattern of overgeneralization as well as undergeneralization with the use of *se*: sometimes they used *se* in contexts where it was not required (e.g. "***Me** amo el campo* 'I love the countryside'" (Elliot 1995:238)), while in other instances they omitted *se* in obligatory contexts (e.g. "**Graduo* 'I graduate'" (Elliot 1995:240)). Elliott (1995) attributed this result to the difference in the way that speakers construe a given event. He points out:

> Non-native construal may be dependent in part on the learner's lack of understanding of the different ways available in the second language for construing an event, as well as in part of the first language (L1) and the means available in the L1 for construing an event. (Elliott 1995:242)

Accordingly, it can be assumed that both the L1 and the L2 systems affect students' interlanguage, based on the fact that "these miscontruals demonstrate the students' awareness of frequent construction types in Spanish and their ability to extend these types to possible, yet unacceptable, Spanish constructions" (Elliot 1995:179).

In sum, from the perspective of cognitive grammar, Elliott (1995) examined the L2 acquisition of Spanish clitic *se* constructions. This study is relevant to the present dissertation in that it conducted a variable rule analysis (i.e. VARBRUL analysis in Elliott (1995) and Rbrul (Johnson 2009) analysis) to explore which factors affect L2 learners' use of *se* with an approach based on cognitive grammar. The VARBRUL analysis revealed that L2 learners were affected by grammatical person and subject position. Although the present study will not include grammatical person as a factor group in the analysis due to the nature of the picture book narration (i.e. third person forms are used almost exclusively because participants narrate a story of a boy, a dog, and a frog), it is of interest to see if subject position will show the same effect in both studies.

3.3.4 Conclusions of L2 studies

A brief summary of L2 studies described so far is provided in Table 9 below. The studies in the table are categorized based on their similarities: error analysis/corpus studies (Araya & Monteserin 2011; Escutia López 2010; Sanromán Vilas 2009), generative studies (in favor of full access to UG: Bailey 2013; Escobar & Teomiro 2016; Gómez Soler 2015a, 2015b *vs.* against full access to UG: Tremblay 2005), and usage-based studies (Elliott 1995; Escutia 2016; Guillén-Solano 2015).

Table 9. L2 studies of the acquisition of Spanish clitic *se*

Researcher (Year)	**L1 of participants**	**Number of participants**	**Type of data**	**Result(s)**
Araya & Monteserin (2011)	-	200 written texts	Written texts	L2 learners acquire productive use of *se* from the beginning level.
Escutia López (2010)	English	7	Written texts	L2 learners produce lots of errors with sentences containing *se*.
Sanromán Vilas (2009)	Finnish	45	Written texts	
Bailey (2013)	English	63	Sentence completion; Translation; GJT	Hierarchy of difficulty: Dispositional middle < Passive < Impersonal; L2 learners have access to UG, when given appropriate input.

Escobar & Teomiro (2016)	English	55	GJT	Hypothesized hierarchy of difficulty not proven; L2 learners have access to UG.
Gómez Soler (2015a, 2015b)	English	65; 65, respectively	GJT	Over- and under-generalization due to multifunctionality of *se* and L1 transfer; L2 learners have access to UG.
Tremblay (2005)	English; French	16; 13, respectively	GJT	Even advanced learners have not (yet) acquired grammatical properties; L2 learners does not have full access to UG.
Elliott (1995)	English	86	Computer-written conferencing messages	L2 learners over- and under-generalize the use of *se*, depending on the contexts.
Escutia (2016)	English	1	Oral production data; Acceptability Judgment Task	Overgeneralization pattern is found; Learners show certain fossilized errors; Clitic *se* constructions are late-acquired, if ever.
Guillén-Solano (2015)	English	6	Guided oral interview data	Passive is acquired before anticausative; *se* is late-acquired, in general.

Although a series of studies on L2 acquisition of *se* have shown that the syntactic and semantic complexity and L1/L2 structural differences hinder successful acquisition by L2 learners, some questions remain unanswered: Can the use of *se* be measured as correct or incorrect when it is varied even among native speakers, except in some cases where researchers give participants a limited set of 'manipulated' sentences including specific grammatical phenomena such as subject-verb agreement, word order, etc.?; How can we generalize the acquisition process of the multifunctional clitic *se* when there is no research investigating all or at least most of the constructions with the use of *se*?; Is there any way to relate findings from different studies (e.g. hierarchies of difficulty from Bailey (2013) and Escobar & Teomiro (2016)) to build more concrete hypotheses that could contribute to the field of SLA? Since Tremblay (2005) and Bailey (2013) attributed L2 learners' ungrammatical behavior to their insufficient level of proficiency and exposure to input, how different will the results be when near-native speakers are included in the experiments?

In order to analyze interpretation and production of a wider range of constructions with *se* by L2 learners, the middle domain proposed by Maldonado (1992 and following), discussed in Chapter 2, will be the basic framework of research for the present study, as it was for Elliott (1995). In particular, by determining the use of *se* as a middle marker, all non-reflexive uses of *se* are generalized into one single category, which minimizes the debate about the type of *se* used in various contexts. Since some research on L1 acquisition provided results that support the assumption that Spanish non-reflexive *se* functions as a middle marker, and that all constructions marked by this *se* belong to the Spanish middle domain, it is of special interest to test if L2 acquisition research shows similar results.

Also, it is claimed here that traditional and pedagogical approaches have been erroneously presenting all instances of *se* as reflexive constructions, hence making it even more difficult for L2 learners to understand that there is no one-to-one correspondence between *se* and a reflexive function. If so, assuming that L2 learners with a lower proficiency level would have more difficulties with middle *se* constructions seems plausible. Therefore, it seems necessary to compare learner groups across different levels of proficiency, which could also shed light on the applicability of the cognitive approach to Spanish middle marker *se*.

In summary, there is clearly a need for L2 acquisition studies of *se* from the perspective of cognitive grammar for various reasons. Cognitive grammar, as a usage-based linguistic model, highlights the viewpoint of the conceptualizer, which is related to the usage of middle marker *se* to a great extent. For second language speakers, the use of *se* reflects their perception of how events are construed and expressed. It also reflects how 'subtle' but meaningful concepts such as dynamicity or energetic-ness are conveyed in Spanish, which is not easily accounted for in GJTs. Hence, examining L2 *se* data from a cognitive perspective is expected to provide important insight into understanding language development by non-native speakers. The cognitive perspective will also be useful for designing an L2 acquisition model, assisted in part by previous L1 acquisition studies which share the same theoretical framework.

The following section analyzes textbooks used in Spanish language courses at the University of Georgia (UGA) and Seoul National University (SNU), where the majority of the participants have attended or are currently attending, in order to get a glimpse of how the clitic *se* is introduced and explained in L2 Spanish classrooms. Textbooks for elementary level classes will demonstrate the way in which the clitic *se* is presented to the students for the first time. The purpose of examining intermediate-level textbooks is to explore how comprehensive the

explanations are, i.e. how many types of *se* constructions are introduced and what kind of explanations are given to guide students' understanding of the multifunctionality of the clitic *se*.

3.4 Textbook analysis on Spanish *se* constructions

As almost all of the L2 learners (32 out of 33 native speakers of English and 28 out of 30 native speakers of Korean) who participated in the current study reported that they learned Spanish in a classroom setting, it is of interest to see how the clitic *se* is taught in language classrooms. One way to understand what kind of input is given to L2 learners is to examine the textbooks used in the classroom. Thus, analyzing how the clitic *se* is introduced to learners and in which way its explanation develops is expected to shed light on the acquisitional process of *se* constructions by L2 Spanish learners.

A total of five textbooks were analyzed: *Experience Spanish* and *Más*, which are designed for English speakers and *Español básico 1, Español básico 2* and *Español intermedio*, which are used by Korean speakers. The criteria for selecting these textbooks is based on the institutions where the majority of the participants were recruited. *Experience Spanish* and *Más* are utilized in elementary and intermediate level courses at the University of Georgia (UGA), respectively. On the other hand, at Seoul National University (SNU), textbooks of elementary level courses are *Español básico 1* and *2*, while *Español intermedio* is used in intermediate level courses. The following table summarizes each textbook's information:

Table 10. Textbooks analyzed for the present study

Title	Editorial	Year	Country	Related courses (Institution)
Experience Spanish	McGraw Hill	2020	US	Elementary Spanish 1 (UGA), Elementary Spanish 2 (UGA), Accelerated Elementary Spanish (UGA)
Más	McGraw Hill	2019	US	Intermediate Spanish 1 (UGA), Intermediate Spanish 2 (UGA)
Español básico 1	SNU Press	2013	South Korea	Elementary Spanish 1 (SNU)
Español básico 2	SNU Press	2013	South Korea	Elementary Spanish 2 (SNU)
Español intermedio	SNU Press	2020	South Korea	Intermediate Spanish 1 (SNU), Intermediate Spanish 2 (SNU)

In *Experience Spanish*, the clitic *se* is first introduced in Chapter 5 as a reflexive pronoun where the reflexive verbs are presented. It explains:

> Spanish has a special category of verbs called reflexive verbs, which are used when speakers talk about what they do to themselves or for themselves. A reflexive verb consists of two parts, a reflexive pronoun followed by a conjugated form of the verb. **The reflexive pronoun always refers to the subject of the verb, who performs the action on him/herself.**
>
> (*Experience Spanish,* p. 148, [emphasis mine])

Also, a list of verbs that carry the clitic *se* is listed as follows:

> **Most verbs about personal-care routine are reflexive in Spanish:** *acostarse* 'to lie down; to go to bed', *afeitarse* 'to shave', *despertarse* 'to wake up', […].
> **There are many other verbs in Spanish that can be used reflexively and that you will learn later in this book. For now, the following are some reflexive verbs not related to personal care that you see in this chapter:** *divertirse* 'to have a good time, have fun', *mudarse* 'to move (from one residence to another)', *relajarse* 'to relax', *sentirse* 'to feel'
>
> (*Experience Spanish*, pp. 148-149, [emphasis mine])

In a section called *Nota comunicativa*, a brief explanation and some examples of the reciprocal uses of *se* are presented. Later, Impersonal and Passive *se* are covered in Chapter 7 of the textbook. It explains that "this structure communicates that the action happens in general, affecting everyone equally" (*Experience Spanish,* p. 221), not distinguishing the two constructions in question. The reason for not presenting them separately is given in the footnote: "Although the impersonal *se* and passive *se* are two different structures, they are very similar. The focus of these guidelines is not to distinguish the two, but rather to help you use both structures correctly" (*Experience Spanish,* p. 221).

Also, a brief introduction to "*se* for unplanned occurrences" (*Experience Spanish,* p. 222) is offered in the *Nota comunicativa* section. Here, it says that "another use of the pronoun *se* is in a special construction that expresses accidental or unplanned events, casting the person or persons affected by the event as victims, not responsible for the occurrence" (*Experience Spanish,* p. 222). This explanation is followed by a list of verbs commonly used in this construction (e.g. *acabar* 'to run out of', *caer* 'to drop', *descomponer* 'to break down', *olvidar* 'to forget', *perder* 'to lose', and *romper* 'to break').

In summary, an analysis of the textbook *Experience Spanish* demonstrated that native speakers of English in beginner level Spanish classes receive information on the clitic *se* in the

following order: (True) Reflexive > Reciprocal > Impersonal and Passive > Unplanned occurrences. One thing to note here is that Impersonal and Passive constructions are undistinguished because of their similarities, according to the textbook and "true reflexives" are not always reflexives (i.e. they don't always allow for the addition of the *a sí mismo* phrase).

Another textbook for English speakers, *Más*, is expected to have a more detailed explanation on various *se* constructions, since it is designed for intermediate-level courses. In Chapter 2, *Más* first introduces *se* in a grammar section where reflexive verbs and pronouns are presented. Since the basic explanation does not differ much from *Experience Spanish*, only noteworthy content will be mentioned here.

Firstly, the list of other verbs, apart from daily routine verbs that are 'reflexive' in Spanish, is longer than that of the elementary level textbook. This list includes: *callar(se)* 'to be quiet', *calmar(se)* 'to calm (oneself) down', *divertir(se)* 'to have fun', *enamorar(se) de* 'to fall in love with', *morir(se)* 'to die', *preparar(se)* 'to prepare (oneself)', *reunir(se)* 'to get together; meet', *sentar(se)* 'to sit (oneself down)', *sentir(se)* 'to feel', of course, not all of which are true reflexives.

Next, there are two additional types of verbs given in this textbook: (i) verbs that change meaning, and (ii) verbs of *becoming*. According to the textbook, these verbs are not necessarily reflexive or have no reflexive meaning in English. The following table shows some examples:

Table 11. Examples of verbs that change meaning and verbs of *becoming* (Adapted from *Más,* p. 56)

Verbs that change meaning			
acordar	to agree	*acordarse*	to remember
beber	to drink	*beberse*	to drink up
dormir	to sleep	*dormirse*	to fall asleep
Verbs of *becoming*			
aburrir → aburrirse		to get/become bored	
alegrar → alegrarse		to get/become happy	
enfadar → enfadarse		to get/become angry	

Finally, in a section called *Nota lingüística*, which seems to play a similar role as *Nota comunicativa* in *Experience Spanish*, the use of *se* with certain verbs that express change is introduced. Here, it says that "in English, oftentimes these changes are expressed with verbs such as *to become, to turn into, to get, to go,* and so on" (*Más*, p. 57). Uses of four particular verbs with *se* is presented in the textbook as follows:

(55) a. *convertirse en/al* + noun → conversion or metamorphosis
b. *hacerse* + adjective/noun → gradual change, implying conscious effort and/or a goal met
c. *ponerse* + adjective → sudden physical or emotional change
d. *volverse* + adjective/noun → physical or emotional change, often sudden, dramatic, and irreversible (*Más*, p. 57)

To sum up, Chapter 2 of *Más* introduced the clitic *se* as a reflexive pronoun, as it was characterized also in *Experience Spanish*. However, additional verbs beyond the category of 'daily routine' and verbs belonging to three new types (i.e. verbs that change meaning, verbs of *becoming*, verbs that express change) are presented. Although the textbook mentions that the last

three types are not strictly related to reflexive meaning in English, students could still consider them reflexive constructions because they are explained in the chapter titled *Los verbos y pronombres reflexivos* 'Reflexive Verbs and Pronouns'.

Later, in Chapter 4, accidental *se* appears in the textbook. This construction type was included in the *Nota comunicativa* section in *Experience Spanish*. Now, in *Más*, it is covered in a separate chapter with a more detailed explanation. The textbook says, "the accidental **-se** construction is grammatically a reflexive action: it appears as if the object of the action does something to itself (e.g. "***Se cayó** la leche* 'The milk fell. (not known how)'" (*Más,* p. 113)). Again, the correlation between the clitic *se* and "reflexivity" is demonstrated in the textbook.

In Chapter 6, impersonal *se* is presented as another *se* construction. Different from *Experience Spanish* where impersonal and passive *se* were covered in one chapter, *Más* presents impersonal *se* separately. Interestingly, one of the explanations compares impersonal *se* constructions to English counterparts: "[…] this construction is often translated as the passive voice in English" (*Más*, p, 177).

Finally, in the last chapter of the textbook, Chapter 12, passive *se* is mentioned as a structure which belongs to the passive voice. Here, the textbook says that "the passive voice with **ser** + *past participle* is used less frequently in Spanish than it is in English […] *Se* construction is much more commonly used, whether or not the agent is known" (*Más,* p. 341). The *Nota lingüística* section following this explanation provides a summary of the uses of *se* covered in the textbook. As can be seen in table below, *Más* actually does not differentiate impersonal and passive *se*, although the impersonal construction was presented in a separate chapter.

Table 12. Summary of the uses of *se* (Adapted from *Más*, p. 342, [translation mine])

Paradigmatic *se*: object pronoun	
Reflexive verbs (Ch. 2)	*Julio no **se** acostó hasta las 11:00.* 'Julio didn't go to bed until 11:00.'
Reciprocal verbs (Ch. 2)	*Tú y yo **nos** vemos tanto como **se** ven José y María.* 'You and I see each other as much as José and María see each other.'
"False" *se* (Ch. 2)	***Se** lo di esta mañana.* 'I gave it to him this morning.'
Nonparadigmatic *se*: substitutes of passive voice	
Impersonal/Passive (Ch. 6)	***Se** habla español.* 'Spanish is spoken.'
Accidental (Ch. 4)	***Se** me perdió la cartera.* 'I lost my wallet.'

To summarize, both English textbooks —*Experience Spanish* and *Más*— introduce clitic *se* to the students as a reflexive pronoun, followed by its impersonal/passive use and accidental *se* constructions. As mentioned before repeatedly, impersonal and passive structures are not distinguished and are equated to English passive voice in that they do not focus on the agent of the sentence. Also, construction types, which would be considered middle construction types in the cognitive perspective, were introduced in the chapter as "reflexive" verbs without any helpful explanation.

On the other hand, in Korean textbooks used in elementary level courses, the clitic *se* is first introduced in Chapter 2 of *Español básico 2*. This means that students who take the Spanish course (i.e. Elementary Spanish 1) do not receive explicit instruction on the clitic *se* during the first semester. It is in the second chapter of the second semester that the explanation about reflexive verbs and reflexive pronouns is given. Specifically, the textbook says that the reflexive pronoun attaches to transitive verbs to make reflexive verbs. Four categories of constructions are provided:

(56) a. Reflexive pronoun as direct object: When a reflexive pronoun is attached to a transitive verb, it becomes a reflexive verb which functions as an intransitive verb (e.g. *levantar* 'to lift' *vs. levantar**se*** 'to wake up')

b. Reflexive pronoun as indirect object: The direct object always comes with an article (e.g. *lavar**se** las manos* 'to wash one's hands')

c. Fixed expressions: *acordar**se** de* 'to remember (something)', *dedicar**se** a* 'to do (something) for a living', *despedir**se** de* 'to say goodbye to (someone)', *encontrar**se** con* 'to meet with (someone)'

d. Reciprocal *se*: Reflexive verbs in which the subject is plural and has the meaning of 'to each other' (e.g. *Juan y Ana **se** aman locamente* 'Juan and Ana love each other madly') (*Español básico 2*, pp. 22-25, [translation mine])

As can be seen above, this textbook mentions the detransitivizing function of the reflexive pronoun *se* and differentiates when the reflexive pronoun is used as a direct object and when it plays the role of an indirect object. However, the explanation given in (56a) is contradictory. If *levantar**se*** is a reflexive verb as it is described above, its literal translation has to be 'to wake oneself up', which results in a transitive verb. Therefore, adopting the term 'detransitivizing *se*' seems more suitable than providing the contradiction that reflexive verb functions as an intransitive verb.

In the next chapter, Chapter 3, the passive *se* construction is introduced to the students. Following a description of the periphrastic passive (i.e. *ser* + past participle) construction, passive *se* is explained as follows: "In passive *se* constructions, the '*por* + agent' phrase tends not to appear. Passive *se* is widely used because it can also be used in present tense. It is usually used when the subject is non-human" (*Español básico 2*, p. 46, [translation mine]).

The last clitic *se* construction presented in *Español básico 2* is the impersonal *se* in Chapter 9. The textbook explains that the impersonal *se*, with its formula '*se* + third person singular verb', is used to express when 'people generally do something' without specifying a

subject. It is worth noting also that the textbook specifies that only the third person singular form of the verb can be used in impersonal *se* constructions. Compared to *Experience Spanish* used at UGA, which does not differentiate passive from impersonal *se*, *Español básico 2* appears to provide students with more metalinguistic explanations.

Español intermedio, the textbook used in intermediate level courses, focuses more on reading comprehension than explicit grammar instruction. Also, the first half of the textbook, which is covered in an 'Intermediate Spanish 1' course, mainly deals with the subjunctive mood. Therefore, no grammar descriptions on any clitic *se* constructions were given. In Chapter 13, the textbook describes various passive voice structures including passive and impersonal *se*. Here, only formulas and some examples are given without any written descriptions of the constructions:

(57) a. Passive *se*: *se* + Verb + Subject

***Se** han escrito unas 200 novelas sobre la guerra civil española.*

b. Impersonal *se*: *se* + le(as) + Verb

*No **se** te entiende cuando hablas con la boca llena.* (*Español intermedio,* p. 137)

Chapter 16 of this textbook is specifically dedicated to various uses of *se*. Here, five different uses are introduced:

(58) a. Indirect object *se*: *Dí**se**lo antes de que sea demasiado tarde.*

b. Reflexive *se* (It turns a transitive verb into an intransitive verb): *Cada vez que **me** equivoco, **me** vuelvo loco.*

c. Reciprocal *se*: ***Se** enfrentaron muy enfadados y finalmente **se** golpearon.*

d. Involuntary/Unintentional *se*: ***Se** me olvidó su nombre.*

e. Meaning-changing *se* (Sentence without *se* is grammatical; Adding *se* slightly changes the meaning):
abandonar 'to abandon' *vs. abandonar**se*** 'to let oneself go; to give in'; […]
referir 'to refer' *vs. referir**se** a* 'to refer to; to be related to'

(*Español intermedio,* pp. 164-166, [translation mine])

Interestingly, the textbook provides extensive examples of different *se* constructions. However, as can be seen in (58e), it does not adequately address the meaning change conveyed by the use of *se*.

In the English textbooks, the term 'reflexive pronoun' is constantly used even for non-reflexive *se* constructions where reflexivity is not apparent. On the other hand, in the Korean textbooks, there is an emphasis on the detransitivizing function of *se* when it is introduced to students. The biggest difference is that the Korean textbooks provide more grammar-centered explanations while the English textbooks compare clitic *se* to similar English constructions. This difference might derive from (i) the relative typological similarity between English and Spanish, and (ii) Korean students' prior exposure to metalinguistic concepts from learning English as L2. The following table summarizes the stages in which *se* constructions are presented in each textbook:

Table 13. Comparison of *se* introduction in four textbooks

		Reflexive	**Reciprocal**	**Passive**	**Impersonal**	**Accidental**	**Meaning change**	**Summary**
Elementary Spanish 1	*Experience Spanish* (Ch. 1-6)	Ch. 5	Ch. 5					
	Español básico 1							
Elementary Spanish 2	*Experience Spanish* (Ch. 7-12)			Ch. 7		Ch. 7		
	Español básico 2	Ch. 2	Ch. 2	Ch. 3	Ch. 9			
Intermediate Spanish 1	*Más* (Ch. 1-7)	Ch. 2			Ch. 6	Ch. 4	Ch. 2	
	Español intermedio (Ch. 1-10)							
Intermediate Spanish 2	*Más* (Ch. 8-12)			Ch. 12				Ch. 12
	Español intermedio (Ch. 11-20)			Ch. 13		Ch. 16	Ch. 16	Ch. 16

Based on the order of exposure, it is plausible to predict that both English and Korean speakers would be most familiar with *se*'s reflexive use in daily routine contexts. Next, the textbooks explain passive and impersonal *se* constructions, but the way they describe them is slightly different: While the English textbooks encourage students to use the two constructions indistinctively, the Korean textbooks differentiate them and provide more metalinguistic explanations. Also, when it comes to accidental and meaning-changing uses, they are introduced in elementary level courses at UGA taken during the second semester, whereas at SNU, they appear in the last chapter of intermediate level courses during the second semester.

Therefore, it is possible to assume the following order of acquisition for native speakers of English and Korean: Reflexive > Passive/Impersonal > Others. However, it is predicted that English speakers will be more familiar with the final category (i.e. Others, including accidental and meaning-changing categories) than Korean speakers. Also, Korean speakers will be more likely to focus on producing grammatically correct sentences, based on the detailed grammatical information provided in the Korean textbooks.

Furthermore, it is plausible that the 'reflexive' terminology prevalent in both English and Korean textbooks misleads L2 learners' understanding of Spanish clitic *se* constructions by falsely introducing non-reflexive *se* constructions as 'reflexive' ones. As Maldonado (2008) suggests, clitic *se*'s true reflexive function needs to be separated from its middle functions. According to Maldonado, a true reflexive is when the clitic *se* is used to mark coreferentiality between subject and object of a transitive verb. All other uses of *se,* where there is a lack of coreferentiality, should be analyzed as middle constructions without any association with reflexivity. To enhance the efficiency of Spanish teaching and learning, Maldonado (2008:158) suggests that "middles must be introduced as opposed to reflexives" and that various middle constructions have to be presented together as middle domain, which is a bigger schema that encompasses all non-reflexive uses of *se*, rather than giving students a list of unrelated exceptions of 'reflexive' function of the clitic *se*.

CHAPTER 4

METHODOLOGY

4.1 Introduction

This chapter outlines the methodological aspects of this study. The primary objective of the present study is to investigate the acquisitional process of the Spanish clitic *se* as a middle marker by L1 English- and L1 Korean-speaking adult learners of L2 Spanish. In order to achieve this goal, an online linguistic survey was conducted to examine learners' and native speakers' interpretation and production of the Spanish middle marker *se*. The survey included a language background questionnaire; a Spanish proficiency evaluation; and three experimental tasks: a written narration, an acceptability judgment task, and a cloze test. The results of the survey will be analyzed and compared across learner groups to determine each group's developmental pattern in the acquisition process. The study also aims to examine how the learners' linguistic behaviors differ from those of native Spanish speakers.

Given the wide range of semantic uses of the Spanish middle domain (cf. Table 4) proposed by Maldonado (2008, among others), which is beyond the scope of this dissertation, the specific constructions that will be tested in the present study will be focused on grooming or body care (i.e. *lavarse* 'to wash', *peinarse* 'to comb one's hair'), emotional reaction (i.e. *alegrarse* 'to be happy', *quejarse* 'to complain'), and spontaneous change of state (i.e. *romperse* 'to break', *caerse* 'to fall'). Two reasons exist for analyzing these three construction types. First, since L2 Spanish classrooms generally introduce 'reflexive *se*' with grooming or body care

verbs, it is expected that participants will already be familiar with these uses even at the lowest level of proficiency. Further research is needed to determine whether they are aware that the clitic *se* in these constructions is in most cases non-reflexive and therefore not compatible with the extended form *a sí mismo*. Second, Maldonado (2008:165) characterizes the emotional reaction construction type as a prototypical middle construction and argues that a spontaneous change of state has a strong connection with prototypical middles. It is plausible to assume that non-native speakers are more familiar with prototypical middle constructions given their ubiquitous character, so including these types of verbs ease participants' task completion. Therefore, in order to measure the extent to which L2 learners interpret *se* as a middle marker and to analyze their use of *se* in its more prototypical constructions, the three construction types (i.e. grooming or body care, emotional reaction, spontaneous change of state) mentioned above will be examined and discussed in detail.

4.2 Research Questions

The general research questions of the present study relate to the interpretation and production of Spanish clitic *se* by adult L1 English and L1 Korean learners of L2 Spanish from the viewpoint of cognitive linguistics. These questions are as follows:

RQ 1: Are there differences between L1 English and L1 Korean speakers' knowledge and acquisition of *se*-marking in Spanish and how does this compare to native Spanish speakers' knowledge and use?

RQ 2: What are the differences among the different proficiency levels in terms of their knowledge and acquisition of *se*-marking and how does this compare to native Spanish speakers' knowledge and use?

In addition to the above two general questions, I have three specific ones which guided the design of the three tasks. These more specific questions are the following:

RQ 3: What are the linguistic and non-linguistic factors or variables that favor *se*-marking among L2 learners and native Spanish speakers?

RQ 4: Do learners and native speakers associate absolute verbal events with non-*se*-marking and dynamic events with *se*-marking?

RQ 5: Do learners and native speakers understand the difference between true *se*-marked reflexives as evidenced by the acceptance of the addition of the phrase *a sí mismo*, and *se*-marked middles as evidenced by the rejection of the phrase?

My hypotheses set up to answer these questions are:

H1a: Not only do English and Korean middle constructions differ from Spanish middle constructions, but there are also differences between the two native languages of the learner groups. If the speakers' L1 has a stronger impact on transfer, the results will also be different for each group of L2 speakers, based on their respective L1s. Also, the discrepancy of the middle system between the three languages will result in non-target-like performance by L2 learners.

H1b: In general, L1 Korean speakers acquire English as their L2 and Spanish as their L3. If L3 interpretation and production is more greatly affected by L2 transfer than by L1 transfer, English-speaking learners' and Korean-speaking learners' uses of *se* will not differ to a great extent. In this case, it will be possible to assume that L2 learners' non-target-like behavior might be attributable to certain linguistic properties of English.

H2: The Spanish clitic *se* is well known for its variable use among native speakers, especially dialectally (Ibarretxe-Antuñano et al. 2016:157). Moreover, some constructions require an obligatory use of *se* (e.g. *quejarse* 'to complain'), while other verbs are possible in both non-*se*-marked and *se*-marked forms (e.g. *ir/irse* 'to go'/'to leave'). The variable use could cause learnability problems for L2 learners; thus, the interpretation and production of middle marker *se* by L2 learners will show differences from native speakers. Especially, the lower-level learners who have received limited input of *se* constructions are expected to demonstrate more non-target-like performance. However, based on the fact that the advanced learners are exposed more to middle constructions, both inside and outside of the classroom, L2 learners' performance will gradually become more native-like as proficiency level increases. Furthermore, L2 learners' production will demonstrate greater divergence than their interpretation, given that active language skills are acquired after passive skills.

H3: Regarding linguistic factors, it is expected that (i) transitive predicates, (ii) null subject form, (iii) postverbal subjects, (iv) past tense forms, and (v) telic verbs will favor the use of *se*. [24] In the same vein, the lack of these factors will disfavor the use of *se*. As for the extralinguistic variables, L2 learners with higher levels of proficiency and/or longer previous experience living in Spanish-speaking countries will produce more native-like *se* constructions, given that previous literature (cf. Escutia López 2010, Tremblay 2005) claims that *se* constructions are late-acquired by non-native speakers.

[24] Detailed discussions on why these predictions are made will be followed in 4.4.

H4: As proven by L1 acquisition studies from the framework of cognitive linguistics (cf. Jackson-Maldonado et al. 1998, Soto & Muñoz 2000), native speakers are expected to associate *se*-marking with the dynamicity of the event. However, since this association is not explicitly taught in the L2 classroom, lower-level L2 learners are more likely to diverge from the expected patterns, while learners with greater proficiency will behave more native-like and will accept *se*-marking in sentences with a dynamic context and accept non-*se*-marking in sentences with an absolute context.

H5: Since the majority of, if not all, Spanish classroom instructions and textbooks introduce *se* as a reflexive marker, L2 learners will first make a one-to-one correspondence between *se* and its identification as a reflexive marker. Therefore, lower-level learners, not being aware of the middle marking function of *se*, will over accept sentences with *a sí mismo* even in the constructions where it is unacceptable or pragmatically infelicitous.

As can be seen above, there are two hypotheses (i.e. H1a and H1b) that respond to Question 1, and this derives from the nature of Third Language Acquisition (L3A). According to García Mayo & Rothman (2012), transfer in L3A is specifically unpredictable because L3 learners are likely to have more advanced metalinguistic skills compared to L2 learners. This unpredictability

gives rise to two disparate assumptions in the field of L3A: some argue in favor of absolute L1 transfer, while others present support for an L2 status factor.[25]

Absolute L1 transfer proposes that "the L1 acts as a filter of sorts, impeding access to acquired L2 properties" (2012:16). Hypothesis 1a stated above partially coincides with absolute L1 transfer in that it predicts that L1 properties have a greater impact on L3 performance than L2 properties. On the other hand, the L2 status factor position suggests that "the L2 takes on a significantly stronger role than the L1" (2012:17). This point of view describes my Hypothesis 1b, which presumes that Korean speakers' linguistic behavior will demonstrate greater influence from their L2, English, than their L1.

However, the discussion of L1 Korean speakers' performance here cannot be considered as a proper indication of L3A for some reasons. First, the linguistic survey conducted in the present study was not designed specifically for a L3A analysis. In order to examine transfer in L3A, the same properties in the L1 and L2 should be examined as well. Although some information on the Korean participants' English proficiency was collected, it is not sufficient to provide any assumptions on how the participants would have interpreted and produced the constructions in question in English.

Also, the aforementioned debate on L1 *vs.* L2 transfer is based on the initial stage of L3A. Since all participants in the present study have at least a low-intermediate level of Spanish proficiency, their acquisitional stage would be classified as a developmental or ultimate stage. In

[25] Rothman (2010) argues that the Typological Primacy Model best describes the process of multilingual transfer. Under the Typological Primacy Model, it is assumed that "proximity in actual or perceived linguistic typology between the target L3/L*n* measured against the grammars of the L1 and L2 is the most deterministic variable to predict which of these previous systems is selected for adult multilingual syntactic transfer" (2010:108).

fact, there is an alternative approach, the Cumulative-Enhancement Model (Flynn et al. 2004, cited in García Mayo & Rothman 2012:18), which accounts for developmental and ultimate stages as well as the initial stage. Its main argument is that "all previously acquired properties are in theory available", therefore, both L1 and L2 transfer may be involved in L3 development. Nonetheless, examination of this model is beyond the scope of this dissertation and will be reserved for future research.

4.3 Participants

Two experimental groups of L2 Spanish learners and one control group of L1 Spanish speakers participated in this research. The first experimental group is composed of 33 L1 English speakers who acquired Spanish as their L2, including one participant (ENG57) who self-identified as an English-Spanish bilingual. The second experimental group consists of 30 L1 Korean speakers who learned English as their L2 except for two participants (KOR06 and KOR42) who self-identified as Korean-Spanish bilinguals that acquired Spanish as their L2; thus, all others learned Spanish as the L3. All self-identified bilingual speakers were grouped with other participants because their performance on all tasks was not notably different from the rest. A group of 30 L1 Spanish speakers from different regional backgrounds is included as a control group.

The present research was conducted online using Qualtrics (Qualtrics, Provo, UT), an online survey software program. Given the nature of the questionnaire, every subject needed internet access to participate in the survey. The questionnaire created in Qualtrics was distributed first via email and through social media, such as Facebook and Instagram, to recruit as many participants as possible. The researcher initially contacted faculty members, instructors and

graduate students at the University of Georgia (UGA) and Seoul National University (SNU) — the researcher's current and former institution, respectively— and asked their help in recruiting their students who were taking intermediate and upper-level Spanish classes at that moment. A link for the questionnaire created in Qualtrics was provided in the email, and the faculty, instructors and graduate students were asked to forward the email to their students and/or upload an announcement in their online course platforms (i.e. eLC and eTL) to give their students access to the survey. For the control group, native Spanish speakers from UGA and SNU were recruited first. Since the number of native speakers was smaller than that of the L2 learners in both institutions, further recruitment measures were taken. For example, information about the survey and its related links was posted in Linguist List (www.linguistlist.org), a website dedicated to providing information on language and linguistics, under the topic "Queries" to recruit more participants outside of the two institutions, especially native speakers of Spanish.

The criteria for participation in the two experimental groups included their first language, Spanish proficiency and their completion of each task. In the first part of the survey where a language background questionnaire was included, each participant was asked what their native language is to determine how they identify themselves in terms of their first language. With regard to Spanish proficiency, although the clitic *se* is first introduced to learners in elementary level classrooms, previous literature has shown that the L2 acquisition of *se* occurs in later stages of language development. Therefore, it was considered that the participants' proficiency level had to be at least intermediate, especially for the purpose of Task 1, a picture book written narration task (cf. Appendix C) which requires participants to produce sentences in Spanish in a spontaneous manner. In order to determine the level of Spanish proficiency more accurately, L2 learner groups completed a proficiency evaluation test, adapted from a subsection of the DELE

(*Diploma de Español como Lengua Extranjera* 'Diploma of Spanish as a Foreign Language'), provided in the questionnaire (cf. Appendix B). Taking these into account, out of a total of 50 L1 English speakers who participated in the survey, 11 participants were eliminated because of their lower proficiency level of Spanish, and another six were not included in the data analysis for not fully completing any of the tasks. For example, participants who started Task 1 but left the survey without completing it were excluded from the analysis. As such, from 44 L1 Korean speakers in total, six lower-level participants were excluded, and another eight were not included because they left the survey without completing the tasks. For L1 Spanish speakers, 19 of 49 participants were removed from the data analysis when task completion was taken into consideration. None of the participants was compensated for their participation in the study.

For L1 English speakers, the following information on personal data and language background was collected: (1) age, (2) country of origin, (3) education level, (4) gender, (5) first language, (6) age when the participant first began to learn Spanish, (7) knowledge of other languages apart from the first language and Spanish, (8) experience living in Spanish speaking countries, (9) method of learning Spanish, and (10) level of Spanish language courses attended.

For L1 Korean speakers, there were 2 more questions added to the aforementioned 10 items: (11) whether they learned English or Spanish first, and (12) their average score in official English proficiency tests. The reason for collecting information on the order of acquisition between English and Spanish was to confirm that Korean speakers were exposed first to English. This was proven to be true, except for the two participants who identified themselves as Korean-Spanish bilinguals. They reported that they acquired Spanish before English. Also, instead of including an English proficiency test in the survey, the average score of various official English proficiency tests was asked. Since the survey conducted in this study included a Spanish

proficiency test, adding a second, English proficiency test ran the risk of overloading the participants. Thus, it was decided that asking Korean participants to report their previous scores would be an acceptable reflection of English proficiency. Specifically, official test scores from one (or more) of the following tests were qualified: TOEFL, Test of English as a Foreign Language; TOEIC, Test of English for International Communication; TEPS, Test of English Proficiency developed by Seoul National University; and New TEPS, an updated version of TEPS which was launched in 2018. The first two tests are most commonly taken in South Korea, and the last two tests are a requirement for students who attend Seoul National University. Every L1 Korean speaker reported at least one of those test scores, with the exception of one 47-year-old female, who was the only high school graduate participant in the present study.

A summary of each participant group is provided in Table 14 below:

Table 14. Summary of participant demographics

	L1 English (*n*=33)	**L1 Korean (*n*=30)**	**L1 Spanish (*n*=30)**
Age (mean)	26.5	26.2	37.5
Country of origin	Canada (1), Mexico (1), US (31)	Japan (1), South Korea (29)	Argentina (3), Colombia (3), El Salvador (1), Mexico (2), Panama (1), Paraguay (1), Peru (2), Spain (16), US (1)
Education level	Some College (15) College Graduate (1) Graduate School (17)	Highschool Graduate (1) Some College (13) College Graduate (3)	Some College (1) College Graduate (5) Graduate School (24)

		Graduate School (13)	
Gender	Male (9) Female (24)	Male (11) Female (19)	Male (10) Female (19) Prefer not to answer (1)
L1	English (31) English-Spanish (2)	Korean (28) Korean-Spanish (2)	Spanish (27) Spanish-Korean (1) Spanish-Other(s) (2)
L2 English or L2 Spanish, where applies	N/A	English (28) Spanish (2)	N/A
Age of Onset of Spanish instruction, where applies (mean)	12.7	16.7	N/A
Other language(s)[26]	Danish (1), Farsi (1), French (4), Galician (1), German (2), Italian (2), Malay (1), Portuguese (12), Quechua (3), No response (17)	Catalan (1), Chinese (3), French (5), German (2), Japanese (3), Latin (1), Portuguese (7), No response (14)	Basque (2), Catalan (5), Czech (2), English (27), French (12), Galician (1), German (6), Greek (1), Hindi (1), Italian (6), Korean (2), Portuguese (6), No response (1)
Experience living in Spanish speaking	Argentina (4) , Bolivia (1), Chile (1), Costa Rica (1), Ecuador (1), Honduras (1),	Argentina (1), Chile (2), Guatemala (2), Mexico (12), Peru (1), Spain (12), Uruguay (1)	N/A

[26] This question was designed differently for each group: For L1 English speakers, it asked participants to include language(s) learned apart from English and Spanish; for L1 Korean speakers, all languages excepting Korean and English; for L1 Spanish speakers, all languages except Spanish.

countries, where applies	Mexico (7), Peru (4), Spain (8), Uruguay (1)		
Length of Residency in Spanish speaking countries (mean), where applies	9.4 months	12.7 months[27]	N/A
Acquisition environment, where applies	Classroom instruction (17) Interaction with people (1) Mixture of both (15)	Classroom instruction (14) Interaction with people (2) Mixture of both (14)	N/A
Level of courses taken[28], where applies	E, I (2); E, I, A (17); I (2); A (12)	E, I (4); E, I, A (14); E, A (1); I (3); I, A (1); A (6); No response (1)	N/A
English proficiency, where applies	N/A	High-intermediate (1) Advanced (17) Near native (11) No response (1)	N/A

Of a total of 93 participants, the number of participants calculated for each task is slightly different, since some participants did not complete all three tasks. Below is a frequency table depicting the number of participants who completed each task:

[27] Two Korean-Spanish bilinguals' information was excluded here. Their lengths of residency were 144 months (KOR06) and 360 months (KOR42), respectively.

[28] E = Elementary, I = Intermediate, A = Advanced.

Table 15. Participant frequencies according to task

	L1 English	**L1 Korean**	**L1 Spanish**
Task 1	33	30	30
Task 2	32	30	28
Task 3	28	29	27

4.4 Survey design and data collection

The first part of the survey included a language background questionnaire (cf. Appendix A). The language of the survey was administered according to each participant's first language. When participants clicked on the survey link, they were directed to a page where it states, 'If you are a native speaker of *x*, please click the arrow below'. Two other options ('If you are a native speaker of *y/z*, please click here.') with the corresponding links to the survey written in other languages were also provided. This measure was used to ensure that all participants completed the survey in their respective first language. Below is an excerpt of this front page from the English version of the survey:

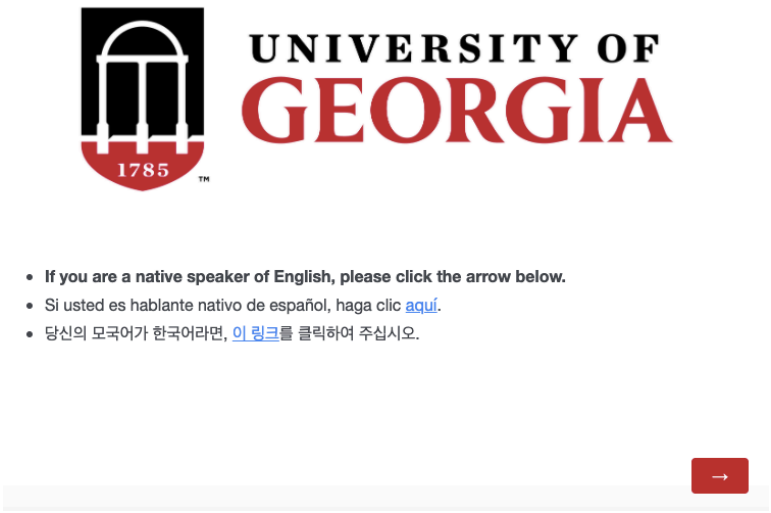


Figure 3. Excerpt of the survey's front page

Next, a letter of consent was provided, which was followed by some questions requesting basic information about the participant, as discussed earlier in 4.3. After the linguistic background questionnaire, a Spanish proficiency evaluation (cf. Appendix B) was administered to all participants, except the L1 Spanish speakers, to determine to which L2 Spanish proficiency group the participants should be assigned. This Spanish proficiency test was adapted from the DELE-based Proficiency Test (http://nhlrc.ucla.edu/data/proficiency-assessments-example-proficiency-exams.asp) by Montrul (2012, published online). The original version consisted of two parts: a Multiple Choice Test and a Cloze Test. In the present survey, only the Multiple Choice Test was conducted to lessen the participants' workload. A total of 30 questions were given, and each question was worth 1 point. Participants who scored between 0 and 17 points were assigned to the Elementary-level group and were excluded from the analysis. The Intermediate-level group refers to participants with scores between 18-23 points, and those who scored between 24-27 points were assigned to the Advanced level; L2 learners who scored above 28 points were grouped as Near-native-level speakers.[29]

Every participant, including those in the control group, was asked to complete three different tasks: a Picture Book Narration Task (cf. Appendix C), an Acceptability Judgment Task (cf. Appendix D), and a Cloze Task (cf. Appendix E).

[29] This assignment of scores to the different levels is also adapted from Montrul's scoring. The maximum total score for the original test is 50, and the levels are divided as follows: Low (0-29), Intermediate (30-49), and Advanced (40-50). Based on this scale, I first divided the levels into three groups: Low (0-17), Intermediate (18-23), and Advanced (24-30). Later, for the purpose of the present study, the last group was further divided into two separate groups: Advanced (24-27) and Near native (28-30).

First, in Task 1, a series of images adapted from a children's picture book, *Frog, where are you?* (Mayer 1969) were presented to participants. Out of 29 pictures from the original story, 10 images were selected in this task. The participants wrote a short description of the story in Spanish based on these images, as is shown in the excerpt below:

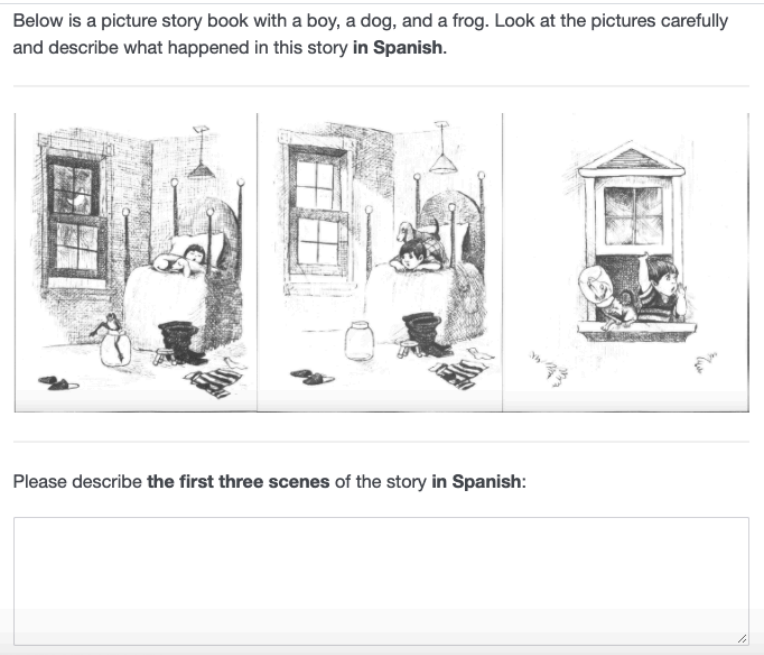


Figure 4. Excerpt of Task 1

The purpose of this task was to analyze differences in the use of *se* in native and non-native written production data. Specifically, uses of *se* in middle contexts including *despertarse, levantarse* (e.g. the boy and the dog wake up in the morning), *escaparse, desaparecerse* (e.g. the frog escapes and disappears), and *caerse, romperse* (e.g. the dog falls out of the window, the jar breaks), were expected and, in fact, found. Other functions were also seen, as discussed below and in Chapter 5.

A mixed-effects multivariate analysis in Rbrul (Johnson 2009) was used in order to analyze the written production data of this first task. Five linguistic variables (i.e. transitivity, subject form, subject position, tense-aspect-mood (TAM), and telicity) and several extralinguistic variables from the participants' personal data and language background questionnaires (e.g. native language, level of proficiency) were coded and analyzed to determine which variables, if any, played a significant role in the use of the middle marker *se* in each speaker group's narrative production. A stepwise variable selection procedure was conducted, in which the set of predictor variables was entered and removed to reach a line of best fit in the logistic regression model.

While designing the task, predictions were made for each independent factor on how certain (extra)linguistic environments will condition participants' use of the Spanish clitic *se* in their picture book narrations. Regarding transitivity, as mentioned earlier, clitic *se* is first introduced as a reflexive marker in L2 classrooms. Since reflexive constructions mostly include transitive predicates, I expected that the learners would favor transitive verbs when they use *se* in their writings.

As for the subject forms, I predict that the null subject form will be favored in both native and non-native speakers' writings. In Spanish, null subject forms are preferred unless the subject is focused or contrastive. Therefore, in a short story where two participants, the boy and his dog, perform similar actions at the same time, speakers are not expected to overtly use the explicit subjects with or without the use of *se*. Furthermore, upon predicting that some speakers will use impersonal *se* constructions (e.g. *Se ve…* 'You see…'), it is expected that a null subject form will condition the use of *se* more than other forms, such as a DP (e.g. *el niño, el perro*) or subject pronoun (e.g. *él, ellos*).

Although the position of the subject is not relevant in null subject constructions, pre-verbal and post-verbal subjects were coded to examine which subject position favors the use of the clitic *se*. A prototypical transitive construction adheres to SVO (Subject-Verb-Object) word order, but in passive *se* constructions, VS order is less marked than SV order (Kelling 2006). Additionally, based on the findings by Bailey (2013) and Guillén-Solano (2015) discussed in 3.3, L2 learners tend to acquire passive *se* constructions relatively early on, compared to impersonal and anticausative types. Therefore, it is predicted that L2 learners will more likely use *se* with post-verbal subjects, since *se* will be produced in numerous passive *se* constructions.

Tense-Aspect-Mood was included in the independent factors following Aaron & Torres Cacoullos' (2006) work discussed in 3.2. They found that the past tense forms, Preterite and Imperfect forms favored *se* marking in the verb *salir*. Their initial prediction was that the Preterite forms would show higher probability of *se*-marking than the Imperfect forms. However, it turned out that the difference between these two forms was not as significant as the difference between past and non-past tense forms more broadly. The authors explain that this is related to the fact that *salirse* is only used for events that have already happened. In other words, if the event had not happened yet or if it was impossible for it to happen, *se* was never used. Therefore, *se* is more likely to be used with past tense forms which describe situations that have taken place. Following Aaron & Torres Cacoullos' (2006) analysis, I predict that the participants in the present study will also favor the use of *se* with past tense forms.

The last linguistic variable examined in the current study is telicity. Among the four types of *Aktionsarten* (i.e. State, Activity, Accomplishment and Achievement), Accomplishments and Achievements seem closely related to the use of *se*. In order to describe spontaneous events or focus on a pivotal moment of a change of state, a verb that belongs to a lexical aspectual

category which bears punctuality is likely to be used, and Achievements are characterized by punctuality. Therefore, it is expected that Achievement verbs will favor the use of the middle marker *se* in the participants' written narrations. Also, as discussed in 2.5, English *get*-passives alternate with anticausatives with Achievement verbs (cf. Alexiadou 2012). Based on this assumption, I predict that L1 English and L1 Korean speakers familiar with English *get*-passive structures will favor the use of *se* with Achievement verbs based on their similarity to Spanish *se* constructions. Considering that Accomplishment and Achievement verbs share the characteristic of telicity, it will be worth investigating whether telic verbs favor the use of *se*.

Along with the independent linguistic factors, I assume that some extralinguistic variables will also condition the use of clitic *se*. Since previous literature discusses that constructions with *se* are acquired late by L2 speakers, it is expected that participants with higher levels of proficiency will perform in a more native-like manner. Also, I predict that non-native speakers with previous experience living in Spanish-speaking countries will produce more target-like sentences based on their greater exposure to Spanish. For L1 Korean speakers, those who are more fluent in English are expected to behave more native-like in Spanish because they are likely to have more advanced metalinguistic skills gained from their L2 acquisition process. Moreover, as even native speakers' use of *se* is variable, it will be of interest to examine which factors condition their use of *se* and to contrast them to those of L2 learners.

In Task 2, an acceptability judgment task (AJT) was designed to examine the participants' interpretation and comprehension of Spanish middle constructions where the *se*-marked verb indicates a dynamic event in contrast to the non-*se*-marked verb that indicated an absolute event. Based on a context provided in the participants' respective native languages, they were then asked to judge the naturalness of two Spanish sentences that described the

contextualized situation. Each situation was intended to lead to either an absolute or dynamic reading. The reader may recall that dynamic contexts are defined by a situation in which middleness (i.e. abruptness, rapidness, out of ordinary-ness) exists, whereas absolute contexts lack middleness. Furthermore, the reading passages were written in a way that no neutral or ambiguous context was provided.

There were 30 questions in total: 18 items were actually related to understanding the clitic *se*, and the other 12 items were distractors. The distractor items were related to Spanish lexical or grammatical concepts which Whitley (2002) considered problematic for L2 learners, especially L1 English speakers. Also, none of the distractor items were related to any type of *se* constructions, and instead included unrelated concepts (e.g. indicative *vs.* subjunctive, preterite *vs.* present perfect, the preposition *por vs. para*). Some examples are given below:

(59) a. This afternoon, when Estefanía turned on her laptop, she saw that it was not working properly. After having it checked by a technician, she was told that she needs to get a new laptop.

A: Estefanía vaya a comprar una nueva computadora portátil.

B: Estefanía va a comprar una nueva computadora portátil.

b. Carlos went out of town to attend a friend's wedding. When he flew back home this afternoon, he saw that the grass, sidewalks and streets were wet.

A: Parece que ha llovido esta mañana.

B: Parece que llovió esta mañana.

The order of the questions, including actual and distractor items, was randomized. Below is an excerpt of the task instructions and an example of two items, the first where the non-*se*-marked verb should be more natural and the second where the *se*-marked verb should be more natural, based on the contexts. Moreover, as shown below, the texts are provided not in Spanish

but in the participants' respective L1s in order to ease their understanding of the situation described in the text in addition to avoiding possible priming effects.

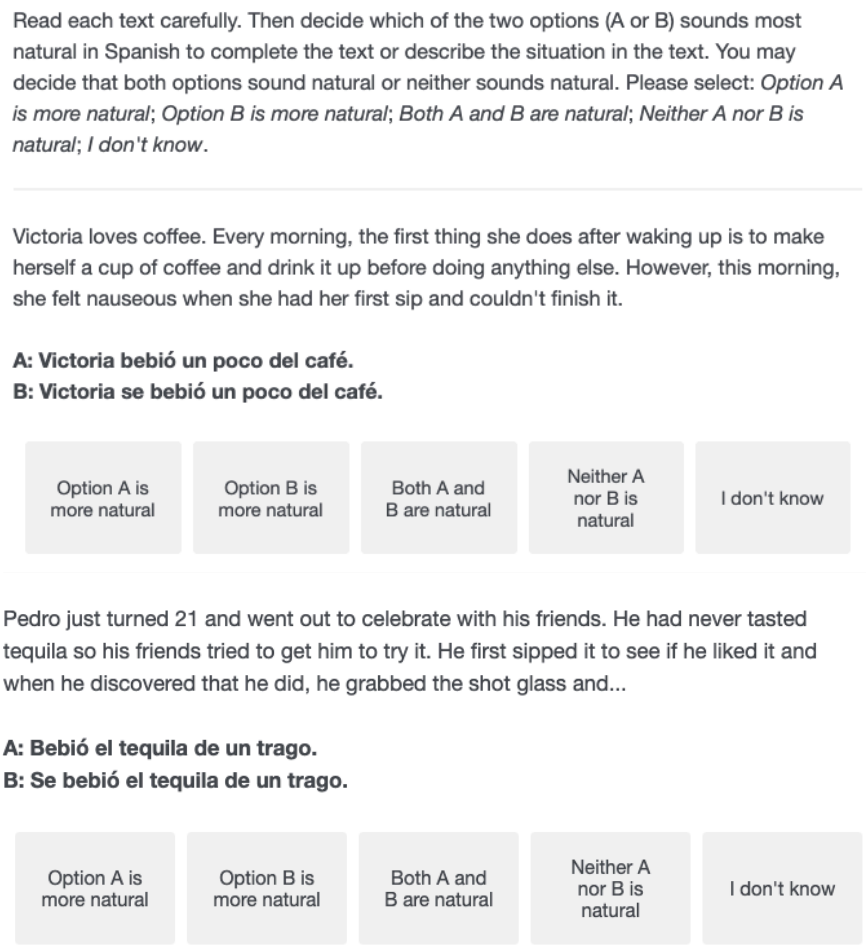


Figure 5. Excerpt of Task 2

The 18 actual questions contained one of the following verbs: *aprender/aprenderse* 'to learn/to learn by heart', *bajar/bajarse* 'to go down/to get off', *beber/beberse* 'to drink/to drink up', *caer/caerse* 'to fall/to fall down', *decidir/decidirse* 'to decide/to make up one's mind', *dormir/dormirse* 'to sleep/to fall asleep', *ir/irse* 'to go/to leave', *saber/saberse* 'to know/to have a full knowledge', and *subir/subirse* 'to go up/to get on'. These verbs can be used with or

without the clitic *se*, and based on discussions from the cognitive perspective, the use of *se* with these verbs marks the middleness of the event. In other words, *se*-marking results in a dynamic reading, which means that the conceptualization of the given event focuses on the crucial moment of the change of state, which is characterized by rapidness, abruptness, and/or 'out of the ordinary'-ness of the event. The non-*se*-marked verb results in an absolute reading of the event.

As exemplified by Figure 5, for the 18 items related to the use of *se*, the sentence in option A was not marked by *se*, while option B consisted of a *se*-marked sentence. By asking which option sounds more natural to the participants, this AJT aimed to examine whether the native and non-native speakers associated the clitic *se* with the middleness of the event. In other words, if the participants answer such that the non-*se*-marked form of the verb sounds more natural in absolute contexts, and the *se*-marked form sounds more natural in dynamic contexts, this leads us to postulate that they recognize the association. In opposite cases (i.e. judging non-*se*-marked constructions dynamic and *se*-marked ones absolute), we interpret these results such that the speakers do not relate the use of *se* to middle contexts.

In order to compare and contrast the general performance of each participant group, I coded their responses based on the expected association and converted the results into numeric values. For example, if a participant chose the non-*se*-marked form (Option A) in an absolute context, the response was scored as 2. If s/he selected the *se*-marked form (Option B) in the same context, it means that the response did not reflect the expected performance, so it was scored as 0. In cases where participants chose 'Both A and B are natural', it was considered a neutral response and was scored as 1. For all other options ('Neither A nor B is natural', 'I don't know'),

a score of 0 was given. Also, a score of 0 was given for all questions for which there was no response.

Based on the descriptions of the middle marker *se* from the standpoint of cognitive linguistics discussed in 2.4, I predict that native speakers of Spanish would show a clear pattern of marking *se* in dynamic contexts. That is, when a specific context favoring either the presence or absence of the clitic *se* is given, native speakers are anticipated to follow the 'expected' response. For example, if a given context expresses a dynamic situation where there is a student (*Juan*) who falls asleep during class because of his lack of sleep, the use of middle marker *se* in *De repente, Juan* ***se*** *durmió en la clase* 'Suddenly, Juan fell asleep in class' is expected to sound more natural, thus the number of the response 'Option B is more natural' should be higher than all other choices.

However, for the non-native speakers, I assume that the level of proficiency would be relevant for their responses. While L2 learners with a higher level of proficiency would perform in a more native-like manner, lower-level learners could still have difficulty interpreting the meaning conveyed by *se* either because of L1 and/or L2 transfer or because of a lack of input. Therefore, their responses are expected to show a gradual resemblance to native-like performance along the development in L2 language proficiency.

On the other hand, when L1 English and L1 Korean speakers are compared, I predict that L1 English speakers will perform in a more native-like manner because they are exposed earlier to various *se* constructions. Based on the textbook analysis (cf. 3.4), it was confirmed that textbooks used in the American institution present non-reflexive *se* constructions earlier in the curriculum than those used in the Korean institution. Therefore, L1 English speakers will be more familiar with the *se* constructions offered in this task, displaying more 'expected'

performance than L1 Korean speakers whose input is mostly limited to reflexive, passive, and impersonal *se* constructions.

Even though there is an 'expected' pattern of responses, it does not necessarily mean that certain responses are grammatical or correct and that others are ungrammatical or wrong. As it was for Task 1, Spanish speakers are likely to mark *se* differently depending on various contexts. In other words, some responses might differ greatly even among the L1 Spanish speaker group. Therefore, it will be of interest to analyze to what extent the predictions by cognitive linguistics align with the results of this experiment among native speakers.

The last task of the current study, Task 3, was a cloze task which specifically tested participants' knowledge or awareness of *se* as a middle marker or as a reflexive marker. It asked the participants whether each use of *se* in a paragraph can be expanded with the phrase *a sí mismo* or not. A total of 10 short paragraphs written in Spanish was provided, and they only included first and third person *se* constructions. Nine texts had two *se* constructions, one had four *se* constructions. None of these short paragraphs was related to each other, and the order of the questions was randomized. Below is an excerpt of the Task 3 instructions and an example of one of the questions:

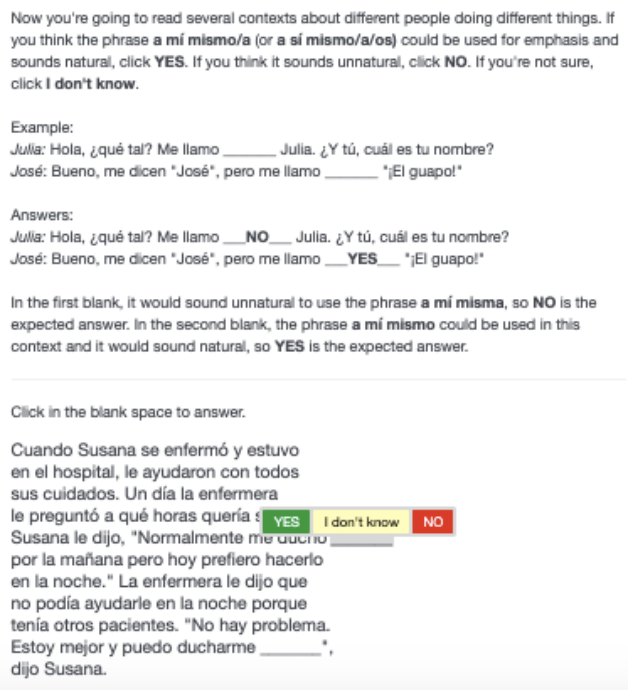


Figure 6. Excerpt of Task 3

As seen above, a blank space was provided next to each *se* construction. When the participants clicked on the blank space, a block with three options appeared (i.e. 'Yes', 'I don't know', 'No'), so they could make a selection by clicking their response.

Half of the texts contained *se*-marked daily routine verbs or grooming/body-care verbs. This latter term 'grooming/body-care verbs' is used by Maldonado (2008); I use the term 'Routine verbs' for this category which includes *afeitarse* 'to shave', *despertarse* 'to wake up', *ducharse* 'to take a shower', *levantarse* 'to get up', *peinarse* 'to comb one's hair', and *vestirse* 'to get dressed'. Verbs used in the other half of the texts were extracted from CREA (*Corpus de Referencia del Español Actual* 'Corpus of Reference of Contemporary Spanish'): *confundirse* 'to get confused', *encontrarse* 'to find oneself', *imaginarse* 'to imagine', *perderse* 'to get lost' and *verse* 'to see oneself'. These specific verbs were selected because their use with the extended form, *a sí mismo*, was explicitly demonstrated in CREA. Each verb listed above appeared twice

in the text. It is expected that, according to the context, one of the two uses is acceptable with *a sí mismo*, while the other would sound unnatural when expanded with *a sí mismo*.

Based on the characteristics of the verbs described above, I classified them into four categories: 'Routine-Reflexive', 'Routine-Middle', 'Non-Routine-Reflexive', and 'Non-Routine-Middle'. The first classification describes whether they belong to the group of daily routine verbs or not, which were then classified as 'Reflexive' or 'Middle', indicating whether it was compatible with *a sí mismo*.

The most distinctive characteristic of Task 3 is that it explicitly asked for participants' judgments on Spanish *se* constructions. In the previous two tasks, there was no explicit mention of Spanish *se* constructions. For example, in Task 1, participants were only asked to narrate a story in Spanish. Participants did not have any idea that the task was examining their production of *se* constructions. In Task 2, although a greater number of questions included *se*-marked sentences (18 *se* items *vs.* 12 distractors), it was never obviously stated that the broader objective of the survey was to analyze their understanding of *se* constructions. Some participants might have noticed that the survey was related to the concept of clitic *se*; however, the distractors and randomized order of questions should have distracted them from focusing exclusively on *se* constructions. Since the participants' awareness of what they were tested on would have affected their performance, Task 3 was specifically reserved as the last part of the survey.

The relation between reflexive *se* and *a sí mismo* can be acquired earlier, given that they are explained (explicitly) from basic-level classroom instruction. However, L2 learners are not exposed to the unacceptability of the coexistence of middle marker *se* and *a sí mismo* in the classroom setting, to my knowledge. None of the textbooks analyzed in 3.4, for example, mentioned anything about *se* constructions in which the phrase *a sí mismo* is not acceptable.

Therefore, it is expected for learners, especially lower-level learners, to overgeneralize the majority of instances of *se* as compatible with the use of *a sí mismo*.

As for the L1 Spanish speakers, I predict that Task 3 will show a greater difference in their responses than in Task 2. This is because Task 3 items are more related to metalinguistic concepts. In other words, determining whether or not to use the clitic *se* in certain sentences would only require the speakers to consider the context or situation type, which would form part of their usual language performance. However, deciding if the phrase *a sí mismo* can be used is beyond what speakers usually do in their daily language use, hence it causes the speakers to think about the linguistic concept itself, putting an extra burden on their workload. Therefore, it is expected that responses will vary depending on participants' subjective and/or dialectally variable judgments on the compatibility of *se* with *a sí mismo*.

In order to statistically analyze the results of Task 2 and Task 3, all responses were analyzed in a software program for statistical computing and graphics, R (R Core Team 2013). Using R, an Analysis of Variance (ANOVA) was performed to determine if there were any significant differences between groups: L1 English group, L1 Korean group, and native speaker control group. Particularly, a regression model was fit to perform an ANOVA test on the data. Then, a post-hoc analysis with Tukey's test was conducted to interpret and analyze the survey results, to examine the significance of the differences among groups.

This chapter has explained the methodological steps that were carried out to analyze the interpretation and production of Spanish *se* constructions by English- and Korean-speaking L2 Spanish learners and L1 Spanish speakers. The next chapter will cover the results of the statistical analysis. Specifically, I will discuss which variables favored the use of *se*-marked constructions in Task 1, and what the results indicate about each participant group's performance

in Tasks 2 and 3 and these learners' knowledge and acquisitional development of various Spanish *se* constructions.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Introduction

In this chapter, I present the results for the three experimental tasks carried out in the present study: (i) the picture book narration task, (ii) the acceptability judgment task (AJT), and (iii) the cloze task, and analyze and discuss these results. The analyses compare and contrast responses by the three participant groups: (i) English speaking learners of L2 Spanish, (ii) Korean speaking learners of L2 Spanish, and (iii) native Spanish speakers, and focus on the findings where statistical significance is shown (although non-statistically significant findings are also discussed).

First, I discuss the results from the written production data for both L2 learners and native speakers regarding the independent variables that show an effect on Spanish clitic *se* marking. For the AJT, I summarize the results and explain to what extent the use of clitic *se* is associated with the dynamicity of the given context by native and non-native speakers of Spanish. Finally, for the cloze task, I present the findings that demonstrate how strongly the clitic *se* is conceptualized as a reflexive marker, which can be expanded with the phrase *a sí mismo*. In presenting the results, I report important findings from both intra- and inter-group analyses.

5.2 Task 1 (Picture book written narration task) Results — identifying variables that favor *se*-marking

In the picture book narration task, participants were asked to write a narration of a short story of a boy, a dog, and a frog in Spanish based on a series of images from a children's picture book, *Frog, where are you?* (Mayer 1969). The purpose of this task was to determine which (extra)linguistic variables affect the use of the middle marker *se* in each speaker group. By analyzing the task results, it was possible to draw some important findings. Firstly, all three groups showed a similar frequency of *se*-marking with an average of 20%. Among the factors expected to predict *se*-marking, telicity demonstrated a statistically significant effect across all groups. That is, both native and non-native speakers favored the use of *se* with telic verbs while atelic verbs were less likely to be *se*-marked. The inconsistencies between the groups were found in Tense-Aspect-Mood: L1 English speakers favored *se*-marking in Preterite, whereas L1 Korean and L1 Spanish speakers did so in Present tense.

All finite sentences (containing a conjugated verb) and non-finite phrases (containing an infinitive, gerund or past participle) were tabulated and subsequently, all constructions were identified as containing a *se*-marked verb or containing a verb not marked with *se* (non-*se*-marked). Table 16 below summarizes the distribution of the responses across all three participant groups, including the frequency and percentage of the total verb tokens; furthermore, the learners are divided further into intermediate, advanced and near-native speaker groups for each L1 language based on their scores on the DELE exam (Int = 18-23; Adv = 24-27; Near = 28-30).

Table 16. Distribution of responses by all participant groups (Task 1)

	ENG (*n*=33)						**KOR (*n*=30)**						**SPAN (*n*=30)**	
	Int (*n*=7)		Adv (*n*=13)		Near (*n*=13)		Int (*n*=3)		Adv (*n*=13)		Near (*n*=14)			
	#	%	#	%	#	%	#	%	#	%	#	%	#	%
se-marked	22	19%	56	22%	61	21%	9	20%	55	19%	70	24%	162	21%
Non-*se*-marked	94	81%	199	78%	231	79%	36	80%	231	81%	226	76%	611	79%
Total	116	100%	255	100%	292	100%	45	100%	286	100%	296	100%	773	100%

The results in Table 16 demonstrate that, in general, the two learner groups produced quite a similar number of verbs: 663 for the L1 English speaker groups combined and 627 for the L1 Korean speaker groups combined. Meanwhile, the native speakers used more verbs than the other two groups, showing a total count of 773. On the other hand, the percentage of participants' production of *se*-marked phrases was similar across all three participant groups (ENG, KOR, SPAN) and also across all Spanish proficiency levels (Int, Adv, Near) ranging from 19% to 24%.

When only *se*-marked verbs are compared across the three different L1 groups, the total number of tokens was 139 for the L1 English group, 134 for the L1 Korean group, and 162 for the L1 Spanish group. In terms of type frequency, 39 types of verbs were *se*-marked by the L1 English group, 47 types by the L1 Korean group, and 40 types by the L1 Spanish group. Despite the high frequency of verb types, most of these verbs occurred only once in the data set. Therefore, those verbs which I consider to be highly frequent are those which occurred 10 or more times. Table 17 below summarizes the types of verbs which were *se*-marked in the production data. The underlined verbs indicate that they occurred fewer than 10 times.

Table 17. Frequently *se*-marked verbs by all participant groups (Task 1)

ENG		**KOR**		**SPAN**	
Verb	# (%)	Verb	# (%)	Verb	# (%)
asomarse	2 (1.44%)	*asomarse*	3 (2.24%)	***asomarse***	**18 (11.11%)**
caerse	23 (16.55%)	*caerse*	20 (14.93%)	*caerse*	14 (8.64%)
darse	12 (8.63%)	*darse*	8 (5.97%)	*darse*	11 (6.79%)
despertarse	18 (12.95%)	*despertarse*	13 (9.70%)	*despertarse*	14 (8.64%)
escaparse	15 (10.79%)	*escaparse*	10 (7.46%)	*escaparse*	15 (9.26%)
irse	**10 (7.19%)**	*irse*	6 (4.48%)	*irse*	4 (2.47%)
romperse	3 (2.16%)	*romperse*	9 (6.72%)	***romperse***	**10 (6.17%)**
verse	0 (0%)	*verse*	1 (0.75%)	***verse***	**13 (8.02%)**

The majority of the verbs listed in the above table shows a difference in meaning depending on whether it is *se*-marked or not: *caer* 'to fall' *vs. caer**se*** 'to fall down'; *dar* 'to give' *vs. dar**se*** 'to occur/exist/initiate/carry out (among other meanings)'; *despertar* 'to wake (someone)' *vs. despertar**se*** 'to wake up'; *ir* 'to go' *vs. ir**se*** 'to go out/leave'; *ver* 'to see' *vs. ver**se*** 'to seem'. The remaining three verbs (i.e. *asomarse* 'to look (out)', *escaparse* 'to escape', *romperse* 'to break') do not induce the same change in meaning (e.g. *caer* 'to fall' vs. *caer**se*** 'to fall down') with *se*-marking; however, using *se* with these verbs gives special emphasis on the pivotal moment of a change of state.

Four verbs showed a high overall frequency of *se*-marking: *caerse* 'to fall down', *darse* 'to occur/exist/initiate/carry out (among other meanings)', *despertarse* 'to wake up', and *escaparse* 'to escape'. Although *darse* appeared only 8 times in L1 Korean speakers' written narrations, it was frequently used by the other two groups. Specifically, for all three groups, the *se*-marked form (*darse*) appeared exclusively in the phrase *darse cuenta* 'to realize', which might be evidence for the acquisition of this construction-specific use of *se* as a lexical chunk.

On the other hand, there are four verbs highlighted in bold in the above table: *asomarse* 'to look (out)', *irse* 'to go out/leave', *romperse* 'to break', and *verse* 'to seem'. These verbs show discrepancies in the frequency of *se*-marking between groups. The first verb, *asomarse*, was produced frequently (i.e. >10) only by the L1 Spanish speakers. The two non-native speaker groups produced *asomarse* between 2 and 3 times. However, this result does not mean that the L2 learners underused the clitic *se* because native speakers also employed the non-*se*-marked form in their writings, as shown in the examples below:

(60) a. *El perro, que **asomaba** también por la ventana, pierde el equilibrio y cae fuera de la casa.* [SPAN01]
'The dog, who was also **looking out** the window, loses his balance and falls outside the house.'

b. *Cuando los dos **asomaban** por la ventana para ver por dónde se habría ido la ranita, el perro se cayó de la ventana y se rompió el jarro.* [KOR10]
'When the two **looked out** the window to see where the frog had gone, the dog fell from the window and broke the jar.'

The *se*-marked form of the second verb, *ir*[30], was highly favored by L1 English speakers. In order to determine whether this was a pattern of overuse of *se*, I calculated and compared the frequency of *se*-marked tokens to non-*se*-marked tokens of *ir*. It was revealed that L1 English speakers used the *se*-marked form 40% of the time (10 out of 25 tokens of *ir*), L1 Korean speakers did so 50% of the time (6 out of 12 tokens of *ir*), and L1 Spanish speakers showed a 24% rate of *se*-marking (4 out of 17 tokens of *ir*). Therefore, it is difficult to conclude that the higher use of *irse* by non-native speakers is due to the overuse of the clitic *se*. Moreover, it is

[30] It would be worth noting that *irse* is considered a highly frequent collocation —especially in the preterite (e.g. *Se fue* 'S/he left')—, and, as such, we might expect a higher use of *se* with this verb.

shown in the following examples that non-native speakers use the *se*-marked form (*irse*) when its intended meaning appears to be 'to go out/leave':

(61) a. *En la tercera, el niño y su perro (que se encuentra con la cabeza metida en el mismo bote) se han sumido a la ventana para ver por dónde la rana* ***se había ido***. [ENG21]
'In the third (image), the boy and his dog (who is with his head stuck in the same jar) have plunged to the window to see where the frog **had left**.'

b. *Después el niño y el perro se despertaron y descubrieron que ya* ***se fue****.* [KOR05]
'Then the boy and the dog woke up and discovered that (the frog) **had** already **left**.'

The remaining two verbs, *romper* and *ver*, were *se*-marked frequently only by native speakers of Spanish in that they occurred 10 or more times. L1 Spanish speakers marked 10 out of 12 tokens of *romper* with *se*. It is worth mentioning that the L1 Korean group preferred the *se*-marked form of the verb *romper*, as 9 out of 10 tokens were *se*-marked. The L1 English group showed the opposite pattern: 3 out of 10 tokens of *romper* were *se*-marked. This result suggests that English speakers prefer traditionally transitive constructions (without *se*) while Korean and Spanish speakers make use of middle constructions, as shown in the examples below:

(62) a. *El perro se cayó de la ventana abierta y* ***rompió*** *el bote.* [ENG08]
'The dog fell from the open window and **broke** the jar.'

b. *Entonces* ***se rompió*** *el frasco de vídrio.* [KOR26]
'Then the glass jar **got broken**.'

c. *Al caer,* ***se rompe*** *el frasco en el que tenía metida su cabeza.* [SPAN25]
'Upon falling, the jar in which his head was stuck **gets broken**.'

Finally, for those instances in which the verb *ver* was *se*-marked, it occurred almost exclusively by L1 Spanish speakers (13 out of 33 tokens of *ver* were *se*-marked). In contrast, the L1 English never *se*-marked *ver* (21 tokens), and the L1 Korean group *se*-marked *ver* only once out of 18 times. This result shows that native speakers were more likely to use passive *se*

constructions and/or used *verse* as copulative verb ('to seem'), whereas non-native speakers favored the transitive use of *ver* 'to see' in active voice:

(63) a. *El niño y su perro, que siempre dormía con él,* ***vieron*** *que la rana no estuvo allí.* [ENG01]
'The boy and his dog, who always slept with him, **saw** that the frog was not there.'

b. *Le abrazó y miró todo el cuerpo y, afortunadamente, no* ***vio*** *ni una gota de sangre.* [KOR02]
'He hugged him and looked around his body and, fortunately, he did not **see** even a drop of blood.'

c. *El perro ya no tiene la cabeza en el bote y* ***se ve*** *que este bote está roto en el suelo.* [SPAN30]
'The dog no longer has his head in the jar and it **seems** that this jar is broken on the ground.'

In general, all participant groups performed similarly in this production task data in that the rate of *se*-marking was almost identical (i.e. ranging from 19% to 24%). However, with certain verbs, it was shown that non-native speakers (especially L1 English speakers) preferred non-*se*-marked forms and made use of transitive constructions when native speakers favored *se*-marked forms. Although the two L1 learner groups in the present study did not demonstrate noteworthy overuse or underuse of *se* in their narration, sometimes the use may be clearly ungrammatical because there is a lexical difference of meaning in the two forms (e.g. *ver* 'to see' *vs.* ***verse*** 'to seem'). On the other hand, sometimes the use is optional, depending on whether the speaker/writer wants to emphasize the suddenness or the moment of change (e.g. *caer* 'to fall' *vs.* ***caerse*** 'to fall down'). Given this situation, the present study does not intend to make any grammatical judgments regarding the use of *se* or not for each verb. It attempts to explain the different usages by identifying certain objective linguistic and extralinguistic factors that can be identified as contributing to the use/non-use of *se* for each group.

As such, to determine which factors influenced the use of the Spanish clitic *se*, a statistical analysis in Rbrul was performed for a mixed-effects variable rule analysis. The statistical findings inform our understanding of the similarities and discrepancies between the factors that favored the use of *se* for each participant group. The table below summarizes the linguistic and extralinguistic variables included in the statistical analyses.

Table 18. Independent variables analyzed for Task 1

Independent Variables	**Participant Group(s)**	**Variables**
Linguistic variables	All	• Transitivity • Telicity • Tense-Aspect-Mood • Subject Form • Subject Position
Extralinguistic variables	All	• Age • Gender • Education Level
	ENG & KOR (Learner groups)	• Spanish Proficiency • Age of Onset • Experience living in Spanish-speaking countries • Learning Environment
	KOR only	• English Proficiency
	SPAN only	• Country of Origin

The reason for choosing these specific variables, as discussed in detail in the previous chapter, is mainly based on the predictions that certain conditions will (dis)favor the *se*-marking. More

specifically, transitive verbs are expected to favor *se*-marking since L2 learners are first exposed to *se* with daily routine verbs, which are mostly transitive verbs. My prediction on the subject forms derives from the nature of the task itself. That is, given that the protagonists of the story perform similar actions at the same time, I assumed that null subject form will be *se*-marked more than any other subject forms (e.g. *Se despertaron* '(The boy and the dog) woke up'). As for the subject position, since L2 learners would acquire passive *se* constructions relatively early on (cf. Bailey 2013, Guillén-Solano 2015), post-verbal subjects will favor the *se*-marking because VS word order is more common in passive *se* constructions. Regarding Tense-Aspect-Mood, I follow the findings of Aaron & Torres Cacoullos (2006), which confirmed that past tense forms will favor the *se*-marking. Also, based on the fact that Accomplishment verbs which bear punctuality and Achievement verbs which are compatible with anticausative constructions are more likely to be *se*-marked, it is expected that telic verbs would favor *se*-marking. On the other hand, regarding extralinguistic variables, L2 learners with higher level of Spanish proficiency and more experience living in Spanish-speaking countries are more likely to perform in a native-like way, given that sufficient input is required for L2 learners to acquire various *se* constructions (cf. Tremblay 2005).

5.2.1 L1 English speakers' Task 1 results

The results of the statistical tests for the first participant group, the L1 English speaker group, are summarized in Table 19. Before analyzing the data, I ran a stepwise variable selection procedure to identify which variables to include in the model. It suggested the inclusion of *Tense-Aspect-Mood, Age of Onset, Length of Residence* (in Spanish-speaking countries), and an interaction between *Subject Position* and *Telicity* as independent variables. Additionally, results

of the stepwise variable selection procedure showed that including *Participant ID* and *Verb* as random effects would produce the best logistic regression model for the data. These factor groups that were expected to be significant predictors of *se*-marking are listed in the table with respective factor weights, counts, percentages, and *p*-values.

In terms of factor weight, a factor weight approaching 1 favors a factor while a weight closer to 0 is said to disfavor it. Based on this, in the present study, a factor weight above 0.5 is interpreted as favoring the use of the clitic *se*. In other words, the closer to 1 the factor weight is, the higher the likelihood is that the verb is *se*-marked. If the factor weight is closer to 0, there is a disfavoring of the usage of the clitic *se*. When a constraint's factor weight is close to 0.5, it indicates that the constraint in question neither favors nor disfavors *se*-marking. As for the count, each number presents the total token frequency in each context, including both *se*-marked and non-*se*-marked verb usages. However, the percentage only counts the proportion of *se*-marked verbs in the data for each context. Therefore, while the sum of counts for each factor group equals 663, which is the total number of tokens by the L1 English speaker group, the percentage does not add up to 100%. For example, in Table 19 below, the first constraint 'Long (> 8.5 mos.)' under the factor group *Length of Residence* shows the count of 235 and the percentage of 28.9%. It means that the participants who belong to the group of 'Long (> 8.5 mos.)' produced 235 tokens (both *se*-marked and non-*se*-marked) and 28.9% of these 235 tokens was *se*-marked. Finally, the *p*-value of each factor group indicates the statistical significance of each group. If the *p*-value is less than .05, the factor group is statistically significant. Conversely, if there is a factor group with a *p*-value higher than .05, the statistical analysis of this factor group lacks significance.

Table 19. Summary of L1 English group's Rbrul results (factors expected to predict *se*-marking)

Model: SE ~ TAM + Onset + LOR + SubPos:Tel + (1 | ID) + (1 | Verb)				
Factor Group	**Factor Weight**	**Count**	**Percentage**	***p*-value**
Length of Residence				3.99e-05
Long (> 8.5 mos.)	0.781	235	28.9%	
None (0 mos.)	0.379	210	18.6%	
Short (< 8.5 mos.)	0.314	218	14.7%	
Subject Position				N/A
Preverbal	0.821	305	31.8%	
Postverbal	0.511	32	12.5%	
Null	0.373	208	11.5%	
N/A	0.259	118	11.9%	
Telicity				N/A
Telic	0.72	287	34.1%	
Atelic	0.28	376	10.9%	
SubPos:Tel				7.46e-03
Preverbal:Telic	0.755	143	51%	
N/A:Atelic	0.671	63	7.94%	
Null:Atelic	0.579	127	8.66%	
Postverbal:Atelic	0.523	24	4.17%	
Postverbal:Telic	0.477	8	37.5%	
Null:Telic	0.421	81	16%	
N/A:Telic	0.329	55	16.4%	
Preverbal:Atelic	0.245	162	14.8%	
Tense-Aspect-Mood				0.0243
Preterite	0.795	221	33%	
Other	0.731	188	18.6%	
Present	0.486	186	16.1%	
Imperfect	0.0915	68	1.47%	
[Age of Onset]				0.0979

Adult (17-28 yrs.)	0.663	121	22.3%	
Adolescent (14-16 yrs.)	0.542	301	19.3%	
Child (1-13 yrs.)	0.3	241	22.4%	

As shown in Table 19, L1 English speakers who have experience living in Spanish-speaking countries for longer than 8.5 months favor *se*-marking with the factor weight of 0.781. English speakers with no experience or a length of residence below 8.5 months showed a pattern of disfavoring the use of *se* with the factor weight of 0.379 and 0.314, respectively.

Although the *p*-value for the individual factor groups *Subject Position* and *Telicity* is not given in the statistics, I provide a brief overview of the findings, which will result in a more comprehensive depiction of the results. As for *Subject Position*, Preverbal was the only constraint that showed a factor weight close to 1 (0.821), followed by Postverbal subjects with a neutral effect (0.511).[31] See examples below:

(64) Preverbal subject with *se*-marking:
*Una noche, el niño **se durmió**, pero la rana **se escapó** de su bote.* [ENG01]
'One night, the boy **fell asleep/slept**, but the frog **escaped** from its jar.'

(65) Postverbal subject with *se*-marking:
*Cuando **se despierta** el chico, ve que la rana desapareció, y va a la ventana con su perro para buscar la rana.* [ENG30]
'When the boy **wakes up**, he sees that the frog disappeared, and he goes to the window with his dog to look for the frog.'

[31] This result is somewhat different from that of Elliot (1995) discussed earlier in 3.3.3, which reported that L2 learners produced more grammatical sentences when the subject was in postverbal position. This discrepancy might be attributable to various reasons: nature of the dataset (written narration *vs.* written conversational texts), level of proficiency (from intermediate to near native *vs.* only intermediate), and researcher's grammatical judgment (not present *vs.* present).

For *Telicity*, it is indicated that Telic verbs favored the use of *se* with the factor weight of 0.72. That telic verbs favor *se*-marking is consistent with the prediction that was discussed in 4.4. That is, based on the punctuality of Accomplishment type verbs and Achievement type verbs' compatibility with anticausative structures, it was expected that telic verbs would favor the use of *se*. This is illustrated in the following example, where an Achievement type verb (i.e. *poner* 'to put') is marked with *se*:

(66) Telic verbs with *se*-marking:
*No sé por qué, pero el perro **se puso** la cabeza dentro de una jarra.* [ENG02]
'I don't know why, but the dog **put** its head into a jug.'

In contrast, Atelic verbs, as shown in the examples below, strongly disfavor *se*-marking, with the factor weight of 0.28:

(67) Atelic verbs without *se*-marking:
*Mientras **dormiá**, una rana se escapó de un bote en el suelo de la habitación del niño.* [ENG03]
'While he **was sleeping**, a frog escaped from a jar on the floor of the boy's room.'

Nonetheless, there were instances of atelic verbs with *se*-marking, most likely due to the confusion among L2 learners regarding the difference in meaning of *dormir* 'to sleep' and *dormir**se*** 'to fall asleep':

(68) Atelic verbs with *se*-marking:
*Una noche, el niño **se durmió**, pero la rana se escapó de su bote.* [ENG01][32]
'One night, the boy **fell asleep/slept**, but the frog escaped from its jar.'

The interaction between the two aforementioned factor groups presents a *p*-value of 7.46e-03, which means that this interaction is statistically significant. Based on the factor weights, the

[32] This is repeated from the previous example, (64).

combination of Preverbal subjects with Telic verbs, as illustrated below, strongly favored the use of *se* with the factor weight of 0.755:

(69) Preverbal subjects with *se*-marked Telic verbs:
Mientras llamaba el niño a Pepe la rana, el perro del niño ***se cayó*** *de cabeza de la ventana a la tierra.* [ENG05]
'While the boy was calling Pepe the frog, the boy's dog **fell** headlong from the window to the ground.'

The factor weight of 0.671 of the following combination, 'N/A:Atelic', means that atelic verbs that were used in non-finite forms (e.g. infinitive, gerund, past participle) favored *se*-marking. See example below:

(70) Atelic verbs in the Non-finite form with *se*-marking:
Le señala al perro de ***quedarse*** *callado, y los dos miran al otro lado del tronco*. [ENG30]
'He signals to the dog **to keep** quiet, and the two look over to the other side of the log.'

Tense-Aspect-Mood indicated that both Preterite and Other verb forms favored conditions of *se*-marking with the factor weights of 0.795 and 0.731, respectively. Regarding 'Other', Conditional (*hablaría*), Future (*hablará*), Subjunctive (*hable; hablara, hablase*), and Non-finite (Infinitive (*hablar*), Gerund (*hablando*), Past Participle (*hablado*)) forms were collapsed into this category due to their low token frequency and/or inapplicability for the current analysis. Meanwhile, Imperfect forms strongly disfavored *se*-marking, a pattern opposite from that of Aaron & Torres Cacoullos (2006). Their study showed that both Preterite and Imperfect favored the use of *se* with the verb *salir* without any significant difference between the two verbal forms. This divergence could be due to methodological differences, such as the participant group under study (Spanish speakers *vs.* English speakers), type of verbs (only *salir* vs. all verbs), and the nature of the data set (corpus data *vs.* survey data).

The last factor group, *Age of Onset*, presents a high *p*-value, .0979, although the best logistic regression model for the data predicted by the stepwise variable selection procedure included it. Based on the factor weight, L1 English speakers who started learning Spanish after the age of 17 favored *se*-marking with a factor weight above 0.5 (0.663). The youngest group showed a clear pattern of disfavoring (0.3), while those who belong to the adolescent group neither favored nor disfavored the use of *se* (0.542).

Thus, for the L1 English speakers, the factors that favor *se*-marking included (i) 8.5 months or longer time spent in Spanish-speaking countries (factor weight 0.781); (ii) preverbal subject position (factor weight 0.821); (iii) telic verbs (factor weight 0.72); (iv) preterite and non-finite forms (factor weight 0.795 and 0.731, respectively); (v) those who started learning Spanish at the age of 17 or older (factor weight 0.663); and (vi) the interaction between preverbal subjects and telic verbs (factor weight 0.755) and that of non-finite phrases that lack the existence of subject forms and atelic verbs (factor weight 0.671).

5.2.2 L1 Korean speakers' Task 1 results

The best logistic regression model for the L1 Korean speakers' data suggested by the stepwise variable selection procedure included fewer indicators than the L1 English speaker group's model. Table 20 below presents the results of the statistical analysis:

Table 20. Summary of L1 Korean group's Rbrul results (factors expected to predict *se*-marking)

Model: SE ~ Tel + TAM + Onset + (1 | ID) + (1 | Verb)				
Factor Group	**Factor Weight**	**Count**	**Percentage**	***p*-value**
Telicity				1.79-04
Telic	0.77	302	37.10%	

Atelic	0.23	325	6.77%	
Tense-Aspect-Mood				0.0233
Present	0.613	118	23.7%	
Preterite	0.585	248	30.6%	
Imperfect	0.499	71	7.04%	
Other	0.31	190	13.2%	
[Age of Onset]				0.0981
Adolescent (14-16 yrs.)	0.63	254	25.6%	
Adult (17-28 yrs.)	0.48	326	19%	
Child (1-13 yrs.)	0.389	47	14.9%	

Three independent variables, *Telicity, Tense-Aspect-Mood,* and *Age of Onset*, were included along with the two random factors, *Participant ID* and *Verb*. In line with previous results from the L1 English speakers' data, Telic verbs favored *se*-marking, and Atelic verbs disfavored the use of *se*, as in the following examples:

(71) Telic verbs with *se*-marking:

*La rana **está escapándose** de la botella mientras la niña duerme.* [KOR01]

'The frog **is escaping** from the bottle while the girl sleeps.'

(72) Atelic verbs without *se*-marking:

*Era una noche tan oscura que hasta la luna **estaba durmiendo** cuando el niño estaba soñando en su cama con su cachorro.* [KOR02]

'It was such a dark night that even the moon **was sleeping** when the boy was dreaming in his bed with his puppy.'

There are, however, slight discrepancies in the results between L1 English and L1 Korean speakers when it comes to *Tense-Aspect-Mood*. It was found that for the L1 Korean speakers, the Present tense showed a favoring condition (factor weight 0.613), followed by the Preterite which is also favored but slightly less so (factor weight 0.585), while Imperfect and Other forms

disfavored *se*-marking. Specifically, *se*-marking occurred in 23.7% of Present forms, and 30.6% in Preterite forms. When these two forms are grouped together, over 50% of the verbs were *se*-marked. Some examples are illustrated below:

(73) Present tense with *se*-marking:
*Afortunadamente el niño salva a su perro pero el frasco **se rompe**.* [KOR23]
'Fortunately the boy saves his dog but the jar **breaks**.'

(74) Other TAM without *se*-marking:
*En el piso, había una jarra de cristal, en la que una rana trataba de **salir** porque el niño había captado y guardado la tarde anterior.* [KOR02]
'On the floor, there was a glass jug, in which a frog was trying **to get out** because the boy had captured and kept (it) the previous afternoon.'

The last factor group, *Age of Onset,* with a high *p*-value (.0981) also diverges from the pattern shown in the L1 English speakers' results. In the case of the L1 Korean speakers, only those who started learning Spanish between the ages of 14 and 16 favored *se*-marking. However, the youngest group showed a disfavoring effect on the use of *se*, which is congruent with the L1 English speaker group's data.

Consequently, the factors that favor *se*-marking for the L1 Korean speakers consisted of the following: (i) telic verbs (factor weight 0.77); (ii) present and preterite forms (factor weight 0.613 and 0.585, respectively); and (iii) those who started learning Spanish between the age of 14 and 16 (factor weight 0.63).

5.2.3 L1 Spanish speakers' Task 1 results

Finally, the statistical analysis of the native speakers' data is summarized below in Table 21. The model formula for the analysis included *Telicity, Tense-Aspect-Mood,* and *Age* as individual factors with *Participant ID* and *Verb* as random effects.

Table 21. Summary of L1 Spanish group's Rbrul results (factors expected to predict *se*-marking)

Model: SE ~ Tel + TAM + Age + (1 | ID) + (1 | Verb)				
Factor Group	**Factor Weight**	**Count**	**Percentage**	***p*-value**
Age Group				7.16e-03
18-24 yrs.	0.792	22	79.2%	
25-30 yrs.	0.541	211	54.1%	
41+ yrs.	0.421	303	42.1%	
31-40 yrs.	0.235	237	23.5%	
Telicity				0.0366
Telic	0.808	344	38.1%	
Atelic	0.192	429	7.23%	
[Tense-Aspect-Mood]				0.0723
Present	0.665	310	26.5%	
Preterite	0.477	175	30.3%	
Other	0.434	219	11.4%	
Imperfect	0.418	69	2.9%	

An interesting finding from the native speakers' data is that *Age Group* showed the lowest *p*-value among all the factor groups. The youngest two groups (18-24 yrs. and 25-30 yrs.) favored *se*-marking with a factor weight of 0.792 and 0.51, respectively, while the other two older groups (31-40 yrs. and 41+ yrs.) used more non-*se*-marked forms with factor weights below 0.5.

The second lowest *p*-value (.0366) was for the factor group *Telicity* which showed results consonant with the previous two participant groups: Telic verbs favored the use of *se*, and Atelic verbs disfavored it, as in the examples below. The factor weight for Telic verbs was even greater, 0.808, for this group. From this, it can be interpreted that there is a stronger favoring pattern for *se*-marking with telic verbs among native speakers than non-native L2 learners.

(75) Telic verbs with *se*-marking:

Internados en el bosque, El Niño [*sic*] *y el perro* ***se detienen*** *a descansar alrededor de un tronco hueco.* [SPAN01]

'Deep in the forest, the boy and the dog **stop** to rest around a hollow log.'

(76) Atelic verbs without *se*-marking:

Un niño y su perro ***duermen*** *plácidamente en la cama mientras que una rana que estaba en un bote de cristal se escapa y sale de la habitación por la ventana, que estaba abierta.* [SPAN02]

'A boy and his dog **sleep** peacefully in bed while a frog that was in a glass jar escapes and leaves the room through the window, which was open.'

The verb *dormir* in example (76) appears non-*se*-marked while it is *se*-marked in example (64) from a L1 English speaker, repeated below for convenience:

(77) *Una noche, el niño* ***se durmió****, pero la rana se escapa de su bote.* [ENG01]

'One night, the boy **fell asleep/slept**, but the frog escaped from its jar.'

In addition to the pattern that atelic verbs like *dormir* disfavored *se*-marking, the use of *se* in this context is semantically infelicitous, as can be inferred from the English translations of the examples: *dormir* 'to sleep' *vs. dormir****se*** 'to fall asleep'. This contrast implies that L2 learners may have memorized certain *se*-marked verbal forms as lexical chunks without a clear understanding of which function the clitic *se* performs in a given situation. For example, since *dormirse* is taught as one of 'reflexive daily routine verbs' along with *despertarse* 'to wake up', *ducharse* 'to take a shower', and many others (cf. the textbook *Experience Spanish* discussed in 3.4), it is likely that L2 learners fail to distinguish the semantic and/or pragmatic differences signaled by the use of clitic *se* in such verbs. Based on this assumption, it is also possible to explain the pattern of native speakers' stronger preference for *se*-marking on telic verbs (factor weight 0.808) than that of L1 English and L1 Korean speakers (factor weight 0.72 and 0.77, respectively).

The last factor group, *Tense-Aspect-Mood* has a high *p*-value of .0723. Here, only the Present tense favored the use of *se* (factor weight 0.665). The second highest factor weight was found for Preterite forms, although it is below the 0.5 significance level. The order of the factor weights in this factor group for native speakers (i.e. highest to lowest: Present > Preterite > Other > Imperfect) is more similar to that of the L1 Korean speaker group (i.e. Present > Preterite > Imperfect > Other), compared to the L1 English speaker group's data (i.e. Preterite > Other > Present > Imperfect). What is shared by all three groups is that the Imperfect strongly disfavors *se*-marking. This finding is somewhat surprising given that Aaron & Torres Cacoullos' (2006) results showed that the distinction of the favoring effect was not between Preterite and Imperfect, but between past and non-past tenses. In fact, these authors predicted that the *se*-marked forms should be favored in Preterite rather than in Imperfect, given the association between perfectivity and *se*-marked dynamic predicates (Bybee et al. 1994:92, cited in Aaron & Torres Cacoullos 2006). Thus, it can be concluded that, while Aaron & Torres Cacoullos' (2006) corpus analysis did not confirm the assumption that perfective aspect expressed by Preterite is strongly associated with *se*-marking, the picture book written narration data in the present study appears to reflect this prediction.

5.2.4 Task 1 results for all participant groups combined

With the individual results of each participant group, I now compare and contrast them in order to get a more in-depth description of how native and non-native Spanish speakers' patterns of *se*-marking is determined. When all three groups were analyzed together, the constraint hierarchies for each factor group were calculated as described in Table 22 below. Due to the fact that Rbrul analysis is based on the in-group comparison, its results for all three L1 groups

combined are not equivalent to the individual group's results. In the analysis, *Participant ID* and *Verb* were included as random effects, *Telicity* was included as an individual factor, and the interaction between *L1* and *Tense-Aspect-Mood* was included in the model.

Table 22. Summary of the Rbrul results for all groups (factors expected to predict *se*-marking)

Model: SE ~ Tel + L1*TAM + (1 | ID) + (1 | Verb)				
Factor Group	**Factor Weight**	**Count**	**Percentage**	***p*-value**
Telicity				7.03e-04
Telic	0.708	933	36.5%	
Atelic	0.292	1130	8.32%	
L1				N/A
Korean	0.561	627	21.4%	
Spanish	0.49	773	21%	
English	0.449	663	21%	
Tense-Aspect-Mood				N/A
Present	0.609	614	22.8%	
Preterite	0.6	644	31.4%	
Other	0.401	597	14.2%	
Imperfect	0.39	208	3.85%	
L1:TAM interaction				0.0523
English:Other	0.66	188	18.6%	
Korean:Imperfect	0.641	71	7.04%	
Spanish:Present	0.584	310	26.5%	
English:Preterite	0.584	221	33%	
Spanish:Imperfect	0.514	69	2.9%	
Korean:Present	0.507	118	23.7%	
Korean:Preterite	0.488	248	30.6%	
Spanish:Other	0.474	219	11.4%	
Spanish:Preterite	0.428	175	30.3%	

English:Present	0.409	186	16.1%	
Korean:Other	0.364	190	13.2%	
English:Imperfect	0.347	68	1.47%	

Regarding *Telicity*, Telic verbs strongly favored *se*-marking with a factor weight of 0.708, and Atelic verbs tended to be non-*se*-marked, showing a factor weight of 0.292. *Telicity* was the factor group that demonstrated the most prominent favoring effect across all three participant groups. The pattern that Telic verbs favor *se*-marking follows the prediction based on previous literature (cf. Alexiadou 2012) and is perhaps attributable to the fact that speakers use the clitic *se* to describe spontaneous events or to convey emphasis on the pivotal moment of change of state, situations which are expressed more with telic verbs than atelic verbs. Furthermore, that *se* is employed to mark dynamicity/energeticness of the event by the participants in the present study is congruent with Soto & Muñoz's (2000) findings (cf. 3.2), which revealed that Spanish-speaking children use *se* to emphasize the pivotal moment of change of state.

Before investigating the findings of the interaction between *L1* and *Tense-Aspect-Mood*, the results of each factor group will be discussed briefly. As shown above, the L1 of the speaker is neither a favoring nor disfavoring factor. Factor weights for each L1 group were not significantly different, ranging around a 0.5 significance level. When it comes to the percentage of *se*-marked verbs across participants' L1, the congruence of the results is even more apparent: of the L1 Spanish and L1 English speaker group data, the percentage of *se*-marked verbs was the same (21%). The L1 Korean speaker group's rate of *se*-marked verbs was only slightly higher (21.4%) with the difference of 0.4%.

The findings for *Tense-Aspect-Mood* align with those of the individual group analyses which reported the Imperfect's disfavoring effect on *se*-marking. Although the order of the favoring effect between Present, Preterite, and Other was inconsistent across the groups, the overall generalization of the three groups was that Present and Preterite showed a favoring effect on *se*-marking, while Other did not.

The stepwise variable selection procedure suggested the interaction between *L1* and *Tense-Aspect-Mood* be included in the analysis despite its relatively higher *p*-value (.0523). The cross tab shown in Table 23 below is a contingency table of the distribution of *se*-marking when *L1* and *Tense-Aspect-Mood* are concerned. The order of the conditions in the column (i.e. TAM-Other > TAM-Imperfect > TAM-Present > TAM-Preterite) reflects the hierarchy of factor weights of the first four interactions: English:Other (0.66), Korean:Imperfect (0.641), Spanish:Present (0.584), and English:Preterite (0.584), respectively. According to these numbers, it was most likely that L1 English speakers would produce *se*-marked verbs in non-Present, non-Preterite, and non-Imperfect Tense-Aspect-Mood. The second most likely L1 and TAM interaction of *se*-marking was that of L1 Korean speakers and Imperfect forms, followed by L1 Spanish speakers with Present tense constructions and L1 English speakers with Preterite forms.

Table 23. Distribution of *se*-marking according to *L1* and *Tense-Aspect-Mood*

	TAM - Other						TAM - Imperfect						TAM - Present						TAM - Preterite					
	ENG		KOR		SPAN		ENG		**KOR**		SPAN		ENG		KOR		**SPAN**		**ENG**		KOR		SPAN	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
se-marked	**35**	**19**	25	13	25	11	1	1	**5**	**7**	2	3	30	16	28	24	**82**	**27**	**73**	**33**	76	31	53	30
Non-*se*-marked	153	81	165	87	194	89	67	99	66	93	67	97	156	84	90	76	228	73	148	67	172	69	122	70

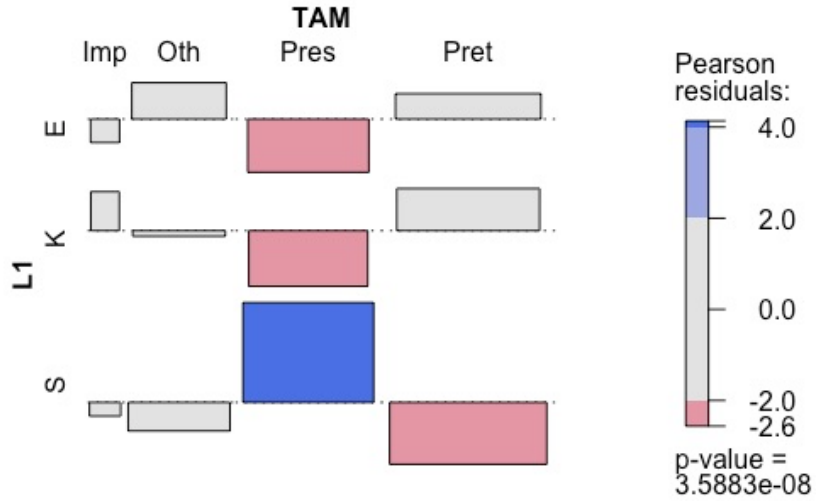


Figure 7. Association plot of *se*-marking according to *L1* and *Tense-Aspect-Mood*

Figure 7 above is a visualization of the distribution shown in Table 23 in the form of an association plot. Here, the wider surfaces of Present and Preterite reflect the relative magnitude of their values, which means that there were more tokens in these two forms than of Imperfect and Other forms. The blue and red colors indicate how much higher or lower the observed value is than the expected value if the data were random. The only L1 and TAM interaction that is colored blue in Figure 7 is between L1 Spanish speakers and Present tense. This result seems to reflect native speakers' use of the historical present or narrative present (RAE 2009:437), a technique commonly used in written texts which narrates events that occurred in the past with present tense structures. Meanwhile, 'English:Present', 'Korean:Present', and 'Spanish:Preterite' were the only interactions indicated in red. From this result, it can be seen that these three interactions are most likely to avoid *se*-marking in written production.

It should be pointed out that the Rbrul results in Table 22 does not align with the result demonstrated in Figure 7. This is because the model used for the Rbrul analysis included *Participant ID* and *Verb* as random effects while the latter did not. However, it was possible to obtain a general overview of the results, such that there is a stronger favoring effect for non-native speakers to use *se*-marked forms in Preterite, while native speakers of Spanish favored *se*-marking in Present tense.

In summary, the statistical analyses of the results of Task 1 revealed that there was not a drastic contrast between the three participant groups. Among eight factor groups which were applied to all three groups, *Telicity* was the only variable that indicated a statistically significant finding: that is, telic verbs strongly favored the use of *se* in these narratives while atelic verbs disfavored *se*-marking. It is surprising that *Transitivity* was not a significant factor that affects the use of *se* in the data. As traditional analyses assumed the primary function of *se* was to

detransitivize verbal argument structure, we would expect that in these narratives, the use of transitive verbs would favor *se*-marking when used intransitively. However, that transitivity was not a significant factor group in the current study leads us to believe that the traditional detransitivization approach is perhaps too limited. Specifically, its narrow view of the clitic *se*'s function fails to capture characteristics of various middle *se* constructions in which the conceptualization of the event is important, especially in the narration of different participants acting out different events of a fixed storyline.

Also, as mentioned above, *Tense-Aspect-Mood* demonstrated the divergence between these native and non-native speaker groups: a favoring condition was found in Present tense for the native speakers but in Preterite for the English- and Korean-speaking L2 learners. Based on these findings, it can be concluded that these L2 learners, at least at this point in their development of the target language, strongly associate *se*-marking with the conceptualization of delimitedness, which implies a certain degree of boundedness (i.e. initiation or completion) of an event expressed by Preterite and/or telic verbs.

5.3 Task 2 (Acceptability Judgment Task, AJT) Results — testing learners' knowledge of absolute *vs.* dynamic verbal events

As described in 4.4, the purpose of Task 2 was to investigate the participants' association between dynamic contexts and *se*-marking. Based on discussions from the cognitive perspective, it was expected that participants would choose the non-*se*-marked form in absolute contexts (78) and the *se*-marked form in dynamic contexts (79).

(78) After school, Laura went upstairs to her room to read. When it was time for dinner, her mom called her to come help set the table. Taking her book and still reading, ***Laura bajó por las escaleras y entró al comedor lentamente*** **'Laura went down the stairs and entered the dining room slowly'**.

(79) Leo and his little brother were playing on the seesaw. When they got tired, his brother slid carefully off the seesaw. But not Leo who does everything in a hurry.
Abruptamente, Leo se bajó del sube y baja de un brinco **'Abruptly, Leo jumped off the seesaw'.**

The opposite cases —favoring the *se*-marked form in absolute contexts and the non-*se*-marked form in dynamic contexts— were unexpected, although this does not necessarily mean that these responses were grammatically incorrect and/or pragmatically infelicitous.

In general, the results showed that the participants of all three groups preferred the non-*se*-marked form when context types were not taken into account. Upon contrasting the participants' performance according to the context types, it demonstrated greater resemblance to the expectation in absolute contexts (i.e. judging that non-*se*-marked forms sound more natural). In dynamic contexts, however, non-native speakers did not show as strong a preference for *se*-marked forms as native speakers. Also, this pattern of favoring *se*-marked forms in dynamic contexts was revealed more in L2 near native speakers' data, suggesting that the development of proficiency in target language leads to a more native-like behavior.

In what follows, I provide statistical analyses of separate participant groups and discuss their results. Also, the possible influence of the type of verb used in each question will be included in the discussion. Since the verb themselves were used for both absolute and dynamic contexts, it is of interest to find some relevance between certain characteristics of verbs (e.g. transitivity, semantic properties) with certain contexts.

5.3.1 L1 English speakers' Task 2 results

After counting the responses of the participants, it was shown that the non-*se*-marked forms were highly preferred by L1 English speakers with the rate above 50% across all proficiency levels. The least chosen response was the option stating that neither the *se*-marked nor the non-*se*-marked form is natural. These findings are reflected in the following table (Table 24), which shows the distribution of responses by L1 English speakers. It is followed by a mosaic plot, which graphically illustrates contrastive patterns of performance between each proficiency group.

Table 24. Distribution of responses by L1 English speakers (Task 2)

	Intermediate (*n*=7)		**Advanced (*n*=12)**		**Near native (*n*=13)**	
	#	%	#	%	#	%
Option A is more natural*	65	51.59%	116	53.70%	121	51.71%
Option B is more natural**	29	23.02%	60	27.78%	62	26.50%
Both A and B are natural	3	2.38%	29	13.43%	33	14.10%
Neither A nor B is natural	3	2.38%	0	0%	5	2.14%
I don't know	1	0.79%	5	2.31%	13	5.55%
No response	25	19.84%	6	2.78%	0	0%
Total	126	100%	216	100%	234	100%

* Option A always included non-*se*-marked forms.
** Option B always included *se*-marked forms.

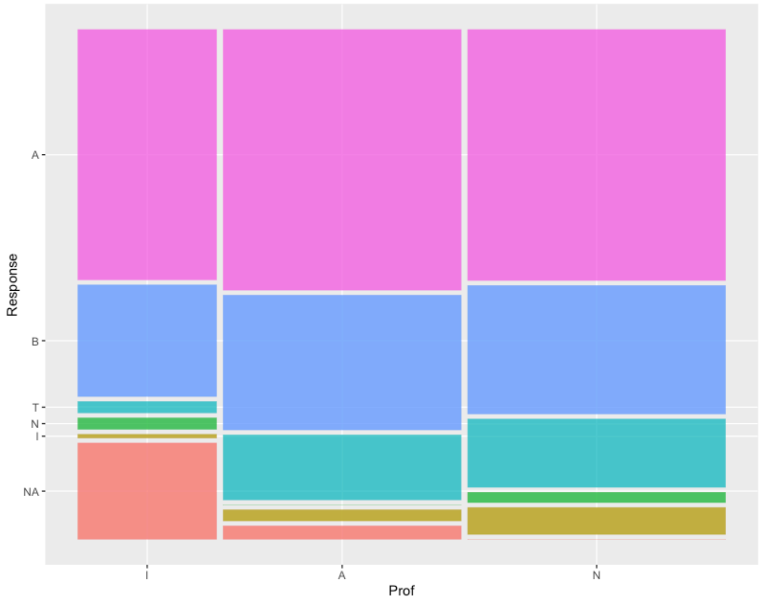


Figure 8. Mosaic plot of responses by L1 English speakers (Task 2)

Firstly, the greater widths of the boxes in the middle and right columns (i.e. A(dvanced) and N(ear native)) indicate that there were more responses, hence more participants, that belonged to the Advanced and Near native groups than the Intermediate group. What the heights of the boxes illustrate is the frequency of each response types (i.e. from top to bottom: Option A is more natural, Option B is more natural, Both A and B are natural, Neither A nor B is natural, I don't know, No response). As illustrated in the greater size of the three boxes on the top row of Figure 8, there is an overall preference for Option A, which consists of the non-*se*-marked verbal forms, across all proficiency levels regardless of absolute or dynamic/energetic context. Specifically, more than 50% of the L1 English speakers across all proficiency groups responded that Option A with non-*se*-marked verbs sounded more natural, while Option B was chosen approximately 25% of the time. This result is not compatible with the expected association between *se*-marking and dynamicity of the event, which would have yielded an identical frequency of the non-*se*-marked

Option A and the *se*-marked Option B, given that half of the contexts was absolute and the other half was dynamic.

Another interesting finding from the distribution of responses is the selection of the option 'Both A and B are natural'. While the advanced and near native speaker groups chose this option around 13.5% of the time (across all nine verbs, with the highest frequency for the verb *ir* 'to go'), this percentage for the intermediate group was only 2.38% (for the verb *subir* 'to go up'). Although this was not a preferred option across all proficiency levels, it was shown that intermediate level speakers disfavored it more strongly. Also, the intermediate group was the only group for which the percentage of 'No response' was relatively high (19.84%) for the verb *subir* 'to go up'.

To determine whether context type had an effect on *se*-marking, a comparison between responses in absolute and dynamic contexts is provided below. Again, the use of the clitic *se* was expected in dynamic contexts as they form part of middle constructions and focus on the dynamicity (suddenness, unexpectedness or pivotal moment of change) of the verbal event.

Table 25. Distribution of responses according to context and proficiency (L1 English group)

	Absolute						**Dynamic**					
	Intermediate		Advanced		Near native		Intermediate		Advanced		Near native	
	#	%	#	%	#	%	#	%	#	%	#	%
Option A is more natural	36	57%	81	75%	87	74%	29	46%	35	31%	34	29%
Option B is more natural	9	14%	13	12%	11	10%	20	32%	47	44%	51	44%
Both	2	3%	11	10%	13	11%	1	2%	18	17%	20	17%
Neither	1	2%	0	0%	1	1%	2	3%	0	0%	4	3%
I don't know	1	2%	1	1%	5	4%	0	0%	4	4%	8	7%
NA	14	22%	2	2%	0	0%	11	17%	4	4%	0	0%
Total	63	100%	108	100%	117	100%	63	100%	108	100%	117	100%

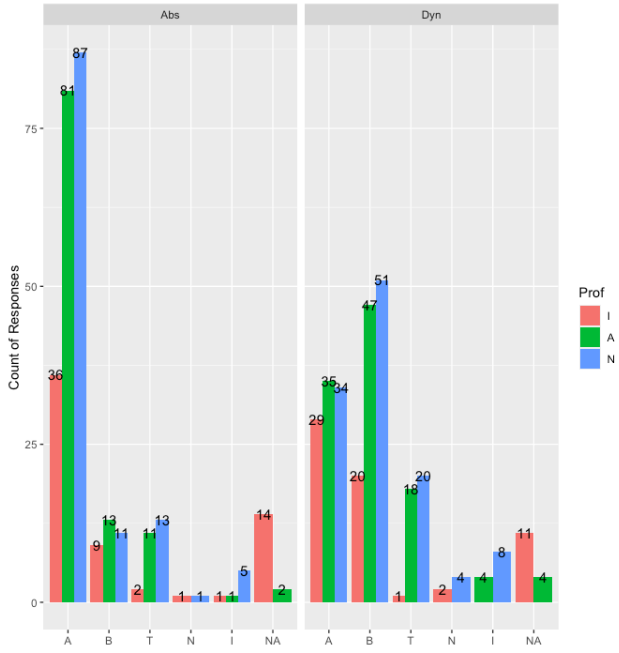


Figure 9. Frequency of responses according to context by L1 English speakers (Task 2)

The most noticeable difference in the analysis when context was taken into account was that the preference for Option A (the non-*se*-marked option) was stronger in absolute contexts than in dynamic ones, especially among advanced and near native speakers (75% and 74%, respectively). When the context yielded a dynamic reading, there was no noticeable favoring of the expected response (i.e. Option B, a preference for the *se*-marked verb). Furthermore, the intermediate group still preferred non-*se*-marked forms, even in dynamic contexts.

Although the advanced and near native participants chose Option B over Option A in dynamic contexts, 17% of them judged that both the non-*se*-marked and *se*-marked constructions sounded natural. This is comparable to the responses in absolute contexts, in which the percentage of choosing 'Both A and B are natural' was just around 10%. Hence, it can be concluded that English speakers —especially those with lower level of proficiency— have a weak association with *se*-marking and dynamicity of the context. Even when they acknowledged

that the situation was not absolute, some participants still judged non-*se*-marked constructions as acceptable. These patterns are illustrated in the table below:

Table 26. L1 English speakers' preference for (non-)*se*-marked option according to contexts

Absolute context: Among all of his friends, Luis is the fastest swimmer. In fact, he used to be a swimmer on his school's champion swimming team. He thinks it's because as a baby, he took swimming lessons.

A: *Luis aprendió a nadar cuando era bebé.* (Expected response)

B: *Luis se aprendió a nadar cuando era bebé.*

	Int		Adv		Near	
	#	%	#	%	#	%
Option A is more natural	**6**	**85.71%**	**11**	**91.67%**	**13**	**100%**
Option B is more natural	0	0%	1	8.33%	0	0%
Both A and B are natural	0	0%	0	0%	0	0%
Other responses*	1	14.29%	0	0%	0	0%

Dynamic context: Daniel has a math exam tomorrow and needs to study hard. He knows that if he can learn the multiplication tables completely, he'll get the A he needs. He studies all night long and by morning,

A: *Daniel había aprendido de memoria las tablas de multiplicación.*

B: *Daniel se había aprendido de memoria las tablas de multiplicación.* (Expected response)

	Int		Adv		Near	
	#	%	#	%	#	%
Option A is more natural	6	85.71%	5	41.66%	6	46.15%
Option B is more natural	**0**	**0%**	**3**	**25%**	**3**	**23.08%**
Both A and B are natural	0	0%	2	16.67%	3	23.08%
Other responses	1	14.29%	2	16.67%	1	7.69%

* The responses 'Neither A nor B is natural', 'I don't know', and 'No response' were combined for convenience.

Based on the description of how L1 English speakers with differing levels of Spanish proficiency responded in absolute and dynamic contexts, I scored and calculated their overall performance and compared it to the expected patterns. Table 27 below demonstrates the mean, median, and standard deviation values of each proficiency group's scores according to context type. Since the maximum score for each context was 18, a mean and median closer to 18 indicate that this performance was more similar to the expected pattern of association (see Chapter 4 for detailed scoring procedure).

Table 27. Mean and median scores of L1 English speakers' performance on Task 2

	Absolute (Max. score = 18)			**Dynamic (Max. score = 18)**		
	Mean	Median	*SD*	Mean	Median	*SD*
Intermediate	10.86	13	6.18	5.86	4	2.97
Advanced	14.42	15.5	3.32	9.33	8.5	4.44
Near native	14.38	15	2.60	9.38	11	4.61

The overall score for absolute contexts was higher than that of dynamic contexts. Also, while the advanced and near native groups performed similarly, intermediate speakers showed a drastic difference. Figure 10 below illustrates the median and interquartile ranges for the selected variables:

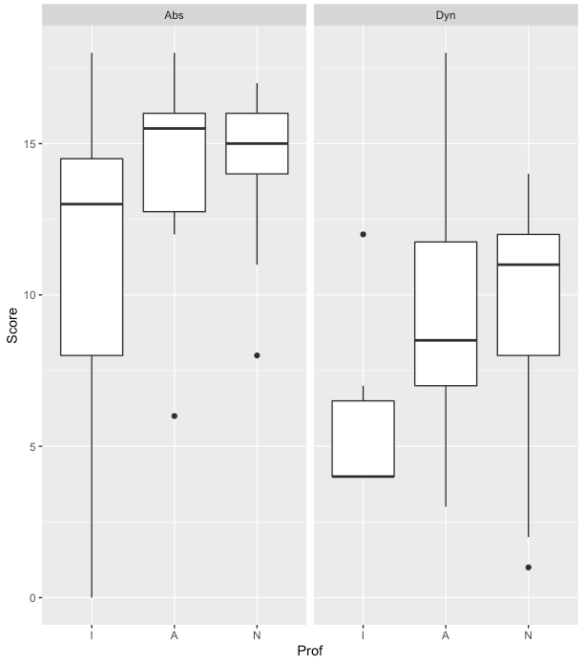


Figure 10. Box plot of scored performance of L1 English speakers (Task 2)

In addition to representing median scores per group, the box plot in Figure 10 also shows how variable each proficiency group's score was. For example, the intermediate group showed a score range of 0 to 18 in absolute contexts, which means that at least one participant of the group scored 0 while (an)other participant(s) scored 18. In comparison, the advanced group showed a score range of 12 to 18 in the same context, of which one outlier scored 6. This means that the performance of the advanced group in absolute contexts was not as varied as that of the intermediate group.

In order to provide further information regarding the statistical significance of these seemingly different patterns between proficiency levels and contexts, an analysis of variance (ANOVA) and a post-hoc analysis with Tukey's test were performed. First, a regression model was fit to perform an ANOVA on the data using the independent variables proficiency and context, as well as an interaction of the two. According to the analysis of variance, there was a

significant effect for context, $F(1, 58) = 24.59$, $p = 6.53e\text{-}06$. The effect of proficiency level was also significant, although not as much as context, $F(2, 58) = 4.12$, $p = .0212$. The interaction effect was not significant, $F(2, 58) = 0$, $p = .9992$.

Based on the ANOVA analysis, Tukey's Test was performed to further investigate differences between each proficiency group. Results of the Tukey's Test showed that the intermediate group differed significantly from both advanced and near native groups at $p < .05$, while the latter two groups were not significantly different from each other, $p = .99$. Below is graphic display of these results:

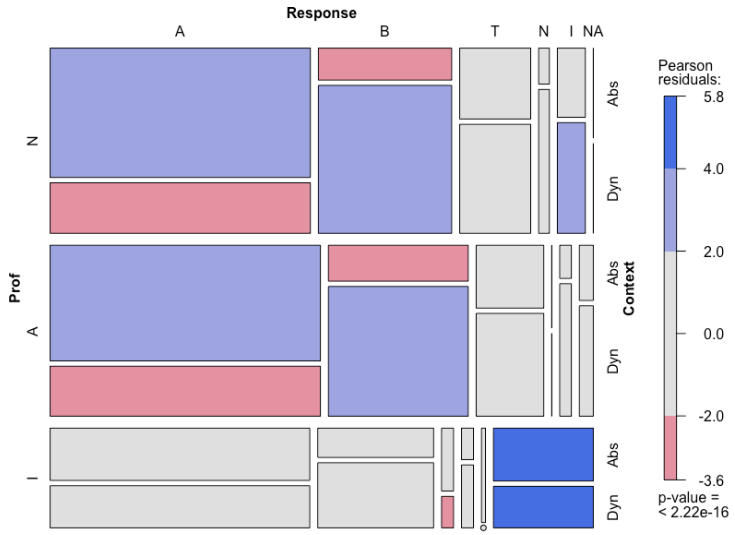


Figure 11. Significance of relevance between factors by L1 English speakers (Task 2)

What is demonstrated in Figure 11 is that the behavior of the advanced and near native groups were similar to each other. That is, in absolute contexts, both groups preferred the non-*se*-marked verb form in Option A (the favoring effect of which is indicated in blue) and disfavored the *se*-marked form in Option B (the disfavoring effect of which is indicated in red). In dynamic

contexts, their responses were the opposite: Option A was dispreferred (red) and Option B was favored (blue). This pattern lines up with the assumptions made in cognitive literature that the use of *se* is related to the dynamicity of the event (cf. Maldonado 1992, 2008, among others). However, L1 English speakers at the intermediate level of proficiency did not show any significant patterns in their responses concerning Option A or Option B. This group showed a significant probability of choosing not to answer ('No response') in both absolute and dynamic contexts. Also, there is a disfavoring effect on the option 'Both A and B are natural' when the given context is dynamic. It is possible to deduce from this result that the association between *se*-marking and context type (i.e. absolute *vs.* dynamic) has not yet been fully developed at this point in the acquisition process of the intermediate-level L1 English learners of Spanish. In other words, these learners have not yet acquired one of the basic semantic functions of the middle marker *se* which is to focus on the dynamic aspect of the verbal event, more specifically, on the pivotal moment of change (Maldonado 2000:159). Their understanding of *se* fails to extend beyond its strictly reflexive or detransitivizing uses, the ones most notably taught in the language classroom. However, in the more advanced levels, learners appear to understand the association of non-*se*-marking with absolute events and *se*-marking with dynamic ones. This is a considerable accomplishment in the second language acquisition process given that these types of association are rarely taught in the foreign language classroom and if so, are done so only implicitly.

Lastly, the following table (Table 28) shows partial results of the participants' verb form selection in dynamic contexts across all proficiency groups. The reason for not further analyzing responses in absolute contexts is that participants' performance in such contexts was in accordance with our expectations: they generally preferred the non-*se*-marked form. This may be

because they are selecting this variant as the default form. Only two unexpected patterns were observed in absolute contexts by the intermediate group: they preferred *se*-marked forms with *bajar* 'to go down' and *caer* 'to fall'. The preference for *se*-marking on these two verbs might be attributable to the fact that L2 learners are often taught these verbs with the *se* form in classrooms (i.e. *bajar**se*** 'to go/jump down', *caer**se*** 'to fall down'), so they are learning them as lexical chunks. On the other hand, in dynamic contexts, intermediate-level participants performed as expected with 3 out of 9 verbs (*bajar* 'to go down', *caer* 'to fall', and *dormir* 'to sleep') with which they preferred the *se*-marked variant. Therefore, only the 6 verbs that behaved unexpectedly in dynamic contexts will be discussed here.

Table 28. L1 English speakers' distribution of responses by verb in dynamic contexts (Task 2)

aprender	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is more natural	6	86%	5	42%	6	46%	17	53%
Option B is more natural	**0**	**0%**	**3**	**25%**	**3**	**23%**	**6**	**19%**
Both A and B are natural	0	0%	2	16.5%	3	23%	5	16%
Other responses*	1	14%	2	16.5%	1	8%	4	12%
beber	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is more natural	6	86%	2	16.5%	3	23%	11	34%
Option B is more natural	**0**	**0%**	**8**	**67%**	**7**	**54%**	**15**	**48%**
Both A and B are natural	0	0%	2	16.5%	1	8%	3	9%
Other responses	1	14%	0	0%	2	15%	3	9%
decidir	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is more natural	3	43%	10	83.5%	8	62%	21	66%
Option B is more natural	**2**	**28.5%**	**2**	**16.5%**	**1**	**8%**	**5**	**16%**

Both A and B are natural	0	0%	0	0%	2	15%	2	6%
Other responses	2	28.5%	0	0%	2	15%	4	12%
ir	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is more natural	4	57.5%	2	16.5%	2	15%	8	25%
Option B is more natural	**2**	**28.5%**	**4**	**33.5%**	**8**	**62%**	**14**	**44%**
Both A and B are natural	0	0%	5	42%	2	15%	7	22%
Other responses	1	14%	1	8%	1	8%	3	9%
saber	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is more natural	5	71.5%	8	67%	6	46%	19	59%
Option B is more natural	**0**	**0%**	**3**	**25%**	**1**	**8%**	**4**	**13%**
Both A and B are natural	0	0%	1	8%	5	38%	6	19%
Other responses	2	28.5%	0	0%	1	8%	3	9%
subir	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is more natural	3	43.5%	4	33.5%	5	38%	12	37%
Option B is more natural	**2**	**28.5%**	**5**	**42%**	**5**	**38%**	**12**	**37%**
Both A and B are natural	1	14%	1	8%	2	15%	4	13%
Other responses	1	14%	2	16.5%	1	8%	4	13%

* The responses 'Neither A nor B is natural', 'I don't know', and 'No response' were combined for convenience.

As indicated in Table 28, the *se*-marked variant (Option B) was selected less than 50% of the time in dynamic contexts, in the descending order of *beber* 'to drink' (48%) > *ir* 'to go' (44%) > *subir* 'to go up' (37%) > *aprender* 'to learn' (19%) > *decidir* 'to decide' (16%) > *saber* 'to know' (13%). *Se*-marking on *beber* 'to drink' was only problematic for intermediate-level speakers. That is, the advanced and near native groups preferred the *se*-marked form *beber**se*** 'to drink up' to any other responses. However, 6 out of 7 intermediate participants judged that the

non-*se*-marked form (Option A) *beber* 'to drink' sounded most natural in dynamic context. The table also indicates that the intermediate speakers generally preferred the non-*se*-marked variants in most cases.

The verb that corresponds to the second highest percentage of choosing the *se*-marked option (Option B) in dynamic context was *ir* 'to go'. This is because the near native group clearly preferred the *se*-marked form *irse* 'to go out/leave'. The non-*se*-marked variant (Option A) was preferred by the intermediate group, while the option 'Both A and B are natural' was favored slightly over the *se*-marked form (Option B) by the advanced group. On the other hand, for the verb *subir* 'to go up', it was the advanced group who followed the expected pattern. However, the near native group's responses were tied between 'Option A is more natural' and 'Option B is more natural'. The intermediate group, again, favored the non-*se*-marked form (Option A). The task item of this verb in dynamic context is given in the example below:

(80) Pedro teaches kindergarten. One day he had planned a party for his kids but it was a disaster because some of them got out of hand and made a terrible mess. Pedro chased them around and tried to catch them, when…
A: *Inesperadamente, los niños traviesos subieron a la mesa de un salto.*
B: *Inesperadamente, los niños traviesos se subieron a la mesa de un salto.* (Expected response)

The remaining three verbs (i.e. *aprender* 'to learn', *decidir* 'to decide', *saber* 'to know') were considered to be more natural when they were not *se*-marked, not only in absolute contexts but also in dynamic contexts. The only exception was *saber* 'to know', for which the near native speakers judged that both the non-*se*-marked and *se*-marked forms sounded natural in dynamic contexts. Based on the fact that all three verbs are transitive, it is possible that L1 English

speakers, even at the near native proficiency level, disfavor *se*-marking with transitive verbs. See example of the task item with the verb *decidir* 'to decide' below:

(81) Juan worked as a manager in a local company, but didn't get along with his boss. He often got job offers from other companies but never wanted to move his family. One day, he had a particularly nasty fight with his boss, and…
A: *En ese instante, Juan decidió a cambiar de trabajo.*
B: *En ese instante, Juan se decidió a cambiar de trabajo.* (Expected response)

In sum, L1 English speakers across all proficiency levels generally preferred the non-*se*-marked forms. When context type was taken into consideration (absolute *vs.* dynamic), the responses of the two upper-level learners (i.e. advanced and near native) conformed to the expected pattern, although their responses aligned more so with our expectations in absolute contexts than in dynamic ones. Furthermore, the intermediate-level speakers' performance did not indicate a discernible contrast according to context types, demonstrating a bias toward the non-*se*-marked form. This may be that for English learners of L2 Spanish, especially at the lower proficiency levels, the non-*se*-marked form may be the default for both contexts. Regarding the effects of individual verbs, it was inferred that transitive verbs were more likely to be non-*se*-marked in both contexts.

5.3.2 L1 Korean speakers' Task 2 results

In what follows, I repeat the same series of analyses with the L1 Korean speaker data with the objective of exploring interesting comparisons and contrasts between the two L2 learner groups, when present. First, L1 Korean speakers also preferred the non-*se*-marked form in both absolute and dynamic contexts. When the participants' responses were contrasted according to context types, a similar pattern with L1 English speakers was found: Korean-speaking learners

also performed more as expected in absolute contexts than dynamic ones. However, that L1 Korean near native speakers responded more as expected was different from the corresponding proficiency group in L1 English speakers. In other words, a clearer developmental pattern in the acquisition of the association between *se*-marking and dynamic nature of the event was demonstrated in the L1 Korean group.

A table of distribution of responses is provided below, followed by a mosaic plot to illustrate similarities and/or differences in the results of each proficiency group:

Table 29. Distribution of responses by L1 Korean speakers (Task 2)

	Intermediate (*n*=3)		**Advanced (*n*=13)**		**Near native (*n*=14)**	
	#	%	#	%	#	%
Option A is more natural	25	46.30%	126	53.85%	129	51.19%
Option B is more natural	15	27.78%	70	29.91%	83	32.94%
Both A and B are natural	3	5.55%	31	13.25%	29	11.51%
Neither A nor B is natural	5	9.27%	4	1.71%	6	2.38%
I don't know	3	5.55%	3	1.28%	5	1.98%
No response	3	5.55%	0	0%	0	0%
Total	54	100%	234	100%	252	100%

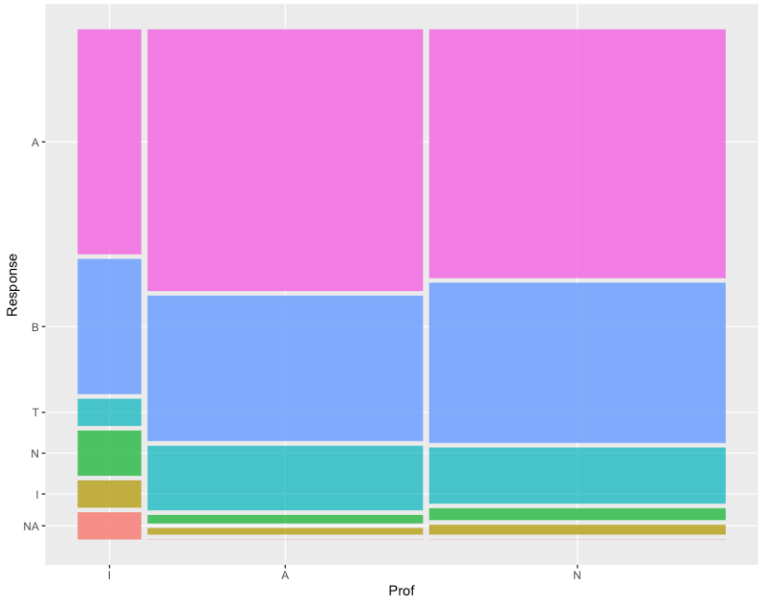


Figure 12. Mosaic plot of responses by L1 Korean speakers (Task 2)

Similar to what was observed in the L1 English speakers' results, there is an overall preference for non-*se*-marked verbal forms (Option A) across all proficiency levels. Except for the intermediate group, whose frequency of Option A responses was slightly below 50%, both advanced and near native speakers judged that the non-*se*-marked form sounded more natural (Option A) in both absolute and dynamic contexts. Again, perhaps these learners, like the L1 English speakers, show a preference for the non-*se*-marked form as the default verb in both contexts.

Additionally, the pattern of choosing the option 'Both A and B are natrual' was also similar to that of L1 English speakers. The two more advanced learner groups (i.e. advanced and near native) selected the 'Both' option more than 10% of the time while the least proficient (i.e. intermediate) speakers' percentage was around 5.55%. Since this option was preferred more in dynamic contexts by L1 English speakers (i.e. 26 times in absolute contexts *vs.* 39 times in

dynamic contexts, cf. Figure 9), it will be of interest to examine the effect of context type for L1 Korean speakers as well. See Table 30 and Figure 13 below.

Table 30. Distribution of responses according to context and proficiency (L1 Korean group)

	Absolute						**Dynamic**					
	Intermediate		Advanced		Near native		Intermediate		Advanced		Near native	
	#	%	#	%	#	%	#	%	#	%	#	%
A	14	52%	89	76%	97	78%	11	41%	37	31%	32	25.5%
B	4	15%	14	12%	15	12%	11	41%	56	48%	68	54%
Both	3	10%	11	9%	8	6%	0	0%	20	17%	21	17%
Neither	4	15%	3	3%	3	2%	1	4%	1	1%	3	2%
I don't know	1	4%	0	0%	3	2%	2	7%	3	3%	2	1.5%
NA	1	4%	0	0%	0	0%	2	7%	0	0%	0	0%
Total	27	100%	117	100%	126	100%	27	100%	117	100%	126	100%

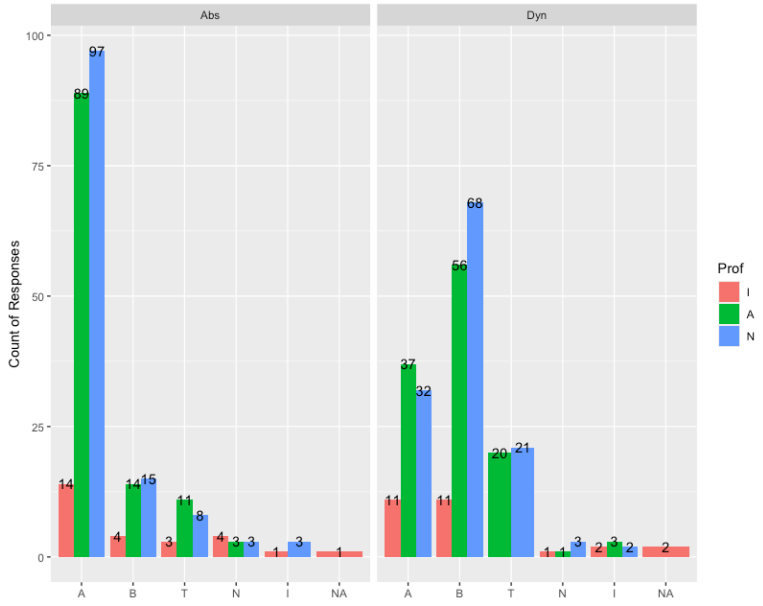


Figure 13. Frequency of responses according to context by L1 Korean speakers (Task 2)

Again, just as with the L1 English speakers, the favoring effect of the non-*se*-marked form (Option A) in absolute contexts was outstanding, especially by advanced (76%) and near native (78%) participants. Although the preference for the *se*-marked variant (Option B) in dynamic contexts was not as discrete, there was still a tendency to favor *se*-marked forms. The intermediate group's count of 'Option A' and 'Option B' responses was tied (41% and 41%, respectively), which diverges from intermediate L1 English speakers' behavior, who preferred Option A over Option B. Another interesting finding here is that near native learners chose the *se*-marked form (Option B) more than half of the time (54%). This conformity with the expected pattern was not observed in the L1 English speakers' results, which showed percentages below 50% across all three proficiency groups (cf. Table 25).

As was assumed based on the results of the L1 English speakers, the token frequency of the response 'Both A and B are natural' was higher in dynamic contexts for the two higher-level proficiency groups. As for the intermediate group, they judged that both options sounded natural only in absolute contexts. Despite this unanticipated behavior of the intermediate speakers, it can still be assumed that L1 Korean speakers found it harder to reject the non-*se*-marked form in dynamic contexts than to reject the *se*-marked form in absolute contexts. In other words, their judgments were more categorical in absolute contexts. It might be the case that L1 Korean speakers are more aware of the fact that the clitic *se* is not likely to be used in contexts where dynamicity is not present. In other words, although they might not have acquired the association between *se*-marking and dynamicity, it is possible that they understand that absoluteness does not trigger the use of *se*.

An outline of how L1 Korean speakers judged the acceptability of (non-)*se*-marking in each context type is demonstrated with examples below:

Table 31. L1 Korean speakers' preference for (non-)*se*-marked option according to contexts

Absolute context: Among all of his friends, Luis is the fastest swimmer. In fact, he used to be a swimmer on his school's champion swimming team. He thinks it's because as a baby, he took swimming lessons.

A: *Luis aprendió a nadar cuando era bebé.* (Expected response)

B: *Luis se aprendió a nadar cuando era bebé.*

	Int		Adv		Near	
	#	%	#	%	#	%
Option A is more natural	**3**	**100%**	**12**	**92.31%**	**13**	**92.86%**
Option B is more natural	0	0%	0	0%	0	0%
Both A and B are natural	0	0%	0	0%	0	0%
Other responses*	0	0%	1	7.69%	1	7.14%

Dynamic context: Daniel has a math exam tomorrow and needs to study hard. He knows that if he can learn the multiplication tables completely, he'll get the A he needs. He studies all night long and by morning,

A: *Daniel había aprendido de memoria las tablas de multiplicación.*

B: *Daniel se había aprendido de memoria las tablas de multiplicación.* (Expected response)

	Int		Adv		Near	
	#	%	#	%	#	%
Option A is more natural	3	100%	7	53.86%	5	35.71%
Option B is more natural	**0**	**0%**	**2**	**15.38%**	**4**	**28.58%**
Both A and B are natural	0	0%	2	15.38%	5	35.71%
Other responses	0	0%	2	15.38%	0	0%

* The responses 'Neither A nor B is natural', 'I don't know', and 'No response' were combined for convenience.

In order to measure the degree of compliance between the expected patterns and the actual responses by L1 Korean speakers, each proficiency group's performance was scored, and its mean, median, and standard deviation were calculated. As before, the closer the score is to the

maximum score (18), the more similar the corresponding performance was to the expected results.

Table 32. Mean and median scores of L1 Korean speakers' performance on Task 2

	Absolute (Max. score = 18)			**Dynamic (Max. score = 18)**		
	Mean	Median	*SD*	Mean	Median	*SD*
Intermediate	10.33	9	3.21	7.33	8	3.06
Advanced	14.54	15	2.37	10.15	10	3.65
Near native	14.62	14	2.99	11	12	3.34

As expected, the overall score in absolute contexts was higher than that of dynamic contexts. While advanced and near native speakers earned over 10 points in both absolute and dynamic contexts, intermediate learners' scores were lower, at 10 and below. It can be predicted that L1 Korean speakers' performance on *se*-marking aligns more with the exptected pattern as their Spanish proficiency level increases. This tendency can be represented graphically as in the following figure:

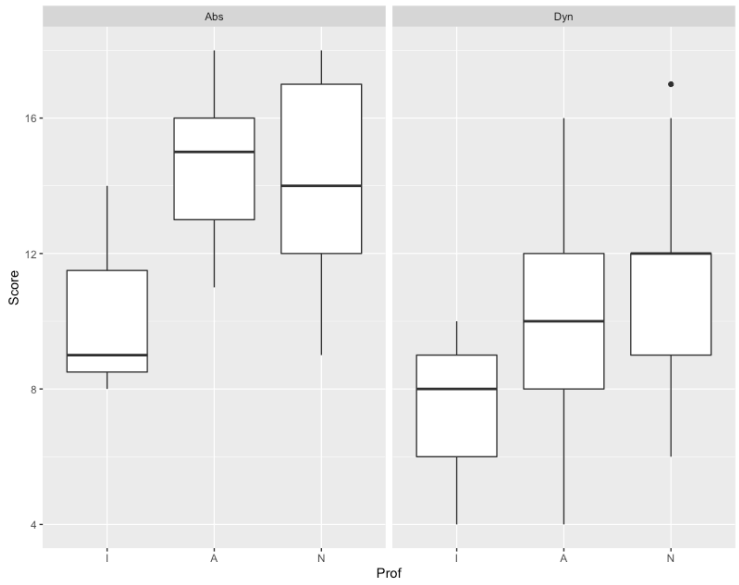


Figure 14. Box plot of scored performance of L1 Korean speakers (Task 2)

The findings so far indicate that (i) the non-*se*-marked form is preferred in general, (ii) it is more likely to see the expected pattern in absolute contexts, and (iii) participants with a higher proficiency level behave more as expected, especially in dynamic contexts. In order to figure out if these findings have any statistical significance, an ANOVA and post-hoc anlaysis with Tukey's test were performed. The same regression model with proficiency, context, and the interaction between these two variables was fit to perform ANOVA on the L1 Korean speakers' data. According to the results, the effect of context turned out to be significant, $F(1, 52) = 22.04$, $p = 1.98e\text{-}05$. Proficiency level also showed significance, to a lesser degree than context, $F(2, 52) = 4.05$, $p = .0233$. However, the interaction was not significant, $F(2, 52) = 0.18$, $p = .8342$.

After the ANOVA, Tukey's Test was performed to determine which proficiency groups were significantly different from each other. The most significant result derived from the difference between the near native and intermediate group ($p < .01$), followed by the contrast between the advanced and intermediate groups ($p < .05$). Meanwhile, the advanced and near

native groups did not indicate any statistical significance, $p = .83$. These results of the statistical analyses can be displayed visually as follows:

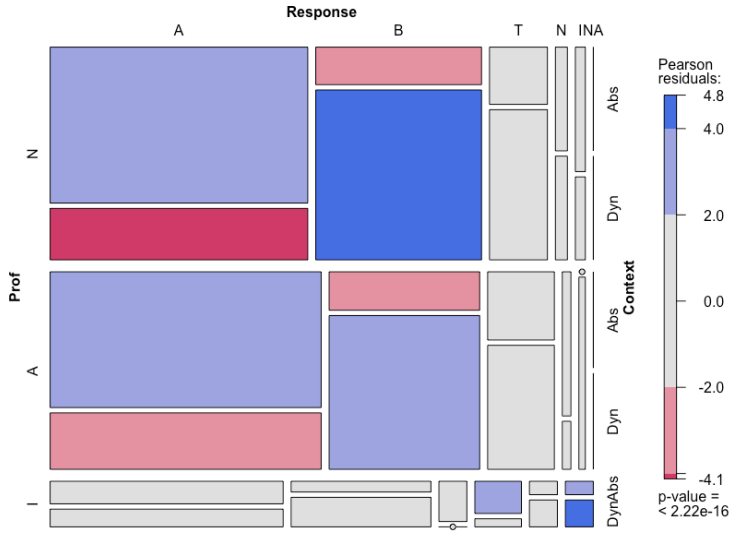


Figure 15. Significance of relevance between factors by L1 Korean speakers (Task 2)

The darkest blue box indicates that near native speakers highly favored *se*-marked verb forms (Option B) in dynamic contexts. Conversely, the darkest red box reflects the opposite pattern: near native speakers disfavored non-*se*-marked forms (Option A) in dynamic contexts. Advanced learners also followed the same behavior, albeit to a lesser degree. Intermediate-level learners' judgments of Option A and Option B were not significant. Instead, that they chose 'No response' in choosing not to respond in dynamic contexts was significant. Also, in absolute contexts, the options 'Both A and B are natural' and 'N/A (No response)' showed some significance, to a lesser degree. This result can be interpreted such that, although the intermediate speakers' judgments were somewhat unsystematic, their accuracy improves as their proficiency of the target language develops.

Once again, like the English speakers, the results for the Korean speakers show a gradual movement toward acquiring the association of *se*-marked verbs with dynamic events and non-*se*-marked verbs with absolute events. This type of knowledge is not explicitly taught in the FL/L2 classroom, but is done implicitly via the explanation of "change of meaning" verbs such as *ir* 'to go' *vs. ir**se*** 'to leave' and *dormir* 'to sleep' *vs. dormir**se*** 'to fall asleep'.

Similar to the analysis of individual verb effects for L1 English speakers, only the responses for a few verbs in the dynamic context will be provided. Despite the low number of intermediate-level participants (n = 3), it is worth mentioning that they all disfavored the non-*se*-marked form of *dormir* 'to sleep' in absolute contexts. One participant chose the *se*-marked form (Option B), while the other two selected the 'Both' option. Also, 9 out of 27 advanced and near native participants chose Option B and 'Both'. On the other hand, in dynamic contexts, there were four verbs that displayed unexpected behavior, which will be discussed here and is illustrated in Table 33 below.

Table 33. L1 Korean speakers' distribution of responses by verb in dynamic contexts (Task 2)

aprender	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is natural	3	100%	7	55%	5	36%	15	50%
Option B is natural	**0**	**0%**	**2**	**15%**	**4**	**28%**	**6**	**20%**
Both A and B are natural	0	0%	2	15%	5	36%	7	24%
Other responses*	0	0%	2	15%	0	0%	2	6%
bajar	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is natural	0	0%	6	46%	4	28%	10	33%
Option B is natural	**1**	**33%**	**4**	**31%**	**7**	**50%**	**12**	**40%**
Both A and B are natural	0	0%	2	15%	3	22%	5	17%

Other responses	2	67%	1	8%	0	0%	3	10%
saber	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is natural	2	67%	7	55%	7	50%	16	52%
Option B is natural	**0**	**0%**	**2**	**15%**	**3**	**22%**	**5**	**17%**
Both A and B are natural	0	0%	3	22%	2	14%	5	17%
Other responses	1	33%	1	8%	2	14%	4	14%
subir	Intermediate		Advanced		Near native		Total	
	#	%	#	%	#	%	#	%
Option A is natural	2	67%	7	55%	6	42%	15	50%
Option B is natural	**1**	**33%**	**3**	**22%**	**5**	**36%**	**9**	**30%**
Both A and B are natural	0	0%	3	22%	3	22%	6	20%
Other responses	0	0%	0	0%	0	0%	0	0%

* The responses 'Neither A nor B is natural', 'I don't know', and 'No response' were combined for convenience.

For all four verbs in Table 33, Option B was selected less than 50% of the time (with the exception of near native speakers' 50% rate of choosing Option B with the verb *bajar* 'to go down') in the participants' responses, even though it was the expected answer for these dynamic contexts. In descending order of the rate of Option B selection, the verbs are: *bajar* 'to go down' (40%) > *subir* 'to go up' (30%) > *aprender* 'to learn' (20%) > *saber* 'to know' (17%). Specifically, it was the advanced learner group which favored the non-*se*-marked form of the verb *bajar* 'to go down' in dynamic contexts (46%). The other two groups, intermediate (33%) and near native (50%), preferred the *se*-marked form over the non-*se*-marked form as expected. Next, *subir* 'to go up' was preferred in its non-*se*-marked form by all three groups, demonstrated by the way in which 50% of the responses selected Option A. A similar pattern was found with the verb *aprender* 'to learn'. The non-*se*-marked form was favored 50% of the time, followed by

the ‘Both’ option (24%). The most unexpected pattern was found with the verb *saber* ‘to know’. Option A was even more preferred (52%), and Option B and ‘Both’ tied at 17%.

The four verbs in question can be divided into two categories based on their semantic characteristics (cf. Maldoando 2008). Firstly, *bajar* ‘to go down’ and *subir* ‘to go up’ can be grouped as verbs of change in location or translational motion. On the other hand, *aprender* ‘to learn’ and *saber* ‘to know’ form part of the verbs of internal change (mental) or cognition middle. Hence, it could be inferred that L1 Korean speakers are more likely to be affected by the semantic characteristics of the verbs than by the contexts upon interpreting the use of *se*.

To sum up, L1 Korean speakers also preferred the non-*se*-marked forms in general, regardless of the context, perhaps like the English speakers, regarding this form as the default verb for both absolute and dynamic contexts. Therefore, their performance aligned more with the expected pattern in absolute contexts than in dynamic contexts. The difference between the three proficiency levels was clearer than that of the English speakers in this study. The probability of the near native group conforming to the expected pattern was high, followed by the advanced group to a lesser degree. On the other hand, the intermediate group’s performance did not demonstrate a noticeable pattern, which implies that their judgment on Spanish *se*-marking beyond the reflexive and detransitivizing function, like the English speakers, has yet to be developed. In terms of the individual verbs’ effects, it appears likely that the semantic properties of verbs affect the interpretation of *se*-marking for L1 Korean speakers.

5.3.3 L1 Spanish speakers’ Task 2 results

Finally, the results of the L1 Spanish speakers as a control group will be described in detail. In general, native speakers also showed an overall preference for non-*se*-marked forms in

all contexts, albeit to a lesser degree than that of non-native speakers. Regarding native speakers' performance according to contexts (i.e. absolute *vs.* dynamic), it was more congruent with the expectations of this research: L1 Spanish speakers favored the non-*se*-marked forms in absolute contexts, and they favored the *se*-marked forms in dynamic contexts. Nonetheless, although native speakers performed as expected in the absolute contexts (preferring the non-*se*-marked sentences), in dynamic contexts, they judged both *se*-marked and non-*se*-marked forms natural, which deviated from our expectations. This reveals that native speakers appear to choose whether or not they focus on or put into profile (Langacker 1991) the pivotal moment of change in these dynamic events. Absolute contexts do not present this option, and thus *se* is not preferred.

Given that dividing the group of speakers into subgroups according to proficiency level is not applicable with native speakers, I will compare the performance of the native speaker group to that of the English and Korean speaker groups. First, I provide a table of distribution of responses with a mosaic plot to graphically represent the contrast.

Table 34. Distribution of responses by L1 Spanish speakers (Task 2)

(*n*=28)	#	%
Option A is more natural	223	44.25%
Option B is more natural	174	34.52%
Both A and B are natural	80	15.87%
Neither A nor B is natural	27	5.36%
I don't know	0	0%
No response	0	0%
Total	504	100%

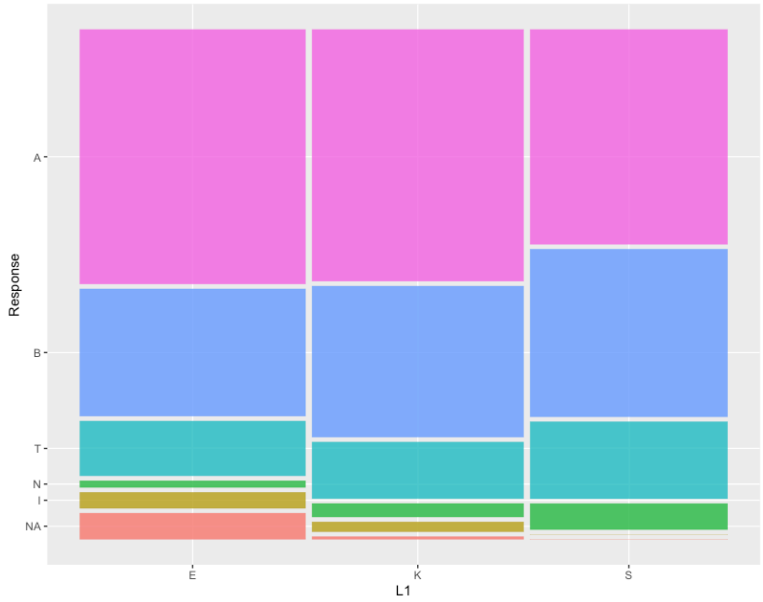


Figure 16. Mosaic plot of responses according to L1 (Task 2)

It is shown above that even native speakers of Spanish preferred the non-*se*-marked verbal forms, with the rate of 44.25%. When compared to the two non-native speaker groups which selected Option A over 50% of the time, native speakers' responses were more evenly distributed. Meanwhile, the last two options (i.e. 'I don't know' and 'No response') were not chosen by the control group.

Below I provide a detailed analysis of the count of responses by native speakers according to context type:

Table 35. Distribution of responses according to context (L1 Spanish group)

	Absolute		**Dynamic**	
	#	%	#	%
Option A is natural	207	82.14%	16	6.35%
Option B is natural	12	4.76%	162	64.28%

Both A and B are natural	23	9.13%	57	22.62%
Neither A nor B is natural	10	3.97%	17	6.75%
I don't know	0	0%	0	0%
No response	0	0%	0	0%
Total	252	100%	252	100%

As shown in Table 35, the favoring effect of Option A in absolute contexts was overwhelmingly high at greater than 82%. However, the preference for Option B in dynamic contexts is not as remarkable, although the percentage remains over 50% at slightly above 64%. In both contexts, the second most preferred response was the 'Both' option. Native speakers were especially more likely to accept both non-*se*-marked and *se*-marked forms in dynamic contexts than they were in absolute contexts.

In order to provide a general sketch of the results, a bar graph which summarizes the count of the responses in each context for all participant groups is presented below:

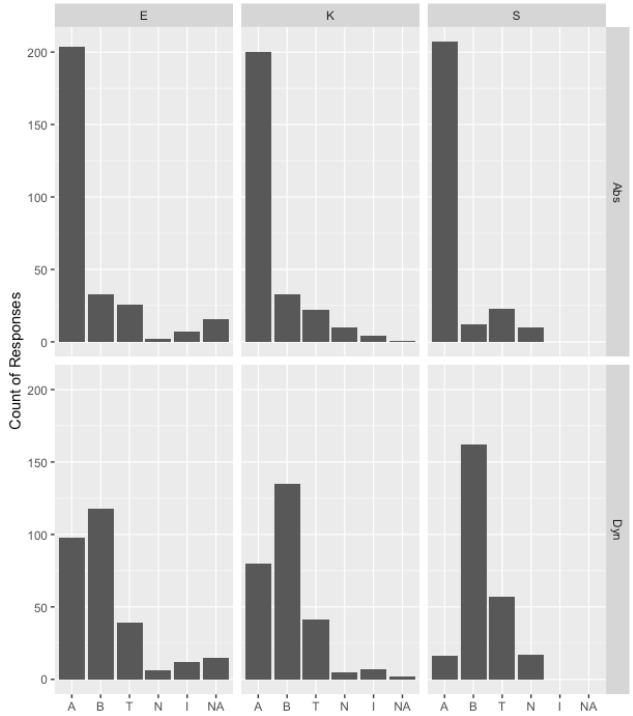


Figure 17. Count of responses by all participant groups (Task 2)

As seen in Figure 17, all three speaker groups showed a similar pattern in absolute contexts: they favored non-*se*-marked sentences, which conforms to the expectation. On the other hand, there is a stronger discrepancy between the non-native L2 learner groups and the native speaker group in dynamic contexts, especially in the frequency of Option A responses. That is, while native speakers clearly disfavored the non-*se*-marked form in dynamic contexts, the disfavoring effect was not as strong for non-native speakers. This seems to demonstrate that native speakers' association between *se*-marking and dynamicity of the event is stronger than that of the less-proficient L2 learners, who are more likely to perceive the non-*se*-marked form as default forms regardless of the contexts. See below example that highlights these findings:

Table 36. Participants' preference for (non-)*se*-marked option according to contexts

Absolute context: Among all of his friends, Luis is the fastest swimmer. In fact, he used to be a swimmer on his school's champion swimming team. He thinks it's because as a baby, he took swimming lessons.

A: *Luis aprendió a nadar cuando era bebé.* (Expected response)

B: *Luis se aprendió a nadar cuando era bebé.*

	ENG		KOR		SPAN	
	#	%	#	%	#	%
Option A is more natural	**30**	**93.75%**	**28**	**93.33%**	**28**	**100%**
Option B is more natural	1	3.125%	0	0%	0	0%
Both A and B are natural	0	0%	0	0%	0	0%
Other responses*	1	3.125%	2	6.67%	0	0%

Dynamic context: Daniel has a math exam tomorrow and needs to study hard. He knows that if he can learn the multiplication tables completely, he'll get the A he needs. He studies all night long and by morning,

A: *Daniel había aprendido de memoria las tablas de multiplicación.*

B: *Daniel se había aprendido de memoria las tablas de multiplicación.* (Expected response)

	ENG		KOR		SPAN	
	#	%	#	%	#	%
Option A is more natural	17	53.12%	15	50%	1	3.57%
Option B is more natural	**6**	**18.75%**	**6**	**20%**	**18**	**64.29%**
Both A and B are natural	5	15.63%	7	23.33%	6	21.43%
Other responses	4	12.5%	2	6.67%	3	10.71%

* The responses 'Neither A nor B is natural', 'I don't know', and 'No response' were combined for convenience.

Overall, native speakers' responses appear to be congruent with the expected patterns. To determine to what extent the similarity lies, I scored and calculated the control group's results in the same way I did with those of the two non-native speaker groups. It will be recalled that the

closer the mean and median are to the maximum score (18), the more in agreement the performance is to the expected pattern.

Table 37. Mean and median scores of L1 Spanish speakers' performance on Task 2

Absolute (Max. score = 18)			**Dynamic (Max. score = 18)**		
Mean	Median	*SD*	Mean	Median	*SD*
15.61	16	2.28	13.61	14.5	3.07

The difference between the two contexts is not as sizeable as the results of non-native speakers. In order to compare the three groups' overall scores, Figure 18 is provided below as a graphic representation of the median and range for the selected variables:

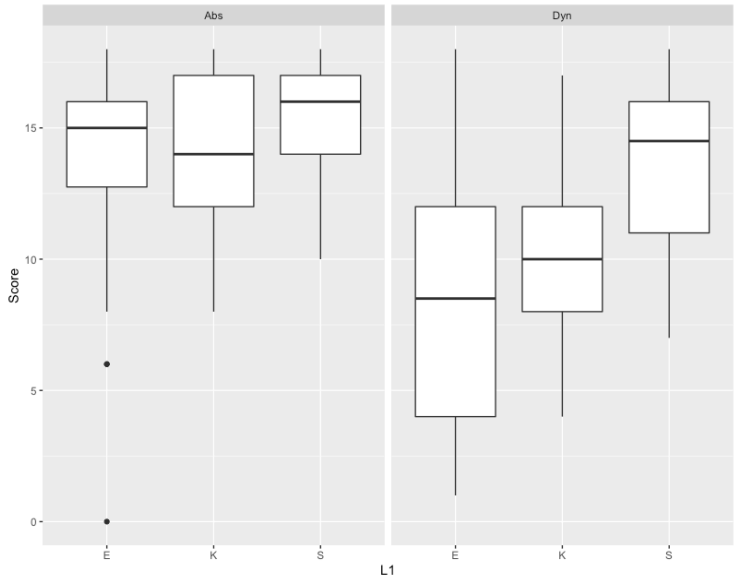


Figure 18. Box plot of scored performance of all participants (Task 2)

Notwithstanding the fact that the median score is highest for the native speaker group, non-native speakers performed fairly native-like in absolute contexts. However, in dynamic contexts, the

two boxes of non-native speakers lie comparatively below the *y*-axis scale, which means that their performance diverged more greatly from the expected pattern.

In order to scrutinize the statistical significance of the differences between the three L1 groups shown so far, ANOVA was performed with a regression model that included L1, context, and the interaction between the two independent variables. The results revealed that both context ($F(1, 172) = 50.29$, $p = 3.30e\text{-}11$) and L1 ($F(2, 172 = 15.572$, $p = 6.09e\text{-}07$) have a statistically significant effect. The interaction was neutrally significant with a *p*-value slightly above .05, $F(2, 172) = 2.861$, $p = .0599$.

The Tukey's Test performed on the same model demonstrated which correlations were significant. When L1 was considered, the difference between the English- and Spanish-speaker groups was the most significant ($p = .0000004$), followed by that of the Korean- and Spanish-spaker groups ($p = .00089$). The two non-native speaker groups were not significantly different from each other, $p > .05$. Also, it was only in the dynamic context where the interaction between context type and L1 was significant. When the context was dynamic, L1 English and L1 Spanish speakers' performance differed significantly, $p < .001$. The difference between L1 Korean and L1 Spanish speakers was also significant in dynamic contexts but to a lesser degree, $p < .01$. However, the interaction between the two non-native speaker groups in dynamic contexts was not significant, $p > .05$. The mosaic plot below (Figure 19) illustrates the results:

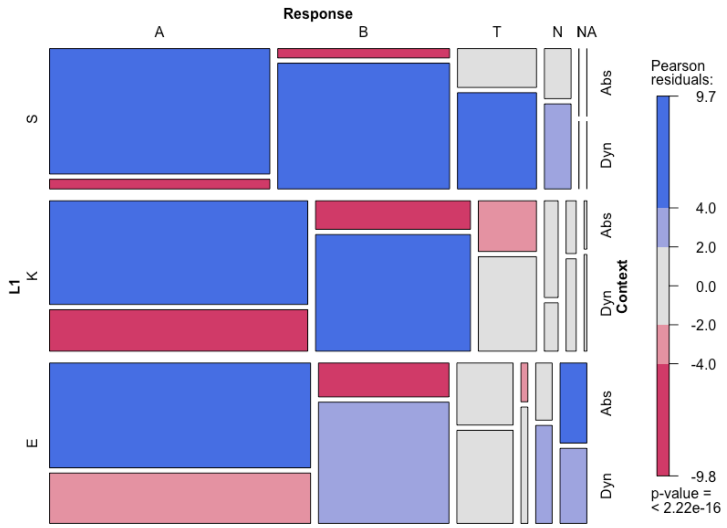


Figure 19. Significance of relevance between response, L1, and context (Task 2)

Based on the darker blue and darker red boxes in the plot, all three groups performed as expected in absolute contexts. However, the level of conformity decreases in L1 English speakers' performance in dynamic contexts, as represented by the brighter blue and brighter red boxes. Also, the 'Both' option was only significantly preferred by the L1 Spanish speaker group in dynamic contexts, while it was disfavored by L1 Korean speakers in aboslute contexts. The 'Neither' option was neutrally favored by L1 Spanish and L1 English speakers in dynamic contexts. Finally, it was only the L1 English speaker group for whom 'No response' was found to be significant in both contexts.

Although the native speakers' performance was generally in line with the expectation, there were three specific verbs in which some interesting findings were observed. They are compared to the count of responses by non-native speakers in Table 38 below:

Table 38. Distribution of responses of verbs that showed unexpected results (Task 2)

	Absolute context						**Dynamic context**											
	saber						*saber*						*bajar*					
	E		K		S		E		K		S		E		K		S	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
A	23	72%	25	83%	9	32%	19	59%	16	52%	2	7%	4	12.5%	10	33%	0	0%
B	6	19%	1	3%	7	25%	4	13%	5	17%	14	50%	20	62.5%	12	40%	12	43%
Both	2	6%	2	7%	9	32%	6	19%	5	17%	12	43%	1	3%	5	17%	12	43%
Neither	0	0%	2	7%	3	11%	0	0%	2	7%	0	0%	1	3%	0	0%	4	14%
I don't know	0	0%	0	0%	0	0%	1	3%	2	7%	0	0%	4	13%	2	7%	0	0%
NA	1	3%	0	0%	0	0%	2	6%	0	0%	0	0%	2	6%	1	3%	0	0%
Total	32	100%	30	100%	28	100%	32	100%	30	100%	28	100%	32	100%	30	100%	28	100%

As shown in Table 38, the native speakers' responses for the verb *saber* 'to know' was unexpected in both absolute and dynamic contexts. While non-native speakers strongly favored the non-*se*-marked form in absolute contexts, the rate of native speakers' responses 'Option A is more natural' and 'Both A and B are natural' were tied at 32%. Option B (*se*-marked) also showed a high frequency of selection (25%). In dynamic contexts, the percentage of selecting Option B was the highest as expected (50%), however, the option 'Both' also showed a high percentage of selection at 43%. Based on these results, it can be inferred that although Spanish speakers did follow the expected pattern with the verb *saber* 'to know', there exists a definite preference for admitting both non-*se*-marked and *se*-marked forms in both contexts. The other verb that showed a similar pattern was *bajar* 'to go down'. With this verb, native speakers' response counts for Option B and the 'Both' option were tied. Again, it is shown that native speakers are more likely to accept both non-*se*-marked and *se*-marked forms in dynamic contexts. Another interpretation is that the native speakers were not bound by the somewhat forced context imposed on them by the task designed to elicit one form or another. But rather, by selecting a preference for one form or another, or by accepting both forms, they gave an absolute or dynamic interpretation to the contexts presented.

In sum, native speakers' responses on Task 2 were more compatible with the assumptions proposed by previous literature in cognitive linguistics, in comparison with the two non-native speaker groups of the present study. However, there was a slight difference depending on the context of the situation. In absolute contexts, there was a stronger preference for the non-*se*-marked form of the verb, as expected. In dynamic contexts, the *se*-marked form was preferred, but the favoring effect was less robust. This is because native speakers judged that both the non-*se*-marked and *se*-marked forms were acceptable in dynamic contexts. By accepting a non-*se*-

marked form, it appears that Spanish speakers are capable of interpreting a dynamic event as an absolute one. By accepting both options in many cases, they signal that they as speakers are not constrained by a particular context but rather by the choice of the use or non-use of the morpheme *se*, they determine whether an event is dynamic or absolute.

In this task, certain items were designed to an abrupt, rapid or unexpected context, usually with an adverb (e.g. *abruptamente* 'abruptly'), or to imply an absolute context with a different kind of adverb (e.g. *lentamente* 'slowly'). What follows are examples with these adverbs and their corresponding distribution of responses by all participant groups:

Table 39. Participants' preference for (non-)*se*-marked option according to contexts with adverbs

Absolute context: After school, Laura went upstairs to her room to read. When it was time for dinner, her mom called her to come help set the table. Taking her book and still reading,
A: *Laura bajó por las escaleras y entró al comedor lentamente.* (Expected response)
B: *Laura se bajó por las escaleras y entró al comedor lentamente.*

	ENG		KOR		SPAN	
	#	%	#	%	#	%
Option A is more natural	**18**	**56.25%**	**23**	**76.67%**	**28**	**100%**
Option B is more natural	7	21.87%	4	13.33%	0	0%
Both A and B are natural	4	12.5%	2	6.67%	0	0%
Other responses*	3	9.38%	1	3.33%	0	0%

Dynamic context: Leo and his little brother were playing on the seesaw. When they got tired, his brother slid carefully off the seesaw. But not Leo who does everything in a hurry.
A: *Abruptamente, Leo bajó del sube y baja de un brinco.*
B: *Abruptamente, Leo se bajó del sube y baja de un brinco.* (Expected response)

	ENG		KOR		SPAN	
	#	%	#	%	#	%
Option A is more natural	4	12.5%	10	33.33%	0	0%

Option B is more natural	**20**	**62.5%**	**12**	**40%**	**12**	**42.86%**
Both A and B are natural	1	3.13%%	5	16.67%	12	42.86%
Other responses	7	21.87%	3	10%	4	14.28%

* The responses 'Neither A nor B is natural', 'I don't know', and 'No response' were combined for convenience.

What is shown in the above table is that when the absoluteness of the context is reinforced by the use of an adverb (i.e. *lentamente* 'slowly'), native Spanish speakers unanimously preferred the non-*se*-marked form. However, in the opposite case where the middleness of the context is highlighted by the use of an adverb (i.e. *abruptamente* 'abruptly'), L1 Spanish speakers' responses were not biased for the *se*-marked form: they rather showed a pattern of allowing both *se*-marked and non-*se*-marked forms.

To summarize, intermediate-level speakers in both L1 English and L1 Korean groups did not show a pattern of differentiating the two context types (i.e. absolute *vs.* dynamic) with their (non-)use of *se*. They overwhelmingly preferred the non-*se*-marked forms in general, suggesting that they might perceive and utilize the non-*se*-marked form as the default for both contexts. However, as proficiency level increases, a stronger association between the use of *se* and the dynamicity of event was revealed. It is an interesting result considering that L2 learners are capable of acquiring those conceptualizations that are not explicitly taught in Spanish classrooms. As for the native speakers, their performance was the most congruent with the expectations. They strongly favored the non-*se*-marked forms in absolute contexts and preferred the *se*-marked forms in dynamic ones. However, the latter pattern was less apparent because Spanish speakers judged that both *se*-marked and non-*se*-marked forms are acceptable in such contexts. From this result, it is possible to assume that L1 Spanish speakers did not necessarily interpret the contexts in a way that was intended by the task. It appears that they don't always

depend on the adverbs to interpret the dynamicity of the event but rather use (or don't use) *se* to give an event different interpretations. Moreover, all groups were better at preferring non-*se*-marked verbs with absolute contexts, showing that these non-abrupt, expected events are not compatible with *se*. However, abrupt and/or unexpected events are compatible with both *se*-marked and non-*se*-marked verbs and thus are open for interpretation.

5.4 Task 3 (Cloze task) Results — testing learners' understanding of reflexives *vs.* middles

Tasks 1 and 2 revealed that the L2 learners in this study have gradually acquired a native-like performance of *se*-marking in both production and interpretation data. Specifically, when interpreting certain contexts, both Spanish speakers and L2 learners showed an understanding of the dynamicity conveyed by the use of *se*. Given that they are aware of the difference between *se*-marked dynamic contexts and non-*se*-marked absolute contexts, it is of interest to examine if they acknowledge the clitic *se* employed in such constructions are not reflexive markers but in fact middle markers. Accordingly, for this task, all participants made judgments on the possibility of using *a sí mismo* when a *se*-marked form appeared in a short text. It will be recalled that there were four types of contexts in this task: 'Routine-Reflexive', 'Routine-Middle', 'Non-Routine-Reflexive', and 'Non-Routine-Middle'. The main purpose of this task was to explore the extent to which the participants associated the construction's reflexivity, or lack thereof, with the use of the clitic *se*. In other words, if a participant were to respond that *a sí mismo* could be used in a context where its presence was not expected, this would indicate that the participant strongly associated the clitic *se* with reflexivity. Conversely, if a participant's response were to correspond with the expectation, it is interpreted that the participant

acknowledged that not all uses of *se* are to mark reflexivity or that *se* functions as a middle marker in certain contexts. The following task items demonstrate the expected patterns:

(82) *Cuando Pilar se puso el nuevo vestido, salió de su cuarto y le preguntó a su hermana, "¿Cómo* ***me veo (*a mí misma)*** *en este vestido?" "Muy guapa", le contestó su hermana. Pero Pilar no estaba segura y dijo, "No sé, ahora que* ***me veo (a mí misma)*** *en el espejo, no me gusta."*

'When Pilar put on the new dress, she came out of her room and asked her sister, "How do I **look (*myself)** in this dress?" "Very pretty," her sister replied. But Pilar wasn't sure and she said, "I don't know, now that I **see/look at (myself)** in the mirror, I don't like it."'

5.4.1 L1 English speakers' Task 3 results

L1 English speakers' task results demonstrated that they are capable of differentiating reflexive *vs.* middle *se* (i.e. accepting *a sí mismo* in reflexive contexts and rejecting it in middle ones), more so with daily routine verbs than non-routine ones. Also, it was observed that this behavior becomes clearer as the proficiency level of L2 learners increases.

Table 40 below reveals the distribution of responses by L1 English speakers according to proficiency level and context type. Expected responses are highlighted in bold. The table will be followed by a bar graph in order to ease the understanding of the data with a visual representation.

Table 40. Distribution of responses by L1 English speakers (Task 3)

		Intermediate (*n*=4)		**Advanced (*n*=10)**		**Near native (*n*=14)**	
		#	%	#	%	#	%
Routine - Reflexive	**Yes**	**16**	**66.66%**	**45**	**75%**	**65**	**77.38%**
	No	7	29.17%	10	16.67%	17	20.24%

	I don't know	1	4.17%	5	8.33%	0	0%
	No response	0	0%	0	0%	2	2.38%
Routine - Middle	Yes	4	16.67%	12	20%	16	19.05%
	No	**18**	**75%**	**42**	**70%**	**67**	**79.76%**
	I don't know	2	8.33%	5	8.33%	1	1.19%
	No response	0	0%	1	1.67%	0	0%
Non-Routine - Reflexive	**Yes**	**11**	**55%**	**35**	**70%**	**50**	**71.43%**
	No	9	45%	10	20%	19	27.14%
	I don't know	0	0%	5	10%	1	1.43%
	No response	0	0%	0	0%	0	0%
Non-Routine - Middle	Yes	4	20%	8	16%	12	17.14%
	No	**15**	**75%**	**36**	**72%**	**58**	**82.86%**
	I don't know	1	5%	5	10%	0	0%
	No response	0	0%	1	2%	0	0%

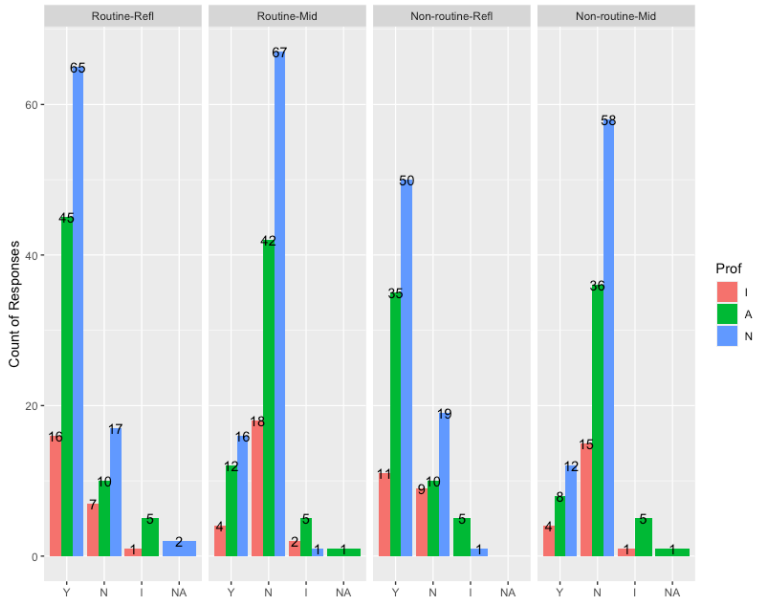


Figure 20. Count of responses by L1 English speakers (Task 3)

Overall, the L1 English speakers behaved according to the expected patterns across different context types. This result indicates that their acceptance of the phrase, *a sí mismo* 'oneself' is higher with daily routine verbs. Also, as the proficiency level increases, the degree of acceptance of this phrase with these types of verbs increases.

The descriptive results presented above were tested for statistical significance with an anlaysis of variance (ANOVA). The regression model fit for the analysis included proficiency, type of context, and the interaction of the two to examine their relevance for the participants' responses. Also, the count for 'No response' was excluded in the analysis for its low frequency. The effect of profieicncy was non-significant, although its *p*-value was lower than that of the interaction effect, $F(2, 604) = 2.753$, $p = .0646$. The only factor that showed a strong significance was context type, $F(3, 604) = 15.37$, $p < 2e\text{-}16$. A mosaic plot is provided below to demonstrate the relevance between the factors:

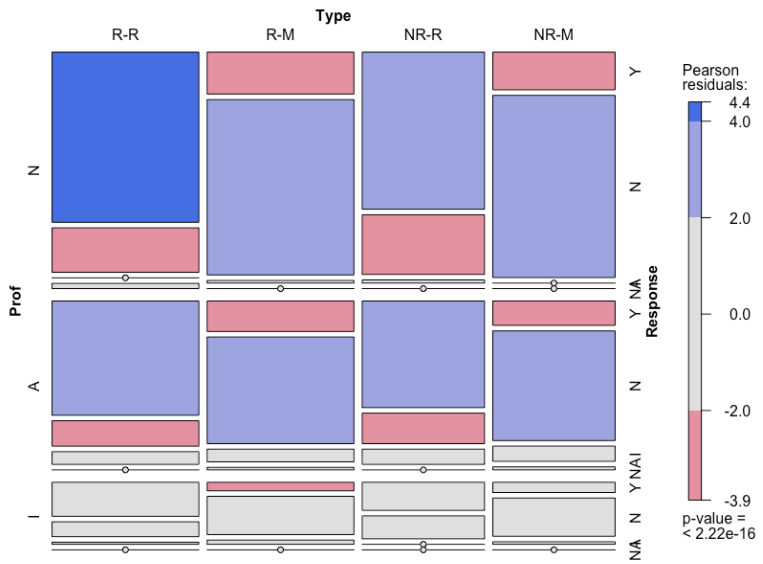


Figure 21. Significance of relevance between factors by L1 English speakers (Task 3)

The only significant relevance shown for the intermediate-level group is their disfavoring of using *a sí mismo* when the context pertained to non-reflexive daily routines, as illustrated in the following example:

Table 41. L1 English speakers' acceptance of *a sí mismo* in daily routine middle context

Middle context: *Cuando suena la alarma,* ***me despierto (*a mí mismo)*** *automáticamente.*						
	Intermediate		Advanced		Near native	
	#	%	#	%	#	%
Yes	0	0%	0	0%	2	14.29%
No	**4**	**100%**	**8**	**80%**	**12**	**85.71%**
I don't know	0	0%	2	20%	0	0%
No response	0	0%	0	0%	0	0%

As shown in Figure 21, the advanced group's responses were moderately in accordance with the expectation in all four context types. Near native speakers also revealed conformity to the expected patterns with a stronger degree of preference for using *a sí mismo* in the context of reflexive daily routines.

One daily routine verb (*despertarse* 'to wake up') and one non-routine verb (*perderse* 'to get lost') diverged from the expected responses. Particularly, in the context in which using *a sí mismo* is considered natural for *despertarse* 'to wake up' and *perderse* 'to get lost' (routine and non-routine contexts, respectively), the L1 English speakers did not favor the inclusion of *a sí mismo* as highly as expected. In opposite contexts, that is, in which the addition of *a sí mismo* is not natural, the participants highly preferred 'No' as expected. These divergent findings of *despertarse* 'to wake up' and *perderse* 'to get lost' are displayed in Table 42 below:

Table 42. L1 English speakers' distribution of responses by verbs (Task 3)

			#	%
despertarse **' to wake up'**	Routine - Reflexive	Yes	16	59.26%
		No	10	37.04%
		I don't know	1	3.7%
	Routine - Middle	Yes	2	7.14%
		No	24	85.72%
		I don't know	2	7.14%
perderse **'to get lost'**	Non-Routine - Reflexive	Yes	10	35.71%
		No	18	64.29%
		I don't know	0	0%
	Non-Routine - Middle	Yes	7	25%
		No	19	67.86%
		I don't know	2	7.14%

The participants' responses were more similar to the expected pattern with the daily routine verb *despertarse* 'to wake up'. Although they selected 'Yes' most of the time (59.26%), approximately 37% of the participants answered that using *a sí mismo* is not acceptable. This divergence from the expectation becomes even greater with the non-routine verb, *perderse* 'to get lost'. As this verb expresses reflexivity, its interpretation can be translated as 'to lose oneself'. The fact that adding *a sí mismo* was refused by the participants indicates that the L1 English speakers might not have perceived the verb's reflexivity, especially considering that it is not a daily routine verb.

In sum, the overall performance of L1 English speakers on Task 3 complies with the expected patterns. Although the effect of proficiency level did not prove to be significant, there was a greater resemblance to the expectation by the most proficient group. There was not a clear pattern in the intermediate-level group's performance, which suggests that their judgments on the

association between the clitic *se* and the phrase *a sí mismo* is considerably weak. Also, the effect of each individual verb was minor.

5.4.2 L1 Korean speakers' Task 3 results

In what follows, I analyze the results of the L1 Korean speakers' responses to the cloze task (i.e. Task 3) using the same descriptive and statistical procedures that were used to examine the L1 English speakers' responses. Overall, Korean-speaking L2 learners performed as expected with deffering degrees according to context types. In general, their favoring effect of accepting *a sí mismo* in reflexive contexts was greater than their rate of rejecting *a sí mismo* in non-reflexive (i.e. middle) contexts. Regarding the difference between each proficiency group, intermediate learners' judgment on use or non-use of *a sí mismo* did not show any concrete pattern (i.e. they neither accepted nor rejected the inclusion of *a sí mismo* in any contexts); however, advanced and near native learners revealed a greater degree of association between reflexivity and *a sí mismo* by including it in reflexive contexts and rejecting it in middle ones.

First, a table of distribution of responses and a bar plot are provided below to illustrate the overall frequencies of each proficiency group's responses.

Table 43. Distribution of responses by L1 Korean speakers (Task 3)

		Intermediate (*n*=3)		**Advanced (*n*=13)**		**Near native (*n*=13)**	
		#	%	#	%	#	%
Routine - Reflexive	**Yes**	**10**	**55.55%**	**63**	**80.77%%**	**67**	**85.9%**
	No	7	38.89%	15	19.23%	11	14.1%
	I don't know	1	5.56%	0	0%	0	0%
	No response	0	0%	0	0%	0	0%
	Yes	3	16.67%	14	17.95%	11	14.1%

Routine - Middle	**No**	**13**	**72.22%**	**61**	**78.2%**	**62**	**79.49%**
	I don't know	2	11.11%	3	3.85%	4	5.13%
	No response	0	0%	0	0%	1	1.28%
Non-Routine - Reflexive	**Yes**	**9**	**60%**	**47**	**72.31%**	**57**	**87.69%**
	No	3	20%	15	23.07%	8	12.31%
	I don't know	3	20%	2	3.08%	0	0%
	No response	0	0%	1	1.54%	0	0%
Non-Routine - Middle	Yes	4	26.67%	5	7.7%	9	13.85%
	No	**8**	**53.33%**	**56**	**86.15%**	**52**	**80%**
	I don't know	3	20%	4	6.15%	1	1.54%
	No response	0	0%	0	0%	3	4.61%

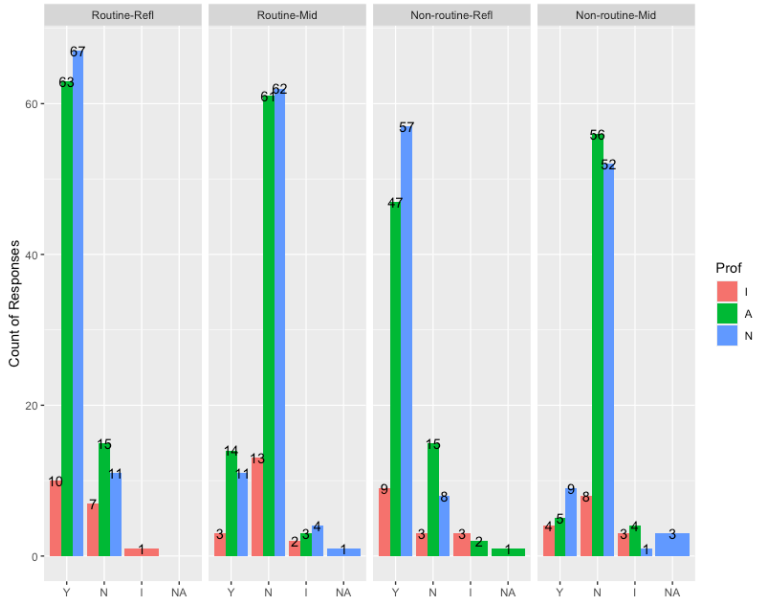


Figure 22. Count of responses by L1 Korean speakers (Task 3)

The L1 Korean speakers' responses also display similarities to the expectation, as was the case for L1 English speakers. Overall, there is a correlation between proficiency level and performance. The higher the speaker's proficiency level is, the more congruent their results are

to the expected pattern. However, there is one exception to this tendency. In the contexts where the phrase *a sí mismo* was not expected to be used for *se*-marked non-routine verbs, the rate of rejection of use of the phrase by advanced speakers (86.15%) slightly exceeded that of the near native level participants (80%).

The results of an ANOVA on a regression model with the factors proficiency, context type, and the interaction of the two demonstrated that the effects of all three factors were significant. Specifically, context type had the strongest significance, $F(3, 626) = 124.35$, $p < 2e\text{-}16$, followed by proficiency level, $F(2, 626) = 6.06$, $p < .01$. The interaction revealed a minor significance, $F(6, 626) = 2.37$, $p < .05$. A post-hoc analysis with Tukey's test showed that the difference between the advanced and near native groups lacked significance, $p > .05$. However, the intermediate and advanced groups differed significantly, $p < .05$. Moreover, there was a stronger significance in the difference between the intermediate and near native groups, $p < .01$. The effect of context type and its interaction with proficiency level is illustrated below in Figure 23.

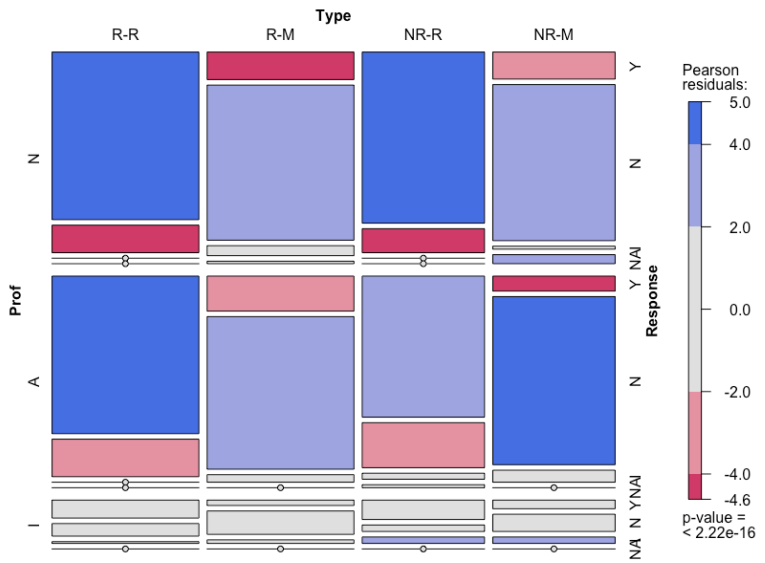


Figure 23. Significance of relevance between factors by L1 Korean speakers (Task 3)

In the above mosaic plot, it is shown that the intermediate group's judgments on the use of *a sí mismo* is not categorical. The only response that showed a slight significance was 'I don't know' with non-daily routine verbs. The more proficient groups' (i.e. advanced and near native) results followed the expected pattern with a differing degree of significance. In the context where daily routine verbs were used as reflexive verbs (i.e. compatible with *a sí mismo*), both groups strongly favored the use of *a sí mismo*. However, the significance of the disfavoring effect differed: the near native group more significantly disfavored the response 'No'. When daily routine verbs lacked reflexivity, only the near native group strongly disfavored the use of *a sí mismo*. However, the overall pattern of responses was congruent with our expectations. With non-daily routine verbs, the near native speakers performed more similar to the expected pattern when the context implied reflexivity, while the advanced group conformed more to the expectation in the opposite context (i.e. the darker blue and darker red boxes in the quartile 'NR-R (Non-Routine-Reflexive) for the proficiency level 'N' (Near native) *vs.* the darker blue and darker red boxes in the quartile 'NR-M' (Non-Routine-Middle) for the proficiency level 'A' (Advanced)). An example of this result is illustrated below:

Table 44. L1 Korean speakers' acceptance of *a sí mismo* with non-routine verbs

Reflexive context: *Cuando veo una película,* ***me imagino (a mí mismo)*** *como un actor principal.*						
	Intermediate		Advanced		Near native	
	#	%	#	%	#	%
Yes	**2**	**66.67%**	**9**	**69.23%**	**13**	**100%**
No	1	33.33%	4	30.77%	0	0%
I don't know	0	0%	0	0%	0	0%
No response	0	0%	0	0%	0	0%

Middle context: *[...] cuando ella me las explica,* ***me confundo (*a mí mismo)*** *aún más.*						
	Intermediate		Advanced		Near native	
	#	%	#	%	#	%
Yes	1	33.33%	0	0%	4	30.77%
No	**1**	**33.33%**	**12**	**92.31%**	**9**	**69.23%**
I don't know	1	33.33%	1	7.69%	0	0%
No response	0	0%	0	0%	0	0%

As was the case for L1 English speakers, individual verbs did not demonstrate notable effects. I present here two daily routine verbs for which participants' responses were less congruent with the expectation, when the reflexivity expressed in the context allowed the use of *a sí mismo*:

Table 45. L1 Korean speakers' distribution of responses by verbs (Task 3)

			#	%
despertarse **' to wake up'**	Routine - Reflexive	Yes	18	62.07%
		No	11	37.93%
		I don't know	0	0%
	Routine - Middle	Yes	5	17.86%
		No	22	78.57%
		I don't know	1	3.57%
peinarse **'to comb one's hair'**	Routine - Reflexive	Yes	18	62.07%
		No	10	34.48%
		I don't know	1	3.45%
	Routine - Middle	Yes	5	17.24%
		No	21	72.41%
		I don't know	3	10.35%

Although the expected answer (i.e. 'Yes' in 'Routine-Reflexive' contexts) was selected over 60% of the time for both verbs, the opposite responses in contexts where it lacked reflexivity (i.e. 'No' in 'Routine-Middle' contexts) was more congruent with the expectation at a selection rate of 70%. This result can be interpreted such that even if L1 Korean speakers had perceived reflexivity and associated it with the use of *a sí mismo*, it was easier for them to recognize the lack of reflexivity and refuse its use.

To sum up, the L1 Korean speakers with advanced and near native proficiency acknowledged the correlation between reflexivity and the use of *a sí mismo* in general, while speakers with a lower level of proficiency (i.e. intermediate-level) have not yet developed sufficiently with non-daily routine verbs. It was shown that the near native group performed more congruent with the expectation when the context implied reflexivity, while the advanced group did so when reflexivity was not present in the context.

5.4.3 L1 Spanish speakers' Task 3 results

Finally, the native speaker group's results will be analyzed in comparison with the two non-native groups' results. It was shown that native speakers were less likely to accept the use of *a sí mismo* in any of the contexts, compared to non-native speakers whose responses reflected the (in)existence of reflexivity of the context. In other words, even in contexts where the use of *a sí mismo* is acceptable, the selection rate of including it by native speakers was lower than that of advanced and near native non-native speakers. See the distribution of responses by L1 Spanish speakers below:

Table 46. Distribution of responses by L1 Spanish speakers (Task 3)

(*n*=27)		#	%
Routine - Reflexive	**Yes**	**93**	**57.41%**
	No	57	35.18%
	I don't know	12	7.41%
	No response	0	0%
Routine - Middle	Yes	16	9.88%
	No	**139**	**85.8%**
	I don't know	6	3.7%
	No response	1	0.62%
Non-Routine - Reflexive	**Yes**	**85**	**62.96%**
	No	44	32.59%
	I don't know	6	4.45%
	No response	0	0%
Non-Routine - Middle	Yes	15	11.11%
	No	**118**	**87.41%**
	I don't know	1	0.74%
	No response	1	0.74%

Interestingly, as mentioned above, the L1 Spanish speakers did not respond as expected in contexts where reflexivity was implied. Even though the percentage of choosing the expected response in reflexive contexts (i.e. 'Yes' in 'Routine-Reflexive' and 'Non-Routine-Reflexive' contexts) was over 50%, the rate of rejecting *a sí mismo* in non-reflexive contexts (i.e. 'No' in 'Routine-Middle' and 'Non-Routine-Middle' contexts) was more explicit, at 85.8% and 87.41%, respectively. In reflexive contexts, conformity to the expectation was smaller with daily routine verbs (57.41%), compared to non-daily routine verbs (62.96%). This unanticipated result is even more surprising when compared to the non-native speakers' responses.

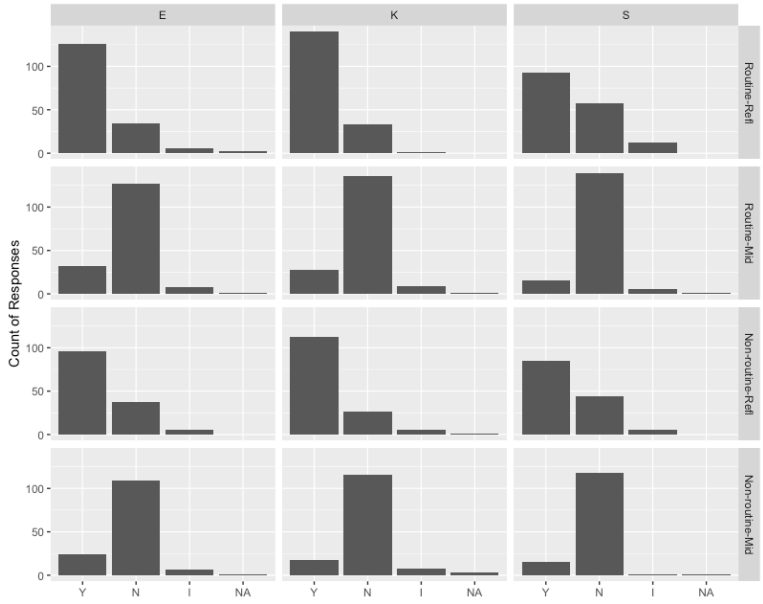


Figure 24. Count of responses by all participant groups (Task 3)

The first and third rows ('Routine-Refl' and 'Non-routine-Refl') reveal how different the pattern of responses was for native speakers (right column 'S' (L1 Spanish)), compared to the two non-native speaker groups' results (left and middle columns 'E' (L1 English) and 'K' (L1 Korean)). While the count of responses of option 'Y' (Yes) is noticeably higher than 'N' (No) in reflexive contexts (row 'Routine-Refl' and row 'Non-routine-Refl') for English- and Korean-speakers (column 'E' (L1 English) and column 'K' (L1 Korean)), this difference between 'Y' (Yes) and 'N' (No) is smaller for native speakers (column 'S' (L1 Spanish)) in the same reflexive contexts (row 'Routine-Refl' and row 'Non-routine-Refl'). On the other hand, the contrast between the native and non-native groups in non-reflexive contexts (row 'Routine-Mid' and row 'Non-routine-Mid') is not remarkable.

An ANOVA was conducted on a regression model with the factors L1, context type, and the interaction between the two and revealed that all three factors had a significant effect on the participants' responses. Context type was the most significant factor, $F(3, 1836) = 205.07$, $p <$

2e-16, followed by L1, $F(2, 1836) = 9.92$, $p = 5.21e\text{-}05$. Finally, a significance of the interaction was also found, $F(6, 1836) = 3.93$, $p < .001$. A post-hoc analysis with Tukey's test revealed that the two non-native speaker groups were not significantly different from each other, $p > .05$. However, the difference between the L1 English and L1 Spanish speakers was significant, $p < .01$. The strongest significance appeared between the L1 Korean and L1 Spanish speaker groups, $p < .001$. The statistical results for context type and its interaction with L1 are illustrated in the mosaic plot below.

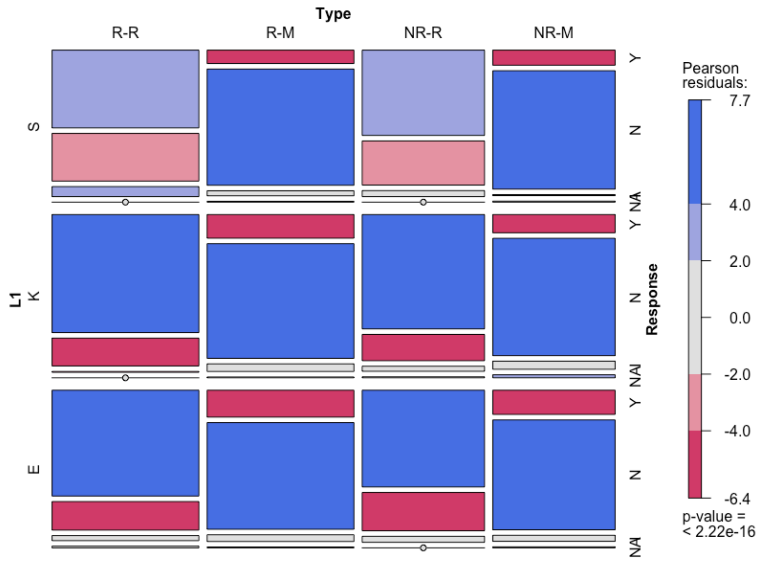


Figure 25. Significance of relevance between response, L1, and context (Task 3)

As can be seen above, the two non-native speaker groups performed uniformly: there was a strong favoring effect on the use of *a sí mismo* when its use was expected (i.e. in reflexive contexts), and a strong disfavoring effect on its use when unexpected (i.e. in middle contexts). The same disfavoring effect was also found in the L1 Spanish speakers' results, albeit to a lesser degree. Also, the native speakers slightly favored the response 'I don't know' when a daily

routine verb expressed reflexivity. Table 47 below showcases how different L1 Spanish speakers' responses were in one of the reflexive daily routine contexts provided in the task:

Table 47. Participants' acceptance of *a sí mismo* in daily routine reflexive context

Reflexive context: *"Vengan niños, les voy a enseñar a* ***peinarse (a sí mismo)****. [...]"*						
	ENG		KOR		SPAN	
	#	%	#	%	#	%
Yes	**19**	**67.86%**	**18**	**62.07%**	**12**	**44.44%**
No	6	21.43%	10	34.48%	12	44.44%
I don't know	3	10.71%	1	3.45%	3	11.12%
No response	0	0%	0	0%	0	0%

This unpredicted result leads us to hypothesize that the L1 Spanish speakers were less likely to associate *se*'s reflexive marking function with its extended form, *a sí mismo*. The results from non-native speakers make it possible to assume that L2 learners behave as they are taught in L2 classroom, even though native speakers find the given structure unnatural. However, at the same time, they gradually reach a point where they acknowledge that not all *se* constructions are reflexive and reject the use of *a sí mismo* in middle contexts, although they have not received any explicit instructions on this matter.

Analyzing the effect of specific verbs might contribute to our understanding of how native speakers interpret the association between reflexivity and the use of *se* with *a sí mismo*. See Table 48 below.

Table 48. Distribution of responses of the verbs that showed unexpected results (Task 3)

			ENG		**KOR**		**SPAN**	
			#	%	#	%	#	%
Routine - Refl	*despertarse* 'to wake up'	Yes	16	59.26%	18	62.07%	**12**	**44.44%**
		No	10	37.04%	11	37.93%	**12**	**44.44%**
		I don't know	1	3.7%	0	0%	3	11.12%
	levantarse 'to get up'	Yes	22	81.48%	27	93.1%	**12**	**44.44%**
		No	4	14.82%	2	6.9%	**13**	**48.15%**
		I don't know	1	3.7%	0	0%	2	7.41%
	peinarse 'to comb one's hair'	Yes	19	67.86%	18	62.07%	**12**	**44.44%**
		No	6	21.43%	10	34.48%	**12**	**44.44%**
		I don't know	3	10.71%	1	3.45%	3	11.12%
Non-Routine - Refl	*perderse* 'to lose oneself'	Yes	0	0%	0	0%	0	0%
		No	**18**	**64.29%**	10	34.48%	**23**	**85.19%**
		I don't know	0	0%	1	3.45%	1	3.7%
	verse 'to see oneself'	Yes	24	85.72%	26	89.65%	**12**	**44.44%**
		No	3	10.71%	2	6.9%	**15**	**55.56%**
		I don't know	1	3.57%	1	3.45%	0	0%

It was only the L1 Spanish speaker group that performed unexpectedly with three daily-routine verbs (*despertarse* 'to wake up', *levantarse* 'to get up' and *peinarse* 'to comb one's hair') listed in the table above. That their 'Yes' and 'No' responses did not differ greatly is interesting given that these two verbs were provided in reflexive contexts and therefore compatible with *a sí mismo*. It is additionally notable that the native speakers' responses diverged from that of the non-native speakers. Moreover, the native speakers' results were even more disparate from the expectation with two non-daily routine verbs (*perderse* 'to lose oneself' and *verse* 'to see oneself'). This outcome can be interpreted such that the Spanish clitic *se* is less likely to be

perceived as a reflexive marker by native speakers than by L2 learners of Spanish. Based on the textbook analysis described in 3.4, non-native speakers' strong association between clitic *se* and reflexivity might derive from the fact that *se* is first (and/or only) introduced as a reflexive marker. Hence, it is plausible that L2 learners mold a one-to-one mapping relationship between reflexivity and *se* with *a sí mismo*.

To sum up, the L1 Spanish speakers' performance on Task 3 was surprisingly different from the expectation. Particularly, their judgments in contexts where reflexivity was present did not behave as expected, due to the way in which their degree of favoring the use of *a sí mismo* in reflexive contexts was not as significant as the expectation. Meanwhile, the L1 English and L1 Korean speakers responded close to the expected pattern, although their degree of conformity differed according to their proficiency level in the target language. As non-native speakers performed more closely to the expected results in contexts where *a sí mismo* can be used and given that these expected results increased as the learners' proficiency level increased, it appears that the association between reflexivity and the use of *se* with *a sí mismo* for native and non-native speakers is contrastive. In other words, L2 learners of Spanish strongly associate reflexivity and *se* with *a sí mismo*, whereas native speakers do not.

Finally, I summarize the findings of the three tasks discussed so far. In Task 1 which was designed to examine L1 and L2 Spanish speakers' production of the clitic *se*, the three participant groups performed in a similar manner: (i) the rate of *se*-marking averaged at 20% with a range from 19% to 24% per group; and (ii) all groups favored *se*-marking with telic verbs while disfavored it with atelic verbs. The difference between native and non-native speakers was that native speakers showed preference for *se*-marking in present tense, whereas non-native speakers disfavored it. It was preterite where non-native speakers were more likely to use *se*-

marked forms; however, native speakers did not. This divergence leads us to assume that native speakers might have employed the historical or narrative present, resulting in more present tense tokens in the data itself. However, as for the non-native speakers, it is plausible that their *se*-marking is affected by the aspectual properties of the event such as boundedness and delimitedness.

The remaining two tasks, Task 2 (AJT testing learners' knowledge of absolute *vs.* dynamic verbal events) and Task 3 (cloze task testing learners' understanding of reflexives *vs.* middles), revealed that L2 learners gradually behave more as expected as their proficiency level increases. Specifically, (i) L2 learners' association between dynamicity of the event and *se*-marking becomes stronger as their L2 competence develops, and (ii) L2 learners differentiate reflexive *se-a sí mismo* to middle *se-*a sí mismo* more categorically in their later stages of L2 acquisition. However, native speakers of Spanish deviated from our expectation to a certain degree in that (i) they accepted both *se*-marked and non-*se*-marked forms in dynamic contexts, and (ii) they refused to include *a sí mismo* in general, regardless of the context. These findings suggest that native speakers' judgment is less likely to be affected by the context provided in task items but they rather express their own perspectives of the event and indicate their own interpretations by (not) using *se* and *a sí mismo*. Meanwhile, L2 learners seem to develop their knowledge of *se* both by implicit (the correlation between *se*-marking and dynamicity and the incompatibility of *a sí mismo* with middle *se* constructions) and explicit (the compatibility of *a sí mismo* with reflexive *se* constructions) input, although the latter is not accepted as natural linguistic behavior by native speakers.

In the next chapter, I will summarize and discuss these findings as they relate to the research questions and confirm or disconfirm the hypotheses presented. I will also consider the

contributions these findings make to the fields of Second Language Acquisition, cognitive linguistics, and foreign language teaching, and point out some of the limitations of this study and suggest possible directions for future research.

CHAPTER 6

CONCLUSIONS

6.1 Summary and general conclusions

The present dissertation has explored how the interpretation and production of the Spanish middle marker *se* are similar and different across three speaker groups: adult English- and Korean-speaking L2 learners of Spanish, and L1 Spanish speakers. One of the main purposes of this dissertation was to report the findings of experimental research into the acquisition of the Spanish clitic *se* as a middle marker by each participant group. Given the lack of discussion on the topic in the literature, the present study is expected to give meaningful insights on the applicability of the cognitive approach to explaining some of the difficulties L2 learners from different L1 backgrounds have in acquiring the Spanish middle marker *se* in the foreign language classroom.

From the framework of cognitive linguistics, Maldonado (2000, 2008, among others) refutes the traditional and generative analyses of the Spanish clitic *se* claiming that not all instances of it are reflexive with a valence reducing or detransitivizing effect; hence the non-reflexive uses of *se* should not be considered as idiosyncratic and exceptional cases that need to be memorized by L2 learners. Specifically, the detransitivization analysis fails to explain the presence of *se* in intransitive constructions where there is no valence that can be reduced (e.g. ***Se*** *cayeron tres vasos* 'Three glasses fell') or in transitive constructions with both subject and object

arguments (e.g. ***Se*** *tomó una copa de vino de un trago* 'S/he drank up a glass of wine in one gulp').

Instead, by determining the use of *se* in non-reflexive contexts as a middle marker, whose main function is to focus on the central properties of events, it is possible to generalize all so-called idiosyncratic *se* constructions into one single category, the middle domain. In particular, middle constructions represent actions, events or states that belong to the domain of the subject, therefore, the function of *se* is to modify the conceptualization of an event in a way that focuses on the change of state designated by the verb. Several L1 studies of the acquisition of Spanish clitic *se* described in this dissertation (cf. 3.2) generally gave support to this concept of middle domain with results showing that clitic *se* is employed mainly to mark dynamic and energetic events and that middle *se* is acquired early on (even before reflexive *se*, for some children).

As opposed to the findings of L1 acquisition studies, L2 studies revealed that clitic *se* constructions are late-acquired, in general. Maldonado attributes the difficulty in acquiring the uses of *se* to the fact that traditional and pedagogical approaches erroneously present all of the instances of *se* as reflexive constructions, making it even more difficult for L2 learners to understand that the one-to-one correspondence of *se* and its reflexive function is not always the case. A textbook analysis conducted in this dissertation (cf. 3.4) demonstrated that English textbooks used the term 'reflexive pronoun' even when describing non-reflexive *se* constructions and that Korean textbooks emphasized the detransitivizing function of *se* in their explanations. Based on these findings, it was possible to assume that the misrepresentation of the linguistic nature of the clitic *se* might mislead the L2 learners' understanding of Spanish *se* constructions, as Maldonado suggested.

In addition, the discrepancy between the middle system of L1 and the target language is also expected to affect how L2 learners interpret and produce Spanish middle *se* constructions. The two languages in question, English and Korean, possess a quite distinct middle system from that of Spanish. In English, there is no explicit morpheme which marks middle, although it shows other similar characteristics such as the presence of adverbs (e.g. *easily, well*) in dispositional middle constructions. On the other hand, Korean is relatively closer to Spanish in that there is a distinct middle marker, *-i*, despite the controversial debate regarding whether it is a middle or a passive marker. However, Spanish and Korean are different in that Korean uses one morpheme (*-i* and its phonological variants) to mark middles and a different morpheme (*casin*) to mark reflexivity.

With the purpose of determining the L2 acquisitional pattern of the middle marker *se*, the present dissertation conducted a linguistic survey to compare two L2 Spanish learner groups with different L1 backgrounds, English and Korean, with a group of native speakers of Spanish. Through a statistical analysis of the interpretation and the production data of the Spanish clitic *se* among 33 L1 English speakers, 30 L1 Korean speakers, and 30 L1 Spanish speakers, I was able to draw several conclusions. In order to discuss these conclusions, I revisit the research questions and hypotheses posed in this dissertation and offer answers to help us understand the production and interpretation of the Spanish clitic *se* by English- and Korean-speaking L2 learners of Spanish, and native Spanish speakers. Each of the research questions and their corresponding hypotheses are repeated below for convenience. First, I revisit the three specific questions and hypotheses, each of which was related to the three experimental tasks of the present study, with an objective of summarizing the interesting results of each task. Later, I discuss the first two

general questions to discuss the L2 acquisition process of Spanish *se*-marking observed in the survey data in broader and more general terms.

RQ 3: What are the linguistic and non-linguistic factors or variables that favor *se*-marking among L2 learners and native Spanish speakers?

H3: Regarding linguistic factors, it is expected that (i) transitive predicates, (ii) null subject form, (iii) postverbal subjects, (iv) past tense forms, and (v) telic verbs will favor the use of *se*. In the same vein, the lack of these factors will disfavor the use of *se*. As for the extralinguistic variables, L2 learners with higher levels of proficiency and/or longer previous experience living in Spanish-speaking countries will produce more native-like *se* constructions, given that previous literature (cf. Escutia López 2010, Tremblay 2005) claims that *se* constructions are late-acquired by non-native speakers.

Task 1, the picture book written narration task, gave insight into the production of the middle marker *se*. The three participant groups exhibited similar frequencies of *se*-marking, although the native speakers produced more tokens (both *se*-marked and non-*se*-marked) than the non-native speakers (i.e. 663 and 627 for English- and Korean-speaker groups, respectively *vs.* 773 for Spanish-speaker group). Through a mixed-effects multivariate analysis via Rbrul, it was revealed that only one factor group, *Telicity*, indicated a statistically significant finding: telic verbs were significantly more likely to be marked with *se* than atelic verbs, across all speaker groups. This pattern was stronger for non-native speakers, a finding which leads us to believe that non-native speakers tend to be affected by the bounded or delimited nature of the event and associate it with the use of *se*. Furthermore, the results that telic verbs favor *se*-marking suggests that both native and non-native Spanish speakers employ the clitic *se* to describe spontaneous events or to convey

emphasis on the pivotal moment of change of state, supporting the thesis of cognitive linguistics on clitic *se*'s middle marking function.

On the other hand, the differences among the three groups were found in the factor *Tense-Aspect-Mood*. The two non-native speaker groups disfavored *se*-marking in the present tense, whereas it was favored by the Spanish speakers. Conversely, the Spanish speakers disfavored *se*-marking in the preterite, while the L2 learners favored it, although the favoring effect was not statistically significant. Again, this might be a reflection of the stronger effect of boundedness or delimitedness of the event on non-native speakers. However, it is also attributable to the fact that native speakers produced more tokens in the present tense in general, choosing to narrate the picture story in the present rather the past tense, which resulted in more *se*-marked tokens in present tense.

RQ 4: Do learners and native speakers associate absolute verbal events with non-*se*-marking and dynamic events with *se*-marking?

H4: As shown by L1 acquisition studies from the framework of cognitive linguistics (cf. Jackson-Maldonado et al. 1988, Soto & Muñoz 2000), native speakers are expected to associate *se*-marking with the dynamicity of the event. However, since this association is not explicitly taught in the L2 classroom, lower-level L2 learners are more likely to diverge from the expected patterns, while learners with greater proficiency will behave more expected or native-like and will accept *se*-marking in sentences with a dynamic context and accept non-*se*-marking in sentences with an absolute context.

Answers to this research question are rooted in the results and discussion of Task 2, the acceptability judgment task. Specifically, the results of Task 2 were analyzed with a special

interest in examining if L2 learners are also capable of differentiating the *se*-marked dynamic middles from their non-*se*-marked absolute counterparts as Spanish-speaking children do. The overall pattern of performance by non-native speakers lined up with the expected pattern, although the intermediate-level learners' judgments from both L1 English and L1 Korean groups did not.

The intermediate group highly preferred the non-*se*-marked forms in both absolute and dynamic contexts and this pattern leads us to propose that the non-*se*-marked is their default verbal form. However, in the advanced and near native groups' data, it was shown that the context type significantly affected how participants understood the use of *se*. In absolute contexts, their responses revealed the pattern of favoring the bare verbal form (without *se*-marking). In dynamic contexts, they tended to judge more often that *se*-marked constructions sounded natural, but the rate of this response was even more frequent in native speakers' data. In other words, the association between the middle marker *se* and the dynamicity/middleness of the context was significantly stronger for native speakers. Such an association can be explained by Langacker's (1991) explanation of the cognition of verbal events: in absolute contexts, there is no energy put into profile, hence the use of *se* is limited; however, in dynamic contexts, special attention is focused on the energy involved in the event and speakers can choose to express such focus by *se*-marking. Moreover, this pattern of performance by native speakers is similar to the results of Anderson (1998) discussed earlier in 3.2, which showed that native Spanish speakers' production of *se* varied depending on how they interpret and perceive the context.

RQ 5: Do learners and native speakers understand the difference between true *se*-marked reflexives as evidenced by the acceptance of the addition of the phrase *a sí mismo*, and *se*-marked middles as evidenced by the rejection of the phrase?

H5: Since the majority of, if not all, Spanish classroom instruction and textbooks introduce *se* as a reflexive marker, L2 learners will first make a one-to-one correspondence between *se* and its identification as a reflexive marker. Therefore, lower-level learners, not being aware of the middle marking function of *se*, will over accept sentences with *a sí mismo* even in the constructions where it is unacceptable or pragmatically infelicitous.

The fifth and final research question can be answered with the results of Task 3, the cloze task. Specifically, this task was designed to examine whether or not L2 learners are aware of the association between reflexive marker *se* and its extended form *a sí mismo*, and the incompatibility of the *a sí mismo* phrase and the middle marker *se*. It was shown that L2 learners have acquired the aforementioned awareness of the pattern. With a differing degree of conformity according to their proficiency level of the target language, both L1 English and L1 Korean speakers responded close to the expected pattern. That is, in contexts where reflexivity was implied, they associated the use of *se* with the phrase *a sí mismo*. When reflexivity was not present in a given situation, they refused to add the phrase *a sí mismo* to the *se*-marked verb. Hence, there was no overuse of *a sí mismo* by L2 learners, especially with higher proficiency levels, in the constructions where it was unacceptable. The rate of refusing *a sí mismo* in non-reflexive contexts was also strong for native speakers. However, a divergence between the native and non-native speakers appeared in reflexive contexts, in which the native speakers were less likely to use *a sí mismo*. Therefore, it seems the association between reflexivity and *se* with *a sí mismo* was significantly weaker for the L1 Spanish speakers.

That the native speakers rejected *a sí mismo* in instances in which it was in fact grammatical leads to speculation that they judged the inclusion of *a sí mismo* according to

felicitousness rather than grammaticality. It is plausible that adding *a sí mismo* in such contexts might seem redundant or excessive for native speakers. This finding was surprising given that the contexts presented in Task 3 for the *se*-marked reflexive sentences with *a sí mismo* were taken from excerpts of naturally occurring native language. In other words, native speakers tend to use *a sí mismo* in addition to the reflexive marker *se* only in instances where there was a need for clarification or emphasis that *a sí mismo* offers. On the other hand, it is possible that L1 English and L1 Korean speakers are less likely to perceive *a sí mismo* as redundant, considering that their respective L1s express reflexivity via lexical items (i.e. *oneself* in English, and *casin* in Korean)[33], whereas employing a clitic (i.e. reflexive marker *se*) is sufficient in Spanish as discussed by the RAE (2010) that the pronoun *sí mismo* "is usually interpreted as a way to emphasize in the syntax a (reflexive) relationship that the morphology already expresses" (2010:1189, [translation mine]). Furthermore, this use of the *a sí mismo* phrase is reinforced in language textbooks and classes.

RQ 1: Are there differences between L1 English and L1 Korean speakers' knowledge and acquisition of *se*-marking in Spanish and how does this compare to native Spanish speakers' knowledge and use?

H1a: Not only do English and Korean middle constructions differ from Spanish middle constructions, but there are also differences between the two native languages of the learner groups. If the speakers' L1 has a stronger impact on transfer, the results

[33] Examples (38) and (39) are adapted and repeated here for convenience:

English: María looked at **herself** in the mirror.

Korean: *María-nun* *kewul-eyse* ***casin*-*ul*** *po-ass-ta.*

María-NOM mirror-in **REFL**-ACC see-PST-DECL

Spanish: *María* ***se*** *vio en el espejo* ***(a sí misma)***.

will also be different for each group of L2 speakers, based on their respective L1s. Also, the discrepancy of the middle system between the three languages will result in non-target-like performance by L2 learners.

H1b: In general, L1 Korean speakers acquire English as their L2 and Spanish as their L3. If L3 interpretation and production is more greatly affected by L2 transfer than by L1 transfer, English-speaking learners' and Korean-speaking learners' use of *se* will not differ to a great extent. In this case, it will be possible to assume that L2 learners' non-target-like behavior might be attributable to certain linguistic properties of English.

Overall, there was no significant difference between the performance of L1 English speakers and that of L1 Korean speakers found in the results of all three tasks. As I mentioned in 4.2, that English- and Korean-speaking L2 Spanish learners performed similarly does not necessarily indicate that Korean learners' performance is affected by English, as the present study is not designed to be an L3 acquisition analysis. It might be the case that there are certain properties of English that have a stronger impact on Korean speakers' production and interpretation of Spanish clitic *se.* However, it is entirely possible that the resemblance between the two non-native groups' performance derives from the fact that both English and Korean do not employ clitics for marking reflexivity and middleness. Another explanation for the lack of difference between the two L1 groups is that the L1 does not play a significant role at the later stages of L2 development, as Bruhn de Garavito (1999a:289) suggests.

Regarding the difference between non-native and native speakers' knowledge of *se*-marking, a certain degree of unexpectedness was shown in the native speaker data while the most proficient non-native speakers demonstrated a pattern more similar to the expectations (i.e.

strong association between *se*-marking and dynamicity/middleness of the event and categorical distinction between reflexive *se* and the *a sí mismo* phrase and non-reflexive middle *se* and its non-compatibility with *a sí mismo*). This discrepancy leads us to assume that non-native speakers' judgment is more likely to be bound by the context, whereas native speakers' judgment is based on the way they prefer to express their own interpretation of the event. In other words, native speakers are not constrained by the context but rather use the morphological marker *se* to render a conceptualization of the context that they as speakers determine; however, for non-native speakers, certain linguistic properties such as adverbs used in the passage and lexical/aspectual characteristic of the predicate might aid in their interpretation of the context, which affects their judgment on *se*-marking and the *a sí mismo* phrase.

RQ 2: What are the differences among the different proficiency levels in terms of their knowledge and acquisition of *se*-marking and how does this compare to native Spanish speakers' knowledge and use?

H2: The Spanish clitic *se* is well known for its variable use among native speakers, especially dialectally (Ibarretxe-Antuñano et al. 2016:157). Moreover, some constructions require an obligatory use of *se* (e.g. *quejarse* 'to complain'), while other verbs are possible in both non-*se*-marked and *se*-marked forms (e.g. *ir/irse* 'to go'/'to leave'). The inconsistency or variable use could cause learnability problems for L2 learners; thus, the interpretation and production of middle marker *se* will show differences from native speakers. Especially, the lower-level learners who have received limited input of *se* constructions are expected to demonstrate more non-target-like performance. However, based on the fact that the advanced learners are exposed more to middle constructions, both inside and outside of the classroom, L2

> learners' performance will gradually become more native-like as the proficiency level increases. Furthermore, L2 learners' production will demonstrate greater divergence than their interpretation, given that active language skills are acquired after passive skills.

First, as for the production data, the amount of *se*-marked phrases was relatively the same across all three participant groups. Besides the evidence that L2 learners may rely on delimitedness, a feature of telic verbs and the preterite, more so than native speakers in *se*-marking, no other remarkable divergence between the non-native and native speakers of Spanish was found in the production data in terms of linguistic factors that favored or disfavored *se*-marking.

Contrary to my hypothesis, it was the interpretation of *se* which showed a greater difference between native and non-native speakers. According to Maldonado (1992, 2008, among others), *se* functions as a middle marker in instances where there is (i) inseparable representation of the subject and object, (ii) lack of agentivity of the experiencer subject, (iii) focus on the event itself, and (iv) indication of abruptness and unexpectedness. This general, overriding function of middle marker *se* makes it possible to connect seemingly idiosyncratic *se* constructions under a single category, the middle domain. However, such an understanding of *se*'s middle marking function is rarely, if not ever, presented to L2 learners, which might be the cause of the non-native-like interpretation on various *se* constructions in the present study. In the interpretation data, the context type significantly affected how participants understood the use of *se*. In absolute contexts, all three groups' responses showed the same pattern of favoring the bare verbal form without *se*-marking. However, native speakers were more likely than L2 learners to judge that *se*-marked constructions sounded more natural in dynamic contexts. The association between the middle marker *se* and the dynamicity of the context was significantly stronger for

native speakers. Also, in accepting *a sí mismo* in reflexive *se* constructions and rejecting it in middle ones, contrary to the expectation, native speakers disfavored the inclusion of *a sí mismo* in general regardless of the contexts; while non-native speakers revealed more categorical judgments, as expected.

Furthermore, it was observed in the interpretation data that the learners' proficiency level of Spanish also affected their performance. The conformity of their responses to the expectation increased as their proficiency level increased. This correlation suggests that the learners' knowledge of the middle function of *se*-marking increases as their competence in Spanish develops. In other words, I posit that non-native speakers' judgments on the use of *se* gradually relies more on the nature of the context as they become more proficient in Spanish.

6.2 Contributions and future directions

The current study aimed to examine the L2 acquisitional process of Spanish middle marker *se*, an area generally understudied in the field of Second Language Acquisition (SLA) and cognitive linguistics. Specifically, this dissertation has contributed to both fields by exploring and answering questions about L2 learners' knowledge of the clitic *se*'s middle marking function in non-reflexive *se*-marked constructions. The present findings support this thesis of middle marker *se* put forth by cognitive linguistics, by showing that L2 learners' production and interpretation of clitic *se* are not limited only to reflexive constructions. In other words, non-native speakers who participated in this study demonstrated their gradual acquisition of the native-like/expected response patterns along the development of their proficiency in the target language: (i) contrastive use of reflexive *vs.* middle *se* by accepting and rejecting the phrase *a sí mismo* according to the context and (ii) understanding of the event-focusing function

of middle marker *se* by differentiating non-*se*-marked absolute contexts to *se*-marked dynamic ones. Moreover, by examining data of Korean-speaking L2 learners of Spanish, the current research has provided some interesting observations on a demographic group which has received little attention in Spanish SLA studies in addition to those of a more widely studied group, L1 English learners of L2 Spanish.

Finally, the pedagogical implications of this research are evident. Learners at the beginning and even intermediate levels have a difficult time conceptually and grammatically understanding the polysemic nature of Spanish *se*-marking. This may have to do with how the clitic *se* is presented and practiced in language textbooks and classrooms. By teaching early on the principal function of *se* is to focus on the central properties of verbal events, perhaps learners would avoid a great deal of confusion. Nonetheless, the results of this study suggest that, despite this confusion, more advanced learners eventually come to understand this principal function and do learn to distinguish between absolute and dynamic events and to distinguish between the reflexive and middle functions of the marker.

However, despite these conclusions, the results are based on the limited amount of data obtained from a small sample size. Thus, it will be of interest to study a larger number of participants in each group to examine in greater detail how both native and non-native speakers produce and interpret the middle marker *se* in Spanish. Specifically, in addition to the picture book written narration data, analyzing other types of production, such as oral narration data or conversational data, will give us greater insights into how *se*-marking differs in written *vs.* spoken narration and narrative *vs.* conversational discourse. Also, more open-ended written and oral data can provide different and interesting results to compare and contrast to the results of the present study. Furthermore, this dissertation research examined the production and interpretation

of the clitic *se* among English- and Korean-speaking L2 learners of Spanish; future research could examine other L1 groups from different language families including those that have morphological middle marking and those that do not have these markers.

REFERENCES

Aaron, Jessi Elena & Rena Torres Cacoullos. 2006. *Te sales:* ¿Puede cuantificarse la subjetivización? *Octavo Encuentro Internacional de Lingüística en el Noroeste, 2º Tomo,* ed. by M. Ortiz Ciscomani, 43-63. Hermosillo, Sonora: Unison.

Alexiadou, Artemis. 2012. Noncanonical passives revisited: Parameters of nonactive Voice. *Linguistics* 50(6). 1079-1110.

Alexiadou, Artemis. 2014. Active, middle, and passive: the morpho-syntax of Voice. *Catalan Journal of Linguistics* 13. 19-40.

Alexiadou, Artemis & Edit Doron. 2012. The syntactic construction of two non-active Voices: Passive and middle. *Journal of Linguistics* 48. 1-34.

Amores, María, Anne Wendel & José Luis Suárez-García. 2020. *Experience Spanish, 3rd Edition*. New York: McGraw-Hill Education.

Anderson, Raquel T. 1998. The use of reflexive constructions by Spanish-speaking children: Differences across functions. *Applied Psycholinguistics* 19. 489-512.

Araya, María Teresa & Anabel Monteserin. 2011. Usos del clítico *se* en textos de distintos niveles de CELU. *Actas del VI Coloquio CELU: "El CELU como meta y punto de partida en la enseñanza de Español como Lengua Segunda o Extranjera"*. Argentina.

Armstrong, Grant. 2013. Agentive reflexive clitics and transitive *se* constructions in Spanish. *Borealis: An international Journal of Hispanic Linguistics* 2(2). 81-128.

Arrese, Juana. 1999. Conceptualization of events, semantic relations between constructions and topology: A catastrophe theoretic study of *get* and *be*. *Journal of English Studies* I. 97-117.

Bailey, Carolina. 2013. *English Native Speakers' L2 Acquisition of the Spanish Clitic* se. Ph.D. dissertation. The University of Wisconsin-Madison.

Balcom, Patricia. 2003. Cross-linguistic Influence of L2 English on Middle Constructions in L1 French. *Effects of the Second Language on the First*, ed. by Vivian Cook, 168-192. Clevedon, UK: Channel View Publications.

Basilico, David. 2010. The *se* clitic and its relationship to paths. *Probus* 22. 271-302.

Bartra, Anna. 2014. Voice phenomena: how many properties behind this label? *Catalan Journal of Linguistics* 13. 3-18.

Bayona, Patricia. 2005. "La pasta se come con salsa": Acquisition of middle constructions in Spanish. *Proceedings of the 2005 annual conference of the Canadian Linguistic Association*, ed. by Claire Gurski, 125-137. The University of Western Ontario.

Bello, Andrés. 1891. *Gramática de la lengua castellana.* Madrid: Librería de Leocadio López.

Bruhn de Garavito, Joyce L. S. 1999a. The *se* construction in Spanish and near-native competence. *Spanish Applied Linguistics* 3. 247-295.

Bruhn de Garavito, Joyce L. S. 1999b. *The syntax of Spanish multifunctional clitics and near-native competence*. Ph.D. dissertation. McGill University.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, aspect, and modality in the languages of the world.* Chicago: Chicago University Press.

Calude, Andreea S. 2017. Testing the boundaries of the middle voice: Observations from English and Romanian. *Cognitive Linguistics* 28(4). 599-629.

Cinque, Guglielmo. 1988. On *Si* Constructions and the Theory of *Arb. Linguistic Inquiry* 19(4). 521-581.

Cuervo, María Cristina. 2014. Alternating unaccusatives and the distribution of roots. *Lingua* 41. 48-70.

Doron, Edit. 2003. Agency and Voice: The Semantics of the Semitic Templates. *Natural Language Semantics* 11. 1-67.

Elliott, Otis Phillip, Jr. 1995. *A Glance at the Syntactic and Semantic Principles Underlying the Spanish Clitic* se*: A Study in Second Language Acquisition.* Ph.D. dissertation, The University of Arizona.

Escobar, Linda & Ismael Teomiro. 2016. The gradual acquisition of clitic "se" in Spanish L2. *Topics in Linguistics* 17(1). 17-29.

Escutia López, Marciano. 2010. El uso de *se* con verbos inacusativos por estudiantes avanzados de español como lengua extranjera: Transferencia y restructuración. *Revista Española de Lingüística Aplicada* 23. 129-151.

Escutia, Marciano. 2016. L2 Spanish preverbal 'se' in analysis and production data. *Revista Española de Lingüística Aplicada* 29(1). 30-63.

Flynn, Suzanne, Claire Foley & Ina Vinnitskaya. 2004. The cumulative-enhancement model for language acquisition. Comparing adults' and children's patterns of development in first, second and third language acquisition. *International Journal of Multilingualism* 1(1). 3-17.

Folli, Raffaella & Heidi Harley. 2005. Flavors of *v*: Consuming Results in Italian & English. *Aspectual Inquires. Studies in Natural Language and Linguistic Theory, vol 62,* ed. by Paula Kempchinsky & Roumyana Slabakova, 95-120. Dordrecht: Springer.

García Mayo, María del Pilar & Jason Rothman. 2012. L3 morphosyntax in the generative tradition: The initial stages and beyond. *Third Language Acquisition in Adulthood*, ed. by Jennifer Cabrelli Amaro, Suzanne Flynn & Jason Rothman, 9-32. Amsterdam: John Benjamins.

Gili y Gaya, Samuel. 1955. *Curso superior de sintaxis española.* Barcelona: Publicaciones y Ediciones Spes.

Givón, Talmy & Lynne Yang. 1994. The rise of the English *get*-passive. *Voice: Form and function*, ed. by Barbara Fox & Paul Hopper, 119-150. Amsterdam/Philadelphia: John Benjamins.

Gómez Soler, Inmaculada. 2015a. The logical problem of second language acquisition of argument structure: Recognizing aspectual distinctions in Spanish psych-predicates. *International Journal of Bilingualism* 19(6). 627-645.

Gómez Soler, Inmaculada. 2015b. Acquisitional patterns of Spanish anticausative *se*: The end of the road. *Spanish Journal of Applied Linguistics* 28(2). 349-381.

Grimshaw, Jane. 1982. On the Lexical Representation of Romance Reflexive Clitics. *Mental Representations of Grammatical Relations,* ed. by Joan Bresnan, 87-148. Cambridge: MIT Press.

Guillén Solano, Patricia. 2015. Adquisición de las construcciones medias en español como segunda lengua. *Revista de Filología y Lingüística de la Universidad de Costa Rica* 41(1). 87-101.

Hatcher, Anna G. 1949. To get/be invited. *Modern Language Notes* 64. 433-446.

Ibarretxe-Antuñano, Iraide, Teresa Cadierno & Alberto Hijazo-Gascón. 2016. The role of force dynamics and intentionality in the reconstruction of L2 verb meanings: A Danish-Spanish bidirectional study. *Review of Cognitive Linguistics* 14(1). 136-160.

Jackson-Maldonado, Donna, Ricardo Maldonado & Donna J. Thal. 1998. Reflexive and Middle Markers in Early Child Language Acquisition: Evidence from Mexican Spanish. *First Language* 18. 403-429.

Jespersen, Otto. 1964. *A Modern English Grammar on Historical Principles, Part VI, Morphology.* London: George Allen and Unwin Ltd.

Johnson, Daniel E. 2009. Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3(1). 359-383.

Kanwit, Matthew & Margaret Lubbers Quesada. 2018. Learner and native-speaker differences in the acceptability of *gustar*-type psychological verbs in Spanish. *International Review of Applied Linguistics in Language Teaching* 56(3). 279-313.

Kelling, Carmen. 2006. Spanish *se*-constructions: The Passive and the Impersonal Construction. *Proceedings of the LFG 06 Conference, Universität Konstanz,* ed. by Miriam Butt & Tracy Holloway King, 275-288. Stanford, CA: CSLI Publications.

Kemmer, Suzanne. 1988. *The Middle Voice: A Typological and Diachronic Study.* Ph.D. dissertation, Stanford University.

Kemmer, Suzanne. 1993. *Middle Voice.* Amsterdam: John Benjamins.

Kemmer, Suzanne. 1994. Middle Voice, Transitivity, and the Elaboration of Events. *Voice: Form and Function*, ed. by P. J. Hopper & B. A. Fox, 179-230. Amsterdam/Philadelphia: John Benjamins.

Kempchinsky, Paula. 2004. Romance *se* as an aspectual element. *Contemporary Approaches to Romance Linguistics: Selected Papers from the 33rd Linguistics Symposium on Romance Languages (LSRL), Bloomington, Indiana, April 2003,* ed. by Julie Auger, J. Clancy Clements & Barbara S. Vance, 238-256. Amsterdam/Philadelphia: John Benjamins.

Kempchinsky, Paula. 2006. Teasing Apart the Middle. *Andolin gogoan: essays in honour of Professor Eguzkitza,* ed. by Itziar Laka & Beatriz Fernández. 531-547. Bilbao: Servicio editorial, Universidad del País Vasco.

Kim, Jae Min. 2013. A study of the Korean Non-Active Voice. *Kenci Inmwunhak* 9. 31-51.

King, Larry D. & Margarita Suñer. 2017. *Gramática española: Análisis y práctica, Third Edition*. Long Grove, IL: Waveland Press.

Labov, William. 1972. *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Langacker, Ronald. 1987. *Foundations of Cognitive Grammar. Vol. 1: Theoretical Prerequisites.* Stanford, California: Stanford University Press.

Langacker, Ronald. 1988. An overview of Cognitive Grammar. *Topics in Cognitive Linguistics*, ed. by Brygida Rudzka-Ostyn, 3-48. Amsterdam/Philadelphia: John Benjamins.

Langacker, Ronald. 1990. *Concept, Image, and Symbol: The Cognitive Basis of Grammar.* Berlin/New York: Walter de Gruyter.

Langacker, Ronald. 1991. *Foundations of Cognitive Grammar. Vol. 2: Descriptive Application.* Stanford, California: Stanford University Press.

Lee, Keedong. 1993. *A Korean Grammar on Semantic-Pragmatic Principles.* Seoul: Hankuk Moon Hwa-Sa.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface.* Cambridge, MA: MIT Press.

Liceras, Juana M. 1999. Linguistic theory, L2 acquisition and pedagogical grammar: On some universal and idiosyncratic properties of Spanish *se*. *Spanish Applied Linguistics* 3(2). 297-309.

MacDonald, Jonathan E. 2017. Spanish aspectual *se* as an indirect object reflexive: The import of atelicity, bare nouns, and leísta PCC repairs. *Probus* 29(1). 73-117.

Macías, Claudia, Hyeon-Kyun Kim, Ho-Joon Yim & Hana Lee Kim. 2020. *Español intermedio*. Seoul: SNU Press.

MacWhinney, Brian. 1995. *The Childes Project: Tools for Analyzing Talk, 2nd edition.* Hillsdale, NJ: Lawrence Eribaum.

MacWhinney, Brian & Catherine Snow. 1985. The child language data exchange system. *Journal of Child Language* 12. 271-296.

Maldonado, Ricardo. 1988. Energetic reflexives in Spanish. *Proceedings of the Annual Meeting of the Berkeley Linguistic Society* 14. 153-165.

Maldonado, Ricardo. 1992. *Middle Voice: The Case of Spanish* se. Ph.D. dissertation, University of California, San Diego.

Maldonado, Ricardo. 1993. Dynamic Construals in Spanish. Studi Italiani di Linguistica Teorica e Applicata XXII 3. 531-566.

Maldonado, Ricardo. 1999. *A media voz: Problemas conceptuales del clítico* se. Mexico: Universidad Nacional Autónoma de México.

Maldonado, Ricardo. 2000. Conceptual distance and transitivity increase in Spanish reflexives. *Reflexives: Forms and functions. Volume 1*, ed. by Zygmunt Frajzyngier & Traci Walker, 153-185. Amsterdam/Philadelphia: John Benjamins.

Maldonado, Ricardo. 2008. Spanish middle syntax: A usage-based proposal for grammar teaching. *Cognitive Approaches to Pedagogical Grammar: A Volume in Honour of René Dirven,* ed. by R. Dirven, T. de Tycker & S. de Knop, 155-196. Berlin/New York: Mouton de Gruyter.

Maldonado, Ricardo. 2009. Middle as a Basic Voice System. *Studies in Role and Reference Grammar,* ed. by L. Guerrero, S. Ibáñez Cerda & V. A. Belloro, 69-109. Mexico: Universidad Nacional Autónoma de México.

Mayer, Mercer. 1969. *Frog, where are you?* New York: Dial Press.

McIntyre, Andrew. 2010. The BECOME=CAUSE hypothesis. *Linguistics* 50(6). 1251-1287.

Montrul, Silvina A. 2004. *The acquisition of Spanish: Morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition.* Philadelphia: John Benjamins.

Montrul, Silvina. 2012. "DELE Proficiency Test". *UCLA International Institute*, https://international.ucla.edu/institute/article/126305

Ordóñez, Francisco & Esthela Treviño. 2011. Impersonals with Passive Morphology. *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, ed. by Luis A. Ortiz-López, 314-324. Somerville, MA: Cascadilla Proceedings Project.

Pérez-Gironés, Ana María & Virginia M. Adán-Lifante. 2019. *Más: español intermedio, 3rd edition*. New York: McGraw-Hill Education.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rand, David & David Sankoff. 1990. Gold Varb 2.1: A Variable Rule Application for the Macintosh. Montréal: Centre de recherches mathématiques, Université de Montréal.

Ranson, Diana L. & Margaret Lubbers Quesada. 2018. *The History of Spanish: A Student's Introduction.* Cambridge, UK: Cambridge University Press.

Raposo, Eduardo & Juan Uriagereka. 1996. Indefinite *Se. Natural Language & Linguistics Theory* 14(4). 749-810.

Real Academia Española. 2009. *Nueva gramática de la lengua española: Manual.* Madrid: Espasa.

Real Academia Española. 2010. *Nueva gramática de la lengua española*. Madrid: Espasa.

Rothman, Jason. 2010. L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27(1). 107-127.

Rothman, Jason & Diego Pascual y Cabo. 2014. Generative Approaches to Spanish Second Language Acquisition. *The Handbook of Spanish Second Language Acquisition,* ed. by Kimberly L. Geeslin, 46-63. Chichester, West Sussex: John Wiley & Sons.

Sánchez, Liliana & Almeida Jacqueline Toribio. 2003. Current Issues in the Generative Study of Spanish Second Language Syntax. *Spanish Second Language Acquisition: State of the science,* ed. by Barbara A. Lafford & Rafael Salaberry, 189-232. Washington, D.C.: Georgetown University Press.

Sanromán Vilas, Begoña. 2009. Las construcciones con *se* en el discurso académico de los aprendices finlandeses de español. *El español en contextos específicos: Enseñanza e investigación. Vol. II*, ed. by A. Vera & I. Martínez, 951-970. Spain: Asociación para la Enseñanza del Español como Lengua Extranjera: Fundación Comillas.

Schwartz, Bonnie D. & Rex A. Sprouse. 1994. Word order and nominative case in nonnative language acquisition: a longitudinal study of (L1 Turkish) German Interlanguage. *Language acquisition studies in generative grammar: papers in honor of Kenneth Wexler from the 1991 GLOW workshops,* ed. by Teun Hoekstra & Bonnie D. Schwartz, 317-68. Philadelphia, PA: John Benjamins.

Schwartz, Bonnie D. & Rex A. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12(1). 40-72.

Seco, Manuel. 1972. *Gramática esencial del español: Introducción al estudio de la lengua.* Madrid: Aguilar.

Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics* 10. 219-231.

Shin, Ja Young, Man-Ki Lee, Un-Kyung Kim & Raimon Blancafort. 2013. *Español básico 1*. Seoul: SNU Press.

Shin, Ja Young, Man-Ki Lee, Un-Kyung Kim & Raimon Blancafort. 2013. *Español básico 2*. Seoul: SNU Press.

Smith, Michael B. 1993. The conceptual import of 'link' in grammar. Paper presented at the Twenty-eight annual Mid-America Linguistics Conference, University of Colorado.

Soto, Guillermo & Daniel Muñoz. 2000. Construcciones medias de alta transitividad en español: Un enfoque cognitivo-discursivo. *Lenguas Modernas* 26-27. 185-208.

Strauss, Susan. 2003. Completive Aspect, Emotion, and the Dynamic Eventive: The case of Korean *V-a/e pelita,* Japanese *V-te shimau,* and Spanish *se*. *Linguistics* 41(4). 653-679.

Suh, Young-Mee. 2013. A Study of Korean College Students Perception on English Middle Verb Construction. *The New Korean Association of English Language and Literature* 55(3). 171-193.

Talmy, Leonard. 1985. Force dynamics in language and thought. *Papers from the Regional Meeting of the Chicago Linguistics Society* 21. 293-337.

Torrego, Esther. 1995. From argumental to non-argumental pronouns: Spanish doubled reflexives. *Probus* 7. 221-241.

Tremblay, Annie. 2005. The L2 Acquisition of Spanish Passive and Impersonal se by French- and English-Speaking Adults. Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference, ed. by L. Dekydtspotter, R. A. Sprouse & A. Liljestrand, 251-268. Somerville, MA: Cascadilla Proceedings Project.

VanPatten, Bill. 2004. *Processing instruction: Theory, research and commentary*. Mahwah, NJ:Lawrence Erlbaum.

Vendler, Zeno. 1967. *Linguistics in philosophy.* Ithaca, NY: Cornell University Press.

Whitley, M. Stanley. 1998. Psych Verbs: Transitivity Adrift. *Hispanic Linguistics* 10. 115-153.

Whitley, M. Stanley. 2002. *Spanish/English Contrasts: A Course in Spanish Linguistics, Second edition*. Washington, D.C.: Georgetown University Press.

Yap, Foong Ha & Mikyung Ahn. 2019. Development of grammatical voice marking in Korean: On the causative, middle and passive uses of suffix *-i. Lingua* 219. 1-23.

APPENDIX A

Language background questionnaire

1. Age (in years): _____

2. Country of origin: _____

3. Education (Degree obtained or current school level attending):

    (a) Some high school
    (b) High school
    (c) Some college
    (d) College graduate
    (e) Graduate School

4. Gender:

    (a) Male
    (b) Female
    (c) Prefer not to answer

5. What is your first language? Please select all that apply:

    (a) English
    (b) Korean
    (c) Spanish
    (d) Other(s)

6. At what age did you first begin to learn Spanish? Please answer in years: _____

7. If you speak other language(s) apart from your first language and Spanish, which are they?

    _____

8. Have you ever lived in any Spanish speaking countries?

(a) Yes

(b) No

9. If you have lived in Spanish speaking countries, where and how long?

Where? _____

How long? (in months) _____

10. How did you learn Spanish up to this point?

(a) Mainly through formal classroom instruction

(b) Mainly through interacting with people

(c) A mixture of both

11. Have you ever attended any Spanish language courses? Please select all that apply.

(a) Beginner level courses

(b) Intermediate level courses

(c) Advanced level courses

(Questions below were only asked to the L1 Korean speakers)

12. Which language did you learn first between English and Spanish?

(a) English

(b) Spanish

(c) N/A

13. Have you ever taken any English proficiency tests? What was your score?

TOEFL	0-31	32-34	35-59	60-93	94-109	110+	N/A
TOEIC	0-120	121-220	221-545	546-780	781-940	941+	N/A
TEPS	10-200	201-400	401-600	601-800	801-900	901+	N/A
New TEPS	0-110	111-211	212-326	327-452	453-525	526+	N/A

APPENDIX B

Spanish proficiency test

Each of the following sentences contains a blank indicating that a word or phrase has been omitted. Select the choice that best completes the sentence.

1. Al oír del accidente de su buen amigo, Paco se puso _____.

    (a) alegre
    (b) fatigado
    (c) hambriento
    (d) desconsolado

2. No puedo comprarlo porque me _____.

    (a) falta
    (b) dan
    (c) presta
    (d) regalan

3. Tuvo que guardar cama por estar _____.

    (a) enfermo
    (b) vestido
    (c) ocupado
    (d) parado

4. Aquí está tu café, Juanito. No te quemas, que está muy _____.

    (a) dulce
    (b) amargo

(c) agrio

(d) caliente

5. Al romper los anteojos, Juan se asustó porque no podía _____ sin ellos.

(a) discutir

(b) oír

(c) ver

(d) entender

6. ¡Pobrecita! Está resfriada y no puede _____.

(a) salir de casa

(b) recibir cartas

(c) respirar con pena

(d) leer las noticias

7. Era una noche oscura sin _____.

(a) estrellas

(b) camas

(c) lágrimas

(d) nubes

8. Cuando don Carlos salió de su casa, saludó a un amigo suyo: – Buenos días, _____.

(a) ¿Qué va?

(b) ¿Cómo es?

(c) ¿Quién es?

(d) ¿Qué tal?

9. ¡Qué ruido había con los gritos de los niños y el _____ de los perros!

(a) olor

(b) sueño

(c) hambre

(d) ladrar

10. Para saber la hora, don Juan miró el _____.

(a) calendario

(b) bolsillo

(c) estante

(d) despertador

11. Yo, que comprendo poco de mecánica, sé que el auto no puede funcionar sin _____.

(a) permiso

(b) comer

(c) aceite

(d) bocina

12. Nos dijo mamá que era hora de comer y por eso _____.

(a) fuimos a nadar

(b) tomamos asiento

(c) comenzamos a fumar

(d) nos acostamos pronto

13. ¡Cuidado con ese cuchillo o vas a _____ el dedo!

(a) cortarte

(b) torcerte

(c) comerte

(d) quemarte

14. Tuvo tanto miedo de caerse que se negó a _____ con nosotros.

(a) almorzar

(b) charlar

(c) cantar

(d) patinar

15. Abrió la ventana y miró: en efecto, grandes lenguas de _____ salían llameando de las casas.

(a) zorros

(b) serpientes

(c) cuero

(d) fuego

16. Compró ejemplares de todos los diarios pero en vano. No halló _____.

(a) los diez centavos

(b) el periódico perdido

(c) la noticia que deseaba

(d) los ejemplos

17. Por varias semanas acudieron colegas del difunto profesor a _____ el dolor de la viuda.

(a) aliviar

(b) dulcificar

(c) embromar

(d) estorbar

18. Sus amigos pudieron haberlo salvado pero lo dejaron _____.

(a) ganar

(b) parecer

(c) perecer

(d) acabar

19. Al salir de la misa me sentía tan caritativo que no pude menos que _____ a un pobre mendigo que había allí sentado.

(a) pegarle

(b) darle una limosna

(c) echar una mirada

(d) maldecir

20. Al lado de la Plaza de Armas había dos limosneros pidiendo _____.

(a) pedazos

(b) paz

(c) monedas

(d) escopetas

21. Siempre maltratado por los niños, el perro no podía acostumbrarse a ___ de sus nuevos amos.

(a) las caricias

(b) los engaños

(c) las locuras

(d) los golpes

22. ¿Dónde estará mi cartera? La dejé aquí mismo hace poco y parece que el necio de mi hermano ha vuelto a _____.

(a) dejármela

(b) deshacérmela

(c) escondérmela

(d) acabármela

23. Permaneció un gran rato abstraído, los ojos clavados en el fogón y el pensamiento _____.

(a) en el bolsillo

(b) en el fuego

(c) lleno de alboroto

(d) Dios sabe dónde

24. En vez de dirigir el tráfico estabas charlando, así que tú mismo _____ del choque.

(a) sabes la gravedad

(b) eres testigo

(c) tuviste la culpa

(d) conociste a las víctimas

25. Posee esta tierra un clima tan propio para la agricultura como para _____.

(a) la construcción de trampas

(b) el fomento de motines

(c) el costo de vida

(d) la cría de reses

26. Aficionado leal de obras teatrales, Juan se entristeció al saber _____ del gran actor.

(a) del fallecimiento

(b) del éxito

(c) de la buena suerte

(d) de la alabanza

27. Se reunieron a menudo para efectuar un tratado pero no pudieron _____.

(a) desavenirse

(b) echarlo a un lado

(c) rechazarlo

(d) llevarlo a cabo

28. Se negaron a embarcarse porque tenían miedo de _____.

(a) los peces

(b) los naufragios

(c) los faros

(d) las playas

29. La mujer no aprobó el cambio de domicilio pues no le gustaba _____.

(a) el callejeo

(b) el puente

(c) esa estación

(d) aquel barrio

30. Era el único que tenía algo que comer pero se negó a _____.

(a) hojearlo

(b) ponérselo

(c) conservarlo

(d) repartirlo

APPENDIX C

Task 1

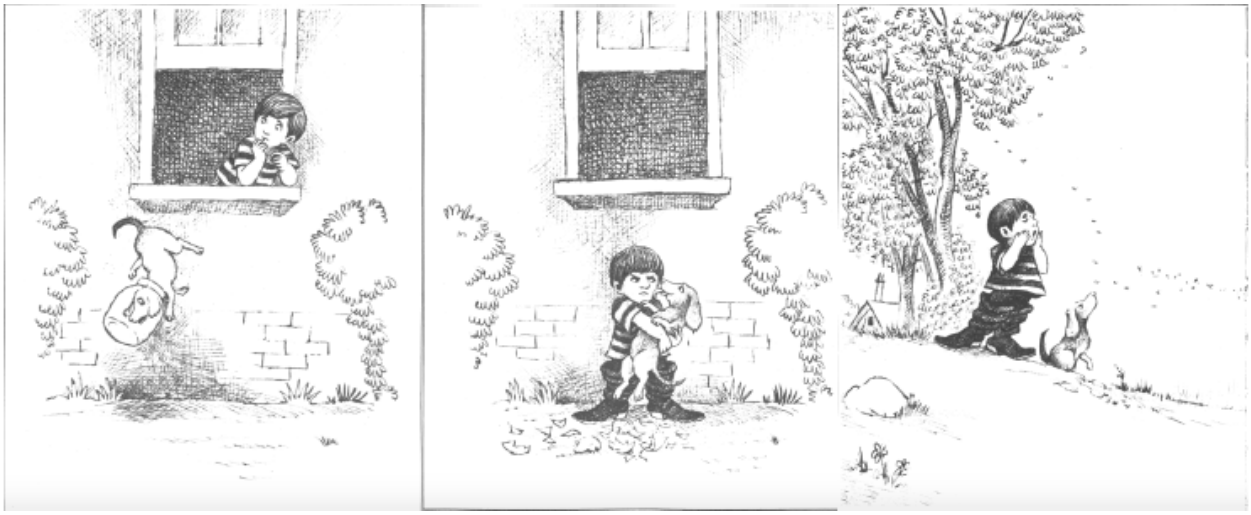
Below is a picture story book with a boy, a dog, and a frog. Look at the pictures carefully and describe what happened in this story ***in Spanish***.

Please describe **the first three scenes** of the story **in Spanish**:

_____

_____

_____



Please describe **the second three scenes** of the story **in Spanish**:

_____

_____

_____

Please describe **the last four scenes** of the story **in Spanish**:

_____

_____

_____

APPENDIX D

Task 2

Read each text carefully. Then decide which of the two options (A or B) sounds most natural in Spanish to complete the text or describe the situation in the text. You may decide that both options sound natural or neither sounds natural. Please select: Option A is more natural; Option B is more natural; Both A and B are natural; Neither A nor B is natural; I don't know.

1. After school, Laura went upstairs to her room to read. When it was time for dinner, her mom called her to come help set the table. Taking her book and still reading,

   A: Laura bajó por las escaleras y entró al comedor lentamente.

   B: Laura se bajó por las escaleras y entró al comedor lentamente.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

2. Leo and his little brother were playing on the seesaw. When they got tired, his brother slid carefully off the seesaw. But not Leo who does everything in a hurry.

   A: Abruptamente, Leo bajó del sube y baja de un brinco.

   B: Abruptamente, Leo se bajó del sube y baja de un brinco.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

3. Inés has been taking diving lessons and on the day of the school competition, she had a perfect dive. The school newspaper reported it like this:

A: "Inés Gómez saltó alto y cayó elegantemente y sin esfuerzo en la piscina."

B: "Inés Gómez saltó alto y se cayó elegantemente y sin esfuerzo en la piscina."

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

4. Hugo works in the library and was putting away the returned books. The last book, a really heavy one, belongs on the top shelf but because Hugo is short, he can hardly get it up there. He finally got it up there and turned to go when…

A: De repente, el libro cayó de la estantería.

B: De repente, el libro se cayó de la estantería.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

5. It was too hot for María to sleep comfortably in her bedroom, so she went to sleep in the living room where the air conditioner is. There,

A: María durmió cómodamente toda la noche.

B: María se durmió cómodamente toda la noche.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

6. Valeria slept badly one night and the next day in math class, she tried hard not to sleep in class. But while listening to her professor, she was overcome with fatigue.

A: De pronto, Valeria durmió en clase.

B: De pronto, Valeria se durmió en clase.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

7. Martín is studying film in college and never misses the weekend film festivals. When Clara invites him to a party on Friday night with her friends, Martín tells Clara:

A: "Sabes que voy al cine todos los fines de semana."

B: "Sabes que me voy al cine todos los fines de semana."

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

8. Teresa's dad has to go to work, so he gets in the car. When Teresa goes to his car to say goodbye to him, he isn't there anymore. She just missed him.

A: Su papá ya había ido al trabajo.

B: Su papá ya se había ido al trabajo.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

9. Alex loves mountain climbing and has been trying to beat his record for climbing the mountain near his house. Yesterday, the weather was nice and sunny and out he went. Later when his neighbor asked his roommate how Alex did the day before, he said:

A: "Alex subió la montaña en cuatro horas."

B: "Alex se subió la montaña en cuatro horas."

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

10. Pedro teaches kindergarten. One day he had planned a party for his kids, but it was a disaster because some of them got out of hand and made a terrible mess. Pedro chased them around and tried to catch them, when…

A: Inesperadamente, los niños traviesos subieron a la mesa de un salto.

B: Inesperadamente, los niños traviesos se subieron a la mesa de un salto.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

11. Among all of his friends, Luis is the fastest swimmer. In fact, he used to be a swimmer on his school's champion swimming team. He thinks it's because as a baby, he took swimming lessons.

A: Luis aprendió a nadar cuando era bebé.

B: Luis se aprendió a nadar cuando era bebé.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

12. Daniel has a math exam tomorrow and needs to study hard. He knows that if he can learn the multiplication tables completely, he'll get the A he needs. He studies all night long and by morning,

A: Daniel había aprendido de memoria las tablas de multiplicación.

B: Daniel se había aprendido de memoria las tablas de multiplicación.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

13. Victoria loves coffee. Every morning, the first thing she does after waking up is to make herself a cup of coffee and drink it up before doing anything else. However, this morning, she felt nauseous when she had her first sip and couldn't finish it.

A: Victoria bebió un poco del café.

B: Victoria se bebió un poco del café.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
◯	◯	◯	◯	◯

14. Pedro just turned 21 and went out to celebrate with his friends. He had never tasted tequila so his friends tried to get him to try it. He first sipped it to see if he liked it and when he discovered that he did, he grabbed the shot glass and...

A: Bebió el tequila de un trago.

B: Se bebió el tequila de un trago.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
◯	◯	◯	◯	◯

15. José has been a struggling musician for 10 years without a breakthrough. Last week, for the tenth time this year, he couldn't pay all of his bills. After thinking long and hard about his life,

A: José decidió cambiar de trabajo una semana después.

B: José se decidió cambiar de trabajo una semana después.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
◯	◯	◯	◯	◯

16. Juan worked as a manager in a local company but didn't get along with his boss. He often got job offers from other companies but never wanted to move his family. One day, he had a particularly nasty fight with his boos, and…

A: En ese instante, Juan decidió a cambiar de trabajo.

B: En ese instante, Juan se decidió a cambiar de trabajo.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

17. Paula is going to audition for the chorus of a musical play. She doesn't know which song they'll ask her to sing so she's a little worried because…

A: Paula sabe canciones de obras musicales pero no todas.

B: Paula se sabe canciones de obras musicales pero no todas.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

18. Pilar is a big fan of a famous singer, Billie Eilish. Billie just had a great hit and Pilar listens to it every day and can sing it, including its melody and its lyrics, from beginning to end without stopping.

A: Pilar sabe la nueva canción de Billie de memoria.

B: Pilar se sabe la nueva canción de Billie de memoria.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

19. One day, Amelia took her 10-year-old son Miguel to work. Miguel had to write a school essay about how people work. When he asked his mom why there are so many female employees, Amelia tells him that:

A: "Las mujeres que trabajan aquí son muy inteligentes."

B: "Las mujeres que trabajando aquí son muy inteligentes."

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

20. Alfonso recently started dating a girl who is slightly taller than he is. Although Alfonso's friends think that she is not that tall, he insists that she is very tall.

A: La novia de Alfonso no es muy alta, pero él cree que lo es.

B: La novia de Alfonso no es muy alta, pero él cree que la es.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

21. When Marisol was preparing to go out, her mom told her to be home by midnight and Marisol promised that she would. However, the time got away from her and she got home at 12:20. Her mom catches her sneaking into her room and says:

A: "¡Llegaste tarde!"

B: "¡Estuviste tarde!"

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

22. Federico doesn't want to go to school today and tells his dad that he is not feeling well, but his dad already knows that he is lying.

A: Papá dice que Federico tenga que ir a la escuela.

B: Papá dice que Federico tiene que ir a la escuela.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

23. Juan woke up early and went to the library because he has a Spanish midterm next week. He needs to get a good grade on this midterm because he didn't do very well in his last quiz. To get a good grade,

A: Lo es importante estudiar para el examen.

B: Es importante estudiar para el examen.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

24. This afternoon, when Estefanía turned on her laptop, she saw that it was not working properly. After having it checked by a techinician, she was told that she needs to get a new laptop.

A: Estefanía vaya a comprar una nueva computadora portátil.

B: Estefanía va a comprar una nueva computadora portátil.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

25. Carlos was out of town to attend a friend's wedding. When he flew back home this afternoon, he saw the grass, sidewalks and streets were wet.

A: Parece que ha llovido esta mañana.

B: Parece que llovió esta mañana.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

26. Paco was doing a group project for his journalism class. After spending a lot of time researching and preparing for the project, they got a lot of compliments from their professor and got a very good grade.

A: El proyecto salió estupendo.

B: El proyecto resultó estupendo.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

27. Alejandro is a football player in his school. While getting ready for practice, he suddenly felt dizzy. He thought that it was not a big deal, but it got worse. His coach noticed that he was not feeling well, so he asked him:

A: "Estás enfermo, ¿no?"

B: "Estás enfermo, ¿verdad?"

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

28. Jorge woke up late this morning and is late for class. He is in a hurry, but he can't figure out where his book is.

A: Jorge está buscando por su libro.

B: Jorge está buscando para su libro.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

29. Antonio is throwing a party tonight and asks some of his friends to help with the decorations. They ask Antonio how he wants them done but he's very busy preparing the food and tells them:

A: "Háganlo como ustedes quieren."

B: "Háganlo como ustedes prefieren."

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

30. This semester, I am taking five difficult classes and have tons of assignments every day. But I'm not stressed out because I always have my friends who study with me.

A: Mis amigos y yo estudiemos juntos.

B: Mis amigos y yo estudiaríamos juntos.

Option A is more natural	Option B is more natural	Both A and B are natural	Neither A nor B is natural	I don't know
○	○	○	○	○

APPENDIX E

Task 3

Now you're going to read several contexts about different people doing different things. If you think the phrase **a mí mismo/a** (or **a sí mismo/a/os***) could be used for emphasis and sounds natural, click* ***YES****. If you think it sounds unnatural, click* ***NO****. If you're not sure, click* ***I don't know****.*

Example:

Julia: Hola, ¿qué tal? Me llamo _____ Julia. ¿Y tú, cuál es tu nombre?
José: Bueno, me dicen "José", pero me llamo _____ "¡El guapo!"

Answers:

Julia: Hola, ¿qué tal? Me llamo **NO** Julia. ¿Y tú, cuál es tu nombre?
José: Bueno, me dicen "José", pero me llamo **YES** "¡El guapo!"

In the first blank, it would sound unnatural to use the phrase **a mí misma**, *so* ***NO*** *is the expected answer. In the second blank, the phrase* **a mí mismo** *could be used in this context and it would sound natural, so* ***YES*** *is the expected answer.*

1. Mi hermano mayor siempre me llamaba a las 6:00 de la mañana para ir a la escuela, pero con mi nuevo reloj despertador, yo ya me despierto _____. Cuando suena la alarma, me despierto

_____ automáticamente. Mi hermano no tiene que entrar y brincar a la cama porque ahora me levanto _____. Por supuesto, cuando hace frío, me gusta dormir un poco más pero normalmente me levanto _____ a las 6:15 o 6:20 — no muy tarde.

2. Soy Francisco y tengo seis años. Para ir a la escuela tengo que vestirme _____ rápido porque viene el autobús muy temprano. Pero cuando mamá ve que todavía tengo puesto mi pijama, quiere ayudarme. Tengo que decirle: "¡Puedo vestirme _____, mamá — tengo seis años!"

3. Cuando Susana se enfermó y estuvo en el hospital, le ayudaron con todos sus cuidados. Un día la enfermera le preguntó a qué horas quería su ducha. Susana le dijo, "Normalmente me ducho _____ por la mañana pero hoy prefiero hacerlo en la noche". La enfermera le dijo que no podía ayudarle en la noche porque tenía otros pacientes. "No hay problema. Estoy mejor y puedo ducharme _____", dijo Susana.

4. Cuando Jorge tenía quince años le empezaron a salir más pelos en la barba. Su papá le dijo, "Hijo, ya es tiempo de afeitarse _____ todos los días. Ven, te voy a ayudar". Pero su papá no sabía que Jorge ya había aprendido a hacerlo con su hermano. "Papá, yo puedo afeitarme _____. José ya me enseñó", Jorge dijo.

5. En el kínder, un día la maestra les dijo a los niños, "Vengan niños, les voy a enseñar a peinarse _____. No tengo tiempo para peinar a todos y viene el fotógrafo. Tienen que estar guapos para las fotos". Pero Pedro, quien siempre andaba despeinado, protestó y dijo, "Me gusta el *'look'* natural — ¡nunca me peino _____ con cepillo!"

6. Cuando yo era joven no sabía qué iba a hacer con mi vida. Entonces fui a París a estudiar arte y al recorrer los museos, viendo las obras de arte, me encontré _____. ¡Ese era el lugar para mí! Hace poco, hice planes de ir otra vez pero no pude ir porque me encontré _____ demasiado ocupada — ¡con mi propia exhibición de arte!

7. Soy terrible con los números y cuando tengo que explicarles las fórmulas matemáticas a mis hijos, hasta me confundo _____. No puedo pedirle ayuda a mi esposa tampoco porque cuando ella me las explica, me confundo _____ aún más.

8. Me encanta el cine y quiero ser actor algún día. Cuando veo una película, me imagino _____ como el actor principal. ¿Es una carrera difícil? Me imagino _____ que sí, pero aún quiero intentarlo.

9. Cuando Pilar se puso el nuevo vestido, salió de su cuarto y le preguntó a su hermana, "¿Cómo me veo _____ en este vestido?" "Muy guapa", le contestó su hermana. Pero Pilar no estaba segura y dijo, "No sé, ahora que me veo _____ en el espejo, no me gusta".

10. Casi siempre cuando voy a una nueva ciudad, me pierdo _____ y tengo que pedir ayuda para llegar a algún lugar. Pero a veces, me gusta caminar y caminar por las calles sin ningún destino. Así, me pierdo _____ en mis pensamientos y me siento totalmente libre.