



PhD THESIS

A MEANING-BASED ANALYSIS OF L2 TEXTS DERIVED
FROM STRUCTURAL FUNCTIONAL LINGUISTICS AND
MÉTODO DE LOS RELOJES R2

TESIS DOCTORAL

UN ANÁLISIS DE TEXTOS L2 BASADO EN EL
SIGNIFICADO Y DERIVADO DE LA LINGÜÍSTICA
ESTRUCTURAL FUNCIONAL Y EL MÉTODO DE LOS
RELOJES R2

AUTOR

TIMOTHY WILLIAM LAWRENCE

DIRECTORES

MANUEL PÉREZ SAIZ

JUAN RAFAEL ZAMORANO-MANSILLA

UNIVERSIDAD DE CANTABRIA

Escuela de **Doctorado** de la Universidad de Cantabria

Santander **2024**

PhD in Geography and History

Department of Philology

University of Cantabria



PhD Thesis

**A Meaning-Based Analysis of L2 Texts Derived from Structural
Functional Linguistics and Método de los Relojos R2**

**Un Análisis de Textos L2 Basado en el Significado y Derivado de
la Lingüística Estructural Funcional y el Método de los Relojos
R2**

Timothy William Lawrence

Submitted to the EDUC as partial fulfillment of the requirements for
Doctor of Geography and History Degree in Philology

PhD Directors

Manuel Pérez Saiz

Juan Rafael Zamorano-Mansilla

July 2024

Abstract

This study presents a large-scale corpus investigation of L2 English and Spanish texts employing meaning-based complexity as a viable alternative to the numerous and varied approaches to syntactic complexity. The research method operationalizes Structural Functional Linguistics, the *Método de los Relojes Reloj 2* and the *Common European Framework of Reference for Languages* to create a fine-grained meaning-based framework. A large data set was used to provide a balanced representative sample consisting of over 1.4 million words extracted from the second version of the *Education First-Cambridge Open Language Database*, *Corpus escrito del español L2* and *Corpus de aprendices de español*. Results indicate that subordination through hypotaxis increases through advanced levels as proficiency rises in L2 English and Spanish texts. Coordination through parataxis was observed having consistent frequencies without significant variation throughout proficiency levels. Findings show the main drivers of high logico-semantic frequency were temporal and causal-conditional indicators in hypotaxis through enhancement. Parataxis through extension demonstrated the highest frequency in the entire study. Tasked-based formulaic sequencing (Alexopoulou et al., 2015; Alhassan & Wood, 2015; Chen et al., 2021; Lewis, 1997; Wray, 2002) appeared to be a small factor in influencing frequency at lower proficiency levels while developmental formulaic sequencing (Boers et al., 2006; Boers & Lindstromberg, 2012; Dai & Ding, 2010; Schmitt, 2008; Stengers et al., 2011) appeared to play a minor role at higher proficiency levels. Findings lead to the introduction of the developmental sequence of meaning-based interlanguage complexification hypothesis which proposes L2 acquisition occurs at advanced proficiency levels in a large grouping of adverbial subordinators which materialize at low frequencies. As frequency increases, the quantity of subordinators decreases substantially with there being a small number of comparable meaning-based units that suggests L1 to L2 transfer.

Resumen

Este estudio presenta una investigación de corpus a gran escala de textos L2 en inglés y español que emplea la complejidad basada en el significado como alternativa viable a los numerosos y variados enfoques de la complejidad sintáctica. El método de investigación operacionaliza la Lingüística Funcional Estructural, el *Método de los Relojes* Reloj 2 y el *Marco Común Europeo de Referencia para las Lenguas* para crear un marco basado en el significado de grano fino. Se incorporaron macrodatos para proporcionar una muestra representativa equilibrada consistente en 1,4 millones de palabras extraídas de la segunda versión de la *Education First-Cambridge Open Language Database*, *Corpus Escrito del español L2* y *Corpus de aprendices de español*. Los resultados indican que la subordinación a través de la hipotaxis aumenta en los niveles avanzados a medida que aumenta el dominio de los textos en inglés y español L2. La coordinación a través de la parataxis se observó en frecuencias consistentes sin variaciones significativas a lo largo de los niveles de competencia. Los resultados muestran que los principales impulsores de la alta frecuencia lógico-semántica fueron los indicadores temporales y causales-condicionales en la hipotaxis a través de la ampliación. La parataxis por extensión mostró la frecuencia más alta de todo el estudio. La secuenciación formulaica basada en tareas (Alexopoulou et al., 2015; Alhassan & Wood, 2015; Chen et al., 2021; Lewis, 1997; Wray, 2002) parecía ser un factor menor que influía en la frecuencia de uso en los niveles de competencia más bajos, mientras que la secuencia formulaica del desarrollo (Boers et al., 2006; Boers & Lindstromberg, 2012; Dai & Ding, 2010; Schmitt, 2008; Stengers et al., 2011) parecía desempeñar un papel menor en los niveles de competencia más altos. Los hallazgos conducen a la introducción de la hipótesis de la secuencia de desarrollo de la complejificación de la interlengua basada en el significado, que propone que la adquisición de la L2 se produce en niveles de competencia avanzados en una gran agrupación de subordinadores adverbiales que se materializan a bajas frecuencias. A medida que aumenta la frecuencia, la cantidad de subordinadores disminuye sustancialmente, existiendo un pequeño número de unidades comparables basadas en el significado, lo que sugiere una transferencia de L1 a L2.

Acknowledgements

I have endless gratitude for my supervisors, Manuel Pérez Saiz and Juan Rafael Zamorano-Mansilla, for their support, guidance and patience throughout this project. I still remember the first lecture with Manolo which inspired me to start this line of research. I am very thankful that Juan-Rafael agreed to take me under his wing after only an email and an online meeting. It shows an incredible dedication to furthering knowledge in higher education.

I would also like to thank members of the Department of Philology at the University of Cantabria. In no certain order, Jesús Ángel González López, Evelyn Gandón-Chapela, Franciso Gallardo del Puerto, Maria del Carmen Camus Camus and the many others, please forgive me if I have not mentioned your name, who have helped me with valuable insight or encouraged me throughout the process. I greatly appreciate the helpful information from the many researchers from numerous universities that I have written to for advice.

Massive thanks to my parents and my family on both sides of the pond for their support, patience, understanding and proofreading. Without you, this voyage would have been impossible.

Last but far from least, a special acknowledgement is reserved for Julia Zeilfelder, Nicolas Fernández Zeilfelder and the cats for keeping me in line, putting up with me and providing endless support. I could not have finished this PhD without you guys.

Table of Contents

Abstract	iii
Resumen	iv
Acknowledgements	v
Table of Contents.....	vi
List of Abbreviations	x
List of Figures	xi
List of Tables	xii
1. Introduction.....	1
1.1 Object of Study.....	1
1.2 The Present Study	5
1.3 Thesis Structure.....	8
2. Literature Review.....	11
2.1 Structural Functional Grammar: A Clause-Based Standpoint	12
2.1.1 Taxis	13
2.1.2 Expansion: Elaborating, Extending and Enhancing	16
2.1.2.1 Elaboration.....	16
2.1.2.2 Extension	18
2.1.2.3 Enhancement	23
2.2 Complexity Measures	31
2.2.1 Syntactic Complexity	33
2.2.2 Syntactic Complexity Viewed Through the CEFR	40
2.2.2.1 What is the CEFR?	40
2.2.2.2 Syntactic Complexity Studies Factoring in EFL Writing Across CEFR Levels	42
2.2.3 Applying Structural Functional Grammar to Syntactic Complexity	50
2.3 Acquisition of L2 Subordinate Clauses: Overview of L1 Spanish and English.....	64
2.3.1 Spanish L1 Acquisition of English L2 Subordinate Clauses	64
2.3.2 English L1 Acquisition of Spanish L2 Subordinate Clauses	66
2.3.3 Interim Summary	67
2.4 Delimiting the Método de los Relojes.....	69
2.4.1 The Four Verb Types.....	70
2.4.2 Versions 1:-5:	70
2.4.3 The Five Basic Verb Forms	71
2.4.4 Reloj 2	73
2.4.4.1 One o'clock	74
2.4.4.2 Two o'clock	75
2.4.4.3 Three o'clock	76
2.4.4.4 Four o'clock	77
2.4.4.5 Five o'clock	79
2.4.4.6 Six o'clock	80
2.4.4.7 Seven o'clock	81
2.4.4.8 Eight o'clock.....	82

2.4.4.9 Nine o'clock	82
2.4.4.10 Ten o'clock	83
2.5 Summary.....	83
3. Operationalizing Parataxis and Hypotaxis Through MR R2 to Form a Meaning-Based Framework.....	87
3.1 Introduction	87
3.2 Methodology	88
3.3 Results and Discussion.....	91
3.3.1 One o'clock	91
3.3.2 Two o'clock	96
3.3.3 Three o'clock	99
3.3.4 Four o'clock	104
3.3.4.1 CI	105
3.3.5 Five o'clock	107
3.3.6 Six o'clock	109
3.3.7 Seven o'clock	114
3.3.8 Eight o'clock.....	117
3.3.9 Nine o'clock	119
3.3.10 Ten o'clock.....	122
3.4 The Meaning-Based Framework.....	123
3.4.1 Extension	124
3.4.2 Enhancement.....	125
3.4.3 Elaboration	130
4. Corpus Study: Application of the Meaning-Based Framework	133
4.1 Introduction	133
4.2 Methodology	135
4.2.1 Corpora	135
4.2.1.1 EFCAMDAT2	136
4.2.1.2 CEDEL2	138
4.2.1.3 CAES	142
4.2.2 Data Summary	144
4.2.3 Technical Details and Evaluation	147
4.3 Experiment 1: Level 1 Taxis.....	154
4.3.1 Results	154
4.3.1.1 Hypotaxis	154
4.3.1.2 Parataxis.....	155
4.3.2 Discussion	156
4.3.2.1 Hypotaxis	156
4.3.2.2 Parataxis.....	158
4.3.3 Conclusions	160
4.4 Experiment 2: Level 2 Expansion	162
4.4.1 Results	162
4.4.1.1 Hypotaxis and Extension.....	162
4.4.1.2 Hypotaxis and Enhancement.....	163
4.4.1.3 Hypotaxis and Elaboration.....	164
4.4.1.4 Parataxis and Extension.....	165
4.4.1.5 Parataxis and Enhancement	166
4.4.1.6 Parataxis and Elaboration	167
4.4.2 Discussion	167
4.4.2.1 Hypotaxis and Extension.....	167

4.4.2.2 Hypotaxis and Enhancement	169
4.4.2.3 Hypotaxis and Elaboration	171
4.4.2.4 Parataxis and Extension	172
4.4.2.5 Parataxis and Enhancement	174
4.4.2.6 Parataxis and Elaboration	175
4.4.3 Conclusions	176
4.5 Experiment 3: Level 3 Meaning-Based Categories	180
4.5.1 Results	180
4.5.1.1 Hypotaxis and Extension Through Addition	182
4.5.1.2 Hypotaxis and Extension Through Variation	182
4.5.1.3 Hypotaxis and Extension Through Alternation	183
4.5.1.4 Hypotaxis and Enhancement Through Temporal Markers	184
4.5.1.5 Hypotaxis and Enhancement Through Spatial Markers	184
4.5.1.6 Hypotaxis and Enhancement Through Manner Markers	185
4.5.1.7 Hypotaxis and Enhancement Through Causal-Conditional Markers	186
4.5.1.8 Parataxis and Extension Through Addition	187
4.5.1.9 Parataxis and Extension Through Variation	188
4.5.1.10 Parataxis and Extension Through Alternation	189
4.5.1.11 Parataxis and Enhancement Through Temporal Markers	190
4.5.1.12 Parataxis and Enhancement Through Spatial Markers	191
4.5.1.13 Parataxis and Enhancement Through Manner Markers	191
4.5.1.14 Parataxis and Enhancement Through Causal-Conditional Markers	192
4.5.1.15 Parataxis and Elaboration Through Exposition	193
4.5.1.16 Parataxis and Elaboration Through Exemplification	193
4.5.1.17 Parataxis and Elaboration Through Clarification	194
4.5.2 Discussion	195
4.5.2.1 Hypotaxis and Extension Through Addition	195
4.5.2.2 Hypotaxis and Extension Through Variation	195
4.5.2.3 Hypotaxis and Extension Through Alternation	195
4.5.2.4 Hypotaxis and Enhancement Through Temporal Markers	196
4.5.2.5 Hypotaxis and Enhancement Through Spatial Markers	198
4.5.2.6 Hypotaxis and Enhancement Through Manner Markers	199
4.5.2.7 Hypotaxis and Enhancement Through Causal-Conditional Markers	200
4.5.2.8 Parataxis and Extension Through Addition	201
4.5.2.9 Parataxis and Extension Through Variation	203
4.5.2.10 Parataxis and Extension Through Alternation	203
4.5.2.11 Parataxis and Enhancement Through Temporal Markers	204
4.5.2.12 Parataxis and Enhancement Through Spatial Markers	206
4.5.2.13 Parataxis and Enhancement Through Manner Markers	207
4.5.2.14 Parataxis and Enhancement Through Causal-Conditional Markers	208
4.5.2.15 Parataxis and Elaboration Through Exposition	209
4.5.2.16 Parataxis and Elaboration Through Exemplification	210
4.5.2.17 Parataxis and Elaboration Through Clarification	211
4.5.3 Conclusions	212
4.6 Experiment 4: Level 4 Meaning-Based Units	218
4.6.1 Results	219
4.6.1.1 Hypotaxis and Extension	219
4.6.1.2 Hypotaxis and Enhancement	222
4.6.1.3 Hypotaxis and Elaboration	233
4.6.1.4 Parataxis and Extension	235
4.6.1.5 Parataxis and Enhancement	238
4.6.1.6 Parataxis and Elaboration	242
4.6.2 Discussion	246
4.6.2.1 Hypotaxis and Extension	246
4.6.2.2 Hypotaxis and Enhancement	251

4.6.2.3 Hypotaxis and Elaboration.....	272
4.6.2.4 Parataxis and Extension.....	275
4.6.2.5 Parataxis and Enhancement	280
4.6.2.6 Parataxis and Elaboration.....	284
4.6.3 Conclusion	288
4.6.3.1 The Developmental Sequence of Meaning-Based Interlanguage Complexification Hypothesis for Adverbial Subordination	295
5. Conclusions.....	300
References.....	304

List of Abbreviations

CAES	<i>Corpus de aprendices de español</i>
CAF	Complexity, Accuracy and Fluency
CEDEL2	<i>Corpus escrito del español L2</i>
CEFR	Common European Framework of Reference for Languages
DSSICH	Developmental Sequence of Syntactic Interlanguage Complexification Hypothesis
DSMBICH	Developmental Sequence of Meaning-based Interlanguage Complexification Hypothesis
EFCAMDAT2	<i>Education First-Cambridge Open Language Database Version 2</i>
EFL	English as a foreign language
HDLF	High Density Low Frequency MBU Category
L1	First Language
L2	Second Language
LDMF	Low Density Medium Frequency MBU Category
LDHF	Low Density High Frequency MBU Category
MBC	Meaning-based Category
MBF	Meaning-based Framework
MBU	Meaning-based Unit
MR	<i>Método de los Relojes</i> (2018)
Reloj 2	R2
SFG	Structural Functional Grammar
SFL	Structural Functional Linguistics
SLA	Second Language Acquisition
V1	First Verb
V2	Second Verb
α	Primary or dominant clause in hypotaxis
β	Secondary or dependent in hypotaxis
<<word>>	Semiotic reference to <u>word</u> meaning
<< EC >>	Semiotic reference to enclosed <u>clause</u> in a sentence
[[EC]]	Embedded clause
#	Grammatical pause (comma)
^	Tonal accent

List of Figures

FIGURE 1 CEFR LEVELS (FIGUERAS, 2012, P. 478).	41
FIGURE 2 LINGUISTIC BACKGROUND AND TASK VARIABLES (CEDEL2, 2023).	139
FIGURE 3 EXAMPLE OF CEDEL2 WEB-BASED INTERFACE (LONZANO, 2022).	140
FIGURE 4 OBJECTIVE MEASUREMENTS PROFICIENCY LEVELS (CEDEL2, 2023).	141
FIGURE 5 EXAMPLE CAES CONCORDANCE LINES WITH KWIC (CAES, 2023).	143
FIGURE 6 EXAMPLE SEARCH CRITERIA FOR PORQUE IN CAES SEARCH FUNCTION.	148
FIGURE 7 EXAMPLE CONCORDANCE LINES AND KWIC FOR PORQUE IN THE CAES.	149
FIGURE 8 EXAMPLE ANALYSIS OF BECAUSE KWIC AND CONCORDANCE LINE (HALLIDAY, 2014).	150
FIGURE 9 EXAMPLE ANALYSIS OF PARATAXIS AND KWIC AND CONCORDANCE LINE (HALLIDAY, 2014).	151
FIGURE 10 EXAMPLE ANALYSIS OF HYPOTACTIC ELABORATION AND KWIC AND CONCORDANCE LINE (HALLIDAY, 2014).	152
FIGURE 11 EXAMPLE MBF WITH CEFR PROFICIENCY LEVELS.	153
FIGURE 12 LEVELS OF ANALYSIS IN THE MBF.	153
FIGURE 13 LEVEL 1 HYPOTAXIS RESULT SUMMARY.	155
FIGURE 14 LEVEL 1 PARATAXIS RESULT SUMMARY.	156
FIGURE 15 LEVEL 2 HYPOTAXIS AND EXTENSION RESULTS SUMMARY.	163
FIGURE 16 LEVEL 2 HYPOTAXIS AND ENHANCEMENT RESULTS SUMMARY.	164
FIGURE 17 LEVEL 2 HYPOTAXIS AND ELABORATION RESULTS SUMMARY.	165
FIGURE 18 LEVEL 2 PARATAXIS AND EXTENSION RESULTS SUMMARY.	166
FIGURE 19 LEVEL 2 PARATAXIS AND ENHANCEMENT RESULTS SUMMARY.	166
FIGURE 20 LEVEL 2 PARATAXIS AND ELABORATION RESULTS SUMMARY.	167
FIGURE 21 LEVEL 3 SUMMARY HYPOTAXIS AND EXTENSION THROUGH ADDITION.	182
FIGURE 22 LEVEL 3 SUMMARY HYPOTAXIS EXTENSION THROUGH VARIATION.	183
FIGURE 23 LEVEL 3 HYPOTAXIS EXTENSION THROUGH ALTERNATION.	183
FIGURE 24 LEVEL 3 HYPOTAXIS AND TEMPORAL ENHANCEMENT SUMMARY.	184
FIGURE 25 LEVEL 3 HYPOTAXIS SPATIAL ENHANCEMENT SUMMARY.	185
FIGURE 26 LEVEL 3 HYPOTAXIS AND MANNER ENHANCEMENT SUMMARY.	186
FIGURE 27 LEVEL 3 HYPOTAXIS AND CAUSAL-CONDITIONAL ENHANCEMENT SUMMARY.	187
FIGURE 28 LEVEL 3 ADDITION THROUGH PARATAXIS AND EXTENSION SUMMARY.	188
FIGURE 29 VARIATION THROUGH PARATAXIS AND EXTENSION SUMMARY.	189
FIGURE 30 LEVEL 3 ALTERNATION THROUGH PARATAXIS AND EXTENSION SUMMARY.	190
FIGURE 31 LEVEL 3 TEMPORAL ENHANCEMENT IN PARATAXIS SUMMARY.	190
FIGURE 32 LEVEL 3 SPATIAL ENHANCEMENT IN PARATAXIS SUMMARY.	191
FIGURE 33 LEVEL 3 MANNER ENHANCEMENT THROUGH PARATAXIS.	192
FIGURE 34 LEVEL 3 CAUSAL-CONDITIONAL ENHANCEMENT THROUGH PARATAXIS SUMMARY.	192
FIGURE 35 LEVEL 3 EXPOSITION THROUGH PARATACTIC ELABORATION RESULTS SUMMARY.	193
FIGURE 36 LEVEL 3 EXEMPLIFICATION THROUGH ELABORATION IN PARATAXIS RESULTS SUMMARY.	194
FIGURE 37 LEVEL 3 CLARIFICATION AND ELABORATION THROUGH PARATAXIS RESULT SUMMARY.	194

List of Tables

TABLE 1 SUMMARY OF EXPANSION CLAUSE COMPLEXES (HALLIDAY, 2014, P. 447).	14
TABLE 2 FINITE AND NON-FINITE DEPENDENT CLAUSES (HALLIDAY, 2014, P. 454).	15
TABLE 3 CATEGORIES OF EXTENSION AND HYPOTACTIC NON-FINITE MARKERS (HALLIDAY, 2014, P. 471).	19
TABLE 4 ENHANCEMENT THROUGH PARATAXIS AND HYPOTAXIS (HALLIDAY, 2014, PP. 477-478).	25
TABLE 5 PARATACTIC ENHANCEMENT CATEGORIES (HALLIDAY, 2014, PP. 478-480).	26
TABLE 6 FINITE HYPOTACTIC ENHANCING CLAUSES (HALLIDAY, 2014, PP. 481-485).	27
TABLE 7 NON-FINITE HYPOTACTIC ENHANCING CLAUSES (HALLIDAY, 2014, PP. 484-486).	29
TABLE 8 CONJUNCTIVE MARKERS USED FOR MORE THAN ONE TYPE OF EXPANSION (HALLIDAY, 2014, P. 487).	31
TABLE 9 BULTÉ AND HOUSEN (2012) AND NORRIS AND ORTEGA (2009) SUMMARY SYNTACTIC COMPLEXITY MEASURES.	35
TABLE 10 SYNTACTIC COMPLEXITY MEASURES LU AND AI (2015).	37
TABLE 11 SAMPLE FINE-GRAINED SYNTACTIC COMPLEXITY INDICES ZHANG AND LU (2022).	39
TABLE 12 EFCAMDAT2 CLAUSE TYPE PER CEFR PROFICIENCY LEVEL (CHEN ET AL., 2021, P. 812).	45
TABLE 13 PREVIOUS SYNTACTIC COMPLEXITY STUDIES IN EFL WRITING ACROSS CEFR LEVELS (KHUSHIK & HUHTA, 2020).	48
TABLE 14 STUDIES USING SFL INDICES TO RESEARCH SYNTACTIC COMPLEXITY.	55
TABLE 15 SULISTYANINGRUM AND RASYID (2015) NUMBER OF KINDS OF TAXIS GROUP 1 AND GROUP 2.	55
TABLE 16 ZARCO-TEJADA ET AL. (2016) EXTENDING CONJUNCTIONS SUMMARY.	57
TABLE 17 ZARCO-TEJADA ET AL. (2016) ENHANCING CONJUNCTIONS SUMMARY.	58
TABLE 18 ZARCO-TEJADA ET AL. (2016) ELABORATING CONJUNCTIONS SUMMARY.	59
TABLE 19 YANG ET AL. (2017) FREQUENCY OF CONJUNCTION APPLICATION BY PARTICIPANTS.	60
TABLE 20 WENHUI XUAN (2019) LOGICO-SEMANTIC RELATIONS IN STUDENT WRITING.	60
TABLE 21 WENHUI XUAN (2019) ENHANCEMENT MARKERS IN L1 CHINESE WRITING IN L2 ENGLISH.	61
TABLE 22 RASOOL & MAHMOOD (2023) OCCURRENCES AND PERCENTAGES OF EXPANSION RELATIONS.	61
TABLE 23 RASOOL & MAHMOOD (2023) OCCURRENCES AND PERCENTAGES OF EXTENSION RELATIONS.	62
TABLE 24 RASOOL & MAHMOOD (2023) OCCURRENCES AND PERCENTAGES OF ELABORATION RELATIONS.	62
TABLE 25 RASOOL & MAHMOOD (2023) OCCURRENCES AND PERCENTAGES OF ENHANCEMENT RELATIONS.	63
TABLE 26 SUBORDINATORS AT R2 ONE O’CLOCK (PÉREZ, 2018, P. 124).	75
TABLE 27 SUBORDINATORS AT R2 TWO O’CLOCK AND MEANING (PÉREZ, 2018, P. 128).SE	76
TABLE 28 PARTICLES IN R2 AT THREE O’CLOCK (PÉREZ, 2018, P. 129).	77
TABLE 29 R2 FOUR O’CLOCK SUBORDINATORS, MEANING AND SUBSTITUTION COMBINATIONS (PÉREZ, 2018, P. 132).	78
TABLE 30 R2 FOUR O’CLOCK CI EXAMPLES (PÉREZ, 2018, PP. 197-198).	79
TABLE 31 R2 AT FIVE O’CLOCK SUBORDINATOR AND MEANING (PÉREZ, 2018, P. 137).	80
TABLE 32 R2 AT SIX O’CLOCK SUBORDINATORS AND MEANINGS (PÉREZ, 2018, P. 139).	81
TABLE 33 SEVEN O’CLOCK SUBORDINATORS AND MEANINGS (PÉREZ, 2018, P. 143).	81
TABLE 34 R2 AT EIGHT O’CLOCK SUBORDINATORS AND MEANINGS (PÉREZ, 2018, P. 149).	82
TABLE 35 R2 AT NINE O’CLOCK SUBORDINATORS AND MEANINGS (PÉREZ, 2018, P. 149)	82
TABLE 36 TEN O’CLOCK SUBORDINATORS AND MEANINGS (PÉREZ, 2018, P. 158).	83
TABLE 37 EXAMPLE OF TRANSLATION METHODOLOGY.	89
TABLE 38 UNITS OF ANALYSIS WITHIN MEANING-BASED FRAMEWORK (HALLIDAY, 2014; PÉREZ, 2018).	90
TABLE 39 ENGLISH R2 ONE O’CLOCK <<CASUAL-CONDITIONAL: CONDITION: CAUSE: PURPOSE: FINITE>>.	95
TABLE 40 ENGLISH R2 ONE O’CLOCK <<TEMPORAL: DIFFERENT TIME: EARLIER>> STRUCTURES.	96
TABLE 41 ENGLISH R2 TWO O’CLOCK <<CAUSAL-CONDITIONAL: CONDITION: CAUSE: REASON>> STRUCTURES.	97
TABLE 42 ENGLISH R2 TWO O’CLOCK <<CAUSAL-CONDITIONAL: CONDITION: CAUSE: RESULT>> STRUCTURES.	99
TABLE 43 ENGLISH R2 THREE O’CLOCK <<TEMPORAL: SAME TIME: [POINT] & [SPREAD]>> STRUCTURES.	100
TABLE 44 ENGLISH R2 THREE O’CLOCK <<TEMPORAL: SAME TIME [POINT]>> STRUCTURES.	101
TABLE 45 ENGLISH R2 THREE O’CLOCK <<TEMPORAL; DIFFERENT TIME LATER>> STRUCTURES.	102

TABLE 46 ENGLISH R2 THREE O’CLOCK <<TEMPORAL: DIFFERENT TIME EARLIER>> STRUCTURES.....	103
TABLE 47 ENGLISH R2 THREE O’CLOCK <<MANNER: COMPARISON>> STRUCTURES.....	103
TABLE 48 ENGLISH R2 THREE O’CLOCK <<SPATIAL POINT >> STRUCTURES.....	104
TABLE 49 ENGLISH R2 FOUR O’CLOCK HYPOTAXIS AND ELABORATION <<NON-DEFINING RELATIVE CLAUSE>>.....	107
TABLE 50 ENGLISH R2 FIVE O’CLOCK <<CASUAL-CONDITIONAL: CONDITION: POSITIVE >> STRUCTURES.....	109
TABLE 51 ENGLISH R2 SIX O’CLOCK <<TEMPORAL: DIFFERENT TIME: LATER >> STRUCTURES.....	110
TABLE 52 ENGLISH R2 SIX O’CLOCK <<MANNER: COMPARISON>> STRUCTURES.....	113
TABLE 53 ENGLISH R2 SIX O’CLOCK <<CAUSAL-CONDITIONAL: CONDITION: CONCESSIVE >> STRUCTURES.....	114
TABLE 54 ENGLISH R2 SIX O’CLOCK <<ADDITION: POSITIVE>> STRUCTURES.....	114
TABLE 55 ENGLISH R2 SEVEN O’CLOCK <<CAUSAL-CONDITIONAL: ALTERNATIVE CONDITION: CONCESSIVE>> STRUCTURES.....	115
TABLE 56 ENGLISH R2 SEVEN O’CLOCK PARATAXIS <<EXTENSION: CLARIFICATION>> STRUCTURES.....	117
TABLE 57 ENGLISH R2 EIGHT O’CLOCK <<CAUSAL CONDITIONAL: CONDITION: CAUSE: REASON>> STRUCTURES.....	118
TABLE 58 ENGLISH R2 EIGHT O’CLOCK <<MANNER: COMPARISON>> STRUCTURES.....	118
TABLE 59 ENGLISH R2 NINE O’CLOCK <<CASUAL-CONDITIONAL: CONDITION: POSITIVE: FINITE>> STRUCTURES.....	120
TABLE 60 ENGLISH R2 FIVE O’CLOCK <<CASUAL-CONDITIONAL: CONDITION: NEGATIVE >> STRUCTURES.....	120
TABLE 61 ENGLISH R2 NINE O’CLOCK <<CASUAL-CONDITIONAL: CAUSE: RESULT>> STRUCTURES.....	121
TABLE 62 ENGLISH R2 NINE O’CLOCK <<MANNER: COMPARISON>> STRUCTURES.....	121
TABLE 63 ENGLISH R2 TEN O’CLOCK IN <<TEMPORAL: SAME TIME: [POINT]>> STRUCTURES.....	122
TABLE 64 MBF FOR HYPOTAXIS AND EXTENSION.....	124
TABLE 65 MBF FOR PARATAXIS AND EXTENSION.....	125
TABLE 66 MBF FOR HYPOTAXIS AND ENHANCEMENT.....	129
TABLE 67 MBF FOR PARATAXIS AND ENHANCEMENT.....	130
TABLE 68 MBF FOR HYPOTAXIS AND ELABORATION.....	131
TABLE 69 MBF FOR PARATAXIS AND ELABORATION.....	132
TABLE 70 CORRESPONDENCE BETWEEN EF TEACHING LEVELS AND CEFR (ALEXOPOULOU ET AL., 2015).....	137
TABLE 71 EXAMPLES OF WRITING TOPICS/UNIT NUMBER ACROSS TEACHING LEVELS (ALEXOPOULOU ET AL., 2015).....	137
TABLE 72 EXAMPLE OF CEDEL2 WRITING TASKS (CEDEL2, 2023).....	142
TABLE 73 EXAMPLE OF CAES WRITING TASKS (CAES, 2023).....	144
TABLE 74 CORPUS DATA SUMMARY.....	144
TABLE 75 DATA SUMMARY BY CEFR PROFICIENCY LEVEL.....	145
TABLE 76 EFCAMDAT2 DATA SUMMARY.....	146
TABLE 77 CEDEL2 DATA SUMMARY.....	146
TABLE 78 CAES DATA SUMMARY.....	147
TABLE 79 LEVEL 1 RESULTS FOR HYPOTAXIS AND PARATAXIS.....	154
TABLE 80 LEVEL 2 NORMALIZED EXPANSION THROUGH EXTENSION, ENHANCEMENT & ELABORATION.....	162
TABLE 81 AVERAGE WORDS PER SCRIPT EFCAMDAT2 & CEDEL2/CAES CEFR LEVELS B2-C2.....	174
TABLE 82 FREQUENCY RANKING PARATACTIC EXPANSION RELATIONS.....	178
TABLE 83 LEVEL 3 RESULTS SUMMARY.....	181
TABLE 84 LEVEL 4 RESULTS HYPOTAXIS EXTENSION AND ADDITION THROUGH POSITIVE ADDITIVE.....	220
TABLE 85 LEVEL 4 RESULTS HYPOTAXIS EXTENSION AND ADDITION THROUGH ADVERSATIVE ADDITION.....	221
TABLE 86 LEVEL 4 RESULTS HYPOTAXIS EXTENSION AND ADDITION THROUGH REPLACIVE VARIATION.....	221
TABLE 87 LEVEL 4 RESULTS HYPOTAXIS EXTENSION AND ADDITION THROUGH SUBTRACTIVE VARIATION.....	221
TABLE 88 LEVEL 4 RESULTS HYPOTAXIS EXTENSION AND ALTERNATION.....	222
TABLE 89 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH SAME TIME ENHANCEMENT.....	223
TABLE 90 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH DIFFERENT TIME ENHANCEMENT.....	224
TABLE 91 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH SAME PLACE SPATIAL ENHANCEMENT.....	226
TABLE 92 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH MANNER AND MEANS ENHANCEMENT.....	226
TABLE 93 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH MANNER AND COMPARISON ENHANCEMENT.....	227
TABLE 94 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH CAUSE AND REASON CAUSAL-CONDITIONAL ENHANCEMENT.....	228
TABLE 95 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH PURPOSE CAUSAL-CONDITIONAL ENHANCEMENT.....	230

TABLE 96 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH POSITIVE CONDITION CAUSAL-CONDITIONAL ENHANCEMENT.	231
TABLE 97 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH NEGATIVE CONDITION CAUSAL-CONDITIONAL ENHANCEMENT.	232
TABLE 98 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH CONCESSIVE CONDITION CAUSAL-CONDITIONAL ENHANCEMENT.	232
TABLE 99 LEVEL 4 RESULTS FOR HYPOTAXIS THROUGH ELABORATION.....	234
TABLE 100 LEVEL 4 RESULTS PARATAXIS EXTENSION AND ADDITION THROUGH POSITIVE ADDITIVE.....	235
TABLE 101 LEVEL 4 RESULTS PARATAXIS EXTENSION AND ADDITION THROUGH NEGATIVE ADDITIVE.	236
TABLE 102 LEVEL 4 RESULTS PARATAXIS EXTENSION AND ADDITION THROUGH ADVERSATIVE.	237
TABLE 103 LEVEL 4 RESULTS PARATAXIS EXTENSION THROUGH VARIATION.	237
TABLE 104 LEVEL 4 RESULTS PARATAXIS EXTENSION THROUGH ALTERNATION.	238
TABLE 105 LEVEL 4 RESULTS FOR PARATAXIS THROUGH TEMPORAL SAME TIME ENHANCEMENT.....	238
TABLE 106 LEVEL 4 RESULTS FOR PARATAXIS THROUGH TEMPORAL DIFFERENT TIME ENHANCEMENT.	239
TABLE 107 LEVEL 4 RESULTS FOR PARATAXIS THROUGH SPATIAL SAME PLACE ENHANCEMENT.....	240
TABLE 108 LEVEL 4 RESULTS FOR PARATAXIS THROUGH MANNER AND MEANS ENHANCEMENT.	240
TABLE 109 LEVEL 4 RESULTS FOR PARATAXIS THROUGH MANNER AND COMPARISON ENHANCEMENT.	240
TABLE 110 LEVEL 4 RESULTS FOR PARATAXIS THROUGH CAUSE AND REASON CAUSAL-CONDITIONAL ENHANCEMENT.	241
TABLE 111 LEVEL 4 RESULTS FOR PARATAXIS THROUGH POSITIVE CONDITION CAUSAL-CONDITIONAL ENHANCEMENT.	241
TABLE 112 LEVEL 4 RESULTS FOR PARATAXIS THROUGH POSITIVE CONDITION CAUSAL-CONDITIONAL ENHANCEMENT.	242
TABLE 113 LEVEL 4 RESULTS FOR PARATAXIS ELABORATION THROUGH EXPOSITION.	242
TABLE 114 LEVEL 4 RESULTS FOR PARATAXIS ELABORATION THROUGH EXEMPLIFICATION.	244
TABLE 115 LEVEL 4 RESULTS FOR PARATAXIS ELABORATION THROUGH CLARIFICATION.	245
TABLE 116 EFCAMDAT2 HYPOTAXIS AND EXTENSION MBU PEAK FREQUENCY PER 5K.	247
TABLE 117 EFCAMDAT2 HYPOTAXIS AND EXTENSION MBU PEAK FREQUENCY CEFR PROFICIENCY LEVEL.	247
TABLE 118 EFCAMDAT2 HYPOTAXIS THROUGH EXTENSION MBU DENSITY ACROSS FREQUENCY RANGES.	248
TABLE 119 CEDEL2/CAES HYPOTAXIS AND EXTENSION MBU PEAK FREQUENCY PER 5K.....	248
TABLE 120 CEDEL2/CAES HYPOTAXIS AND EXTENSION MBU PEAK FREQUENCY CEFR LEVEL.	249
TABLE 121 CEDEL2/CAES HYPOTAXIS THROUGH EXTENSION MBU DENSITY ACROSS FREQUENCY RANGES.....	249
TABLE 122 EFCAMDAT2 HYPOTAXIS AND ENHANCEMENT HIGH DENSITY LOW FREQUENCY MBUS.....	251
TABLE 123 EFCAMDAT2 HYPOTAXIS THROUGH ENHANCEMENT HDLF MBUS ACROSS FREQUENCY RANGES.	252
TABLE 124 EFCAMDAT2 HYPOTAXIS AND ENHANCEMENT HIGH DENSITY LOW FREQUENCY MBUS PER CEFR LEVEL.....	253
TABLE 125 CEDEL2/CAES HIGH DENSITY LOW FREQUENCY MBUS.	254
TABLE 126 CEDEL2/CAES2 HYPOTAXIS THROUGH ENHANCEMENT HDLF MBUS ACROSS FREQUENCY RANGES.	255
TABLE 127 CEDEL2/CAES HYPOTAXIS AND ENHANCEMENT HIGH DENSITY LOW FREQUENCY MBUS PER CEFR LEVEL.....	255
TABLE 128 EFCAMDAT2 HYPOTAXIS AND ENHANCEMENT LOW DENSITY MEDIUM FREQUENCY.	257
TABLE 129 EFCAMDAT2 HYPOTAXIS THROUGH ENHANCEMENT MBU DENSITY ACROSS FREQUENCY RANGES.	258
TABLE 130 EFCAMDAT2 HYPOTAXIS AND ENHANCEMENT LOW DENSITY MEDIUM FREQUENCY.	259
TABLE 131 CEDEL2/CAES HYPOTAXIS AND ENHANCEMENT LOW DENSITY MEDIUM FREQUENCY MBUS.....	260
TABLE 132 CEDEL2/CAES MBU DENSITY ACROSS FREQUENCY RANGES.	261
TABLE 133 CEDEL2/CAES HYPOTAXIS AND ENHANCEMENT LOW DENSITY MEDIUM FREQUENCY MBUS.....	262
TABLE 134 EFCAMDAT2 HYPOTAXIS AND ENHANCEMENT LOW DENSITY HIGH FREQUENCY MBUS.....	264
TABLE 135 EFCAMDAT2 HYPOTAXIS AND ENHANCEMENT LOW DENSITY HIGH FREQUENCY MBUS PER CEFR PROFICIENCY LEVEL.	265
TABLE 136 EFCAMDAT2 HYPOTAXIS AND ENHANCEMENT MBUS PER FREQUENCY AND CEFR LEVEL.....	267
TABLE 137 CEDEL2/CAES HYPOTAXIS AND ENHANCEMENT LOW DENSITY HIGH FREQUENCY MBUS.	267
TABLE 138 CEDEL2/CAES HYPOTAXIS AND ENHANCEMENT LOW DENSITY HIGH FREQUENCY MBUS PER CEFR LEVEL.....	270
TABLE 139 CEDEL2/CAES MBUS PER FREQUENCY AND CEFR LEVEL.	270
TABLE 140 FREQUENCY SUMMARY FOR EFCAMDAT2 & CEDEL2/CAES.	271

TABLE 141 EFCAMDAT2 HYPOTAXIS AND ELABORATION MBU PEAK FREQUENCY.	272
TABLE 142 EFCAMDAT2 HYPOTAXIS AND ELABORATION MBU PER CEFR LEVEL.	273
TABLE 143 CEDEL2/CAES HYPOTAXIS AND ELABORATION MBU PEAK FREQUENCY PER 5K.	274
TABLE 144 CEDEL2/CAES HYPOTAXIS AND ELABORATION MBU PER CEFR LEVEL.	274
TABLE 145 EFCAMDAT2 PARATAXIS AND EXTENSION MBU PEAK FREQUENCY PER 5K.	275
TABLE 146 EFCAMDAT2 PARATAXIS AND EXTENSION MBU PER CEFR LEVEL.	277
TABLE 147 CEDEL2/CAES PARATAXIS AND EXTENSION MBU PEAK FREQUENCY PER 5K.	278
TABLE 148 CEDEL2/CAES PARATAXIS AND EXTENSION MBU PEAK FREQUENCY PER CEFR LEVEL.	279
TABLE 149 EFCAMDAT2 PARATAXIS AND ENHANCEMENT MBU PEAK FREQUENCY PER 5K.	280
TABLE 150 EFCAMDAT2 PARATAXIS AND ENHANCEMENT MBU PEAK FREQUENCY PER CEFR LEVEL.	281
TABLE 151 CEDEL2/CAES PARATAXIS AND ENHANCEMENT MBU PEAK FREQUENCY PER 5K.	282
TABLE 152 CEDEL2/CAES PARATAXIS AND ENHANCEMENT MBU PEAK FREQUENCY PER CEFR LEVEL.	282
TABLE 153 EFCAMDAT2 PARATAXIS AND ELABORATION MBU PEAK FREQUENCY PER 5K.	284
TABLE 154 EFCAMDAT2 PARATAXIS AND ELABORATION MBU PEAK FREQUENCY PER CEFR LEVEL.	285
TABLE 155 CEDEL2/CAES PARATAXIS AND ELABORATION PEAK FREQUENCY PER 5K.	286
TABLE 156 CEDEL2/CAES PARATAXIS AND ELABORATION PEAK FREQUENCY PER CEFR LEVEL.	287

1. Introduction

1.1 Object of Study

Linguistic complexity is thought to be a dynamic property of a learners' interlanguage system and a more stable property of an individuals' linguistic elements which constitute the interlanguage system (Housen & Kuiken, 2009). Within the components that form linguistic complexity, syntactic complexity has been defined as the range of syntactic structures produced by language learners and the degree of sophistication in which syntactic elements are used (Ai & Lu, 2015; Ortega, 2003; Pallotti, 2009). A plethora of studies have highlighted that syntactic complexity is the most frequent and intensively used component in linguistic complexity in second language acquisition research (Kuiken et al., 2019). Researchers have operationalized multiple measurements to gauge the range and sophistication of grammatical resources employed in language production and have found syntactic complexity measures are the most reliable index to gauge proficiency in L2 writing (Ai & Lu, 2013; Bardovi-Harlig & Bofman, 1989; Ferris, 1994; Lu, 2011; Sotillo, 2000).

The meaning-based dimension of linguistic complexity has largely been lacking in research dominated by traditional syntactic complexity measures (Ryshina-Pankova, 2015). Studies have called for more research into meaning in complexity through Structural Functional Linguistics (Bulté & Housen, 2012; Norris & Ortega, 2009; Ortega, 2003, 2012; Pallotti, 2009; Ryshina-Pankova, 2015, henceforth SFL). Theoretical and empirical justifications exist for measurement of syntactic complexity through SFL as it views language development changing from dynamic to synoptic styles of expression through expansion as learners become more sophisticated in their linguistic capabilities (Ortega, 2003; Norris & Ortega, 2009).

An alternative approach to syntactic complexity in L2 writing takes into account the relationship between meaning in text production (Ryshina-Pankova, 2015, p. 52). SFL has been proposed as a theoretical justification for measuring syntactic complexity and would indicate that language development begins through the expression of ideas through parataxis, commonly known as coordination. Expansion through hypotaxis, otherwise known as

subordination, would grow to prominence at intermediate levels to express ideas through grammatically intricate texts. At advanced stages of language proficiency there would be a reliance on grammatical metaphor, in the sense of nominalizing nouns and verbs, which would exhibit lower levels of subordination but higher levels of lexical density with more complex phrases instead of clauses (Halliday & Matthiessen, 2006; Norris & Ortega, 2009).

Bulté and Housen (2018) termed the developmental sequence of syntactic interlanguage complexification hypothesis (DSSICH) referring to the explanation given by Norris and Ortega (2009) which implied expansion from dynamic to synoptic styles would mean parataxis would be a marker indicating coordination at a beginning level (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012). Hypotaxis would indicate complex language use at an intermediate level, since subordination is believed to be a resource used to express the logical connection of ideas via grammatically intricate texts. Subordination would decrease at advanced levels as learners increasingly used grammatical metaphor (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quintero et al., 1998).

Various studies have reported mixed results when investigating the DSSICH. It has been observed that as proficiency increases, the use of subordination and coordination also rises (Kuiken & Vedder, 2019; Lahuerta Martínez, 2018a; Lorenzo & Rodríguez, 2014). Subordination as an independent measurement has been seen to steadily grow as proficiency increases (Hawkins & Filipovic, 2012; Kaushik & Huhta, 2020; Lu, 2011; Wolfe-Quintero et al., 1998). The DSSICH has been partially confirmed as clausal coordination and phrasal sophistication increased but the clausal subordination significantly decreased (Lei et al., 2023; Kuiken & Vedder, 2019). Clausal subordination has also been seen to remain constant through proficiency levels (Asención-Delaney & Collentine, 2011; Bulté & Housen, 2014). In contrast, several studies have reported a slight decrease in clausal subordination as proficiency increases (Ai & Lu, 2013; Lu, 2011). Grammatical metaphor has additionally been confirmed to increase at advanced levels (Byrnes, 2009).

A line of research within the context of SFL which appears to not have received much attention is focused on taxis, expansion and logico-semantic relations to realize meaning-making. Taxis entails two concepts with the first one being parataxis and the second being hypotaxis. Expansion is comprised of extension, elaboration, and enhancement. Logico-

semantic relations such as temporal, spatial and causal-conditional indicators, are classified under each expansion element (Halliday, 2014).

The known studies incorporating parataxis have largely found expansion through extension and addition to be the main driver of L2 coordination (Rasool & Mahmood, 2023; Wenhui Xuan, 2019; Yang et al., 2017; Zarco-Tejada et al., 2016) with the one exception when paratactic elaboration was seen to have a higher use frequency (Sulistyaningrum & Rasyid, 2015). When looking into coordinators as fine-grained measures, *and* and *but* demonstrate the highest use frequencies in L2 written texts (Yang et al., 2017; Zarco-Tejada et al., 2016). Findings for hypotaxis have shown that expansion through enhancement, otherwise known as adverbial subordination, is the most often employed logico-semantic classification with L2 learners showing causal-conditional subordination to be the most often used followed by temporal subordination (Rasool & Mahmood, 2023; Sulistyaningrum & Rasyid, 2015; Wenhui Xuan, 2019).

SFL studies have occurred in a variety of L1s including Indonesian (Sulistyaningrum & Rasyid, 2015), Chinese (Wenhui Xuan, 2019) and Pakistani (Rasool & Mahmood, 2023). The larger percentage of the previously referenced studies did in-depth research into taxis and logico-semantic meaning-based categories. In contrast, Zarco-Tejada et al. (2016) with L1 Spanish learners and Yang et al. (2017) with L1 Chinese learners investigated individual coordinators in English L2 writing. Not a single known study investigated individual subordinators and coordinators in the context of all three expansion classifications and most known research used a small representative sample (Rasool & Mahmood, 2023; Sulistyaningrum & Rasyid, 2015; Wenhui Xuan, 2019). The only known study to overtly apply the *Common European Framework of Reference for Languages* (2001, henceforth CEFR) levels to measures through SFL is Zarco-Tejada et al. (2016) which investigated coordination at proficiency levels A2, B1 and B2.

The CEFR is a widely known and often used resource in creating modern language syllabuses and proficiency tests (Figuras, 2012; Leung & Lewowicz, 2013). Research into syntactic complexity through CEFR proficiency levels has shown clear differences between A2 and B2 levels in a variety of indices (Kim, 2004; Verspoor et al., 2012). A manifold of cross-sectional and longitudinal research has investigated the reliability of the vast number of syntactic complexity measures used to accurately gauge L2 proficiency in texts (Bulté & Housen, 2012; Lu, 2011; Norris & Ortega, 2009; Ortega, 2003; Wolfe-Quintero et al., 1998).

A large number of dissimilar indices has made it difficult to generalize results and the multidimensional aspect of syntactic complexity requires appropriate measures for each dimension thus further complicating the affair (Lu & Ai, 2015, p.19). Many studies have used small amounts of data (Lei et al., 2023) and few studies researching syntactic complexity have investigated fine-grained measures (Chen et al., 2021; Kyle & Crossley, 2018). Fine-grained measures are thought to have a more predictive power to gauge L2 writing quality when compared to traditional indices (Zhang & Lu, 2022). The majority of studies outlined above employed corpus linguistics as a baseline tool for linguistic research.

Corpus linguistics involves research questions which are investigated based upon the complete and systematic analysis of the conditional distribution of a linguistic phenomenon in a representative sample (Stefanowitsch, 2020). The present study incorporates three separate corpora to provide a representative sample indicative of a large data set. The updated second version of the *Education First-Cambridge Open Language Database* was (EFCAMDAT2) used to obtain L2 English texts written by L1 Spanish learners constituting a total of 771,162 words. Two corpora were selected to supply an adequate sampling for L2 Spanish texts written by L1 English learners. The *Corpus escrito del español L2* (CEDEL2) and *Corpus de aprendices de español* (CAES) were merged to come up with a total of 707,330 words. Data for the three separate corpora provided a total of 1,478,492 words. This data set is the largest known representative sample used to date involving Structural Functional Linguistics to analyze L2 Spanish and English texts.

Data extrapolated from corpus linguistic studies for native and L2 learners assist researchers in pinpointing language trends (Phakiti et al., 2018) as well as serve as a basis for comparability to gauge acquisition and L1 to L2 transfer. Previous findings involving L1 Spanish learners writing in L2 English have shown that subordination and coordination have significantly increased in younger Spanish school students (Lahuerta Martínez, 2018a). Steady progression, but not a significant increase in subordination, was seen by professional writers over a four-year period which may suggest Spanish L1 to English L2 transfer (Neff et al., 2004). Studies have called for more research looking into how Spanish L1 affects English L2 writing (Basterrechea & Weinert, 2017; Neff et al., 2004; Reyes & Hernández 2006).

Collentine and Collentine (2020) researching English L1 transfer to Spanish L2 in written texts observed that beginning and intermediate learners produced complexity in the form of causal adverbial phrases while advanced learners tend to complexify relative clauses. Restrepo-Remos (2021) conducted a semester-long study which observed an increase in the production adverbial and relative clauses with a decrease in coordination involving L1 English writing in L2 Spanish at beginning levels. The previously mentioned study found that intermediate use of relative clauses and coordination decreased while adverbial clauses slightly increased. Restrepo-Remos (2021) underlined the importance of further research in this field as few studies have focused on longitudinal growth across proficiency levels in L2 Spanish college level writing classes.

Research is a systematic process used to address a phenomenon in which linguists can apply a theoretical framework as a methodological approach founded on theory (Phakiti et al., 2018). The *Método de los Relojes* (2018, henceforth MR) is a descriptive grammar approach based on a metaphorical map of the Spanish language using three “relojes”, or “clocks” in English. Reloj 2 (R2) portrays how the majority of messages are organized in the Spanish language throughout twelve hours. Each hour shares the commonality of offering a minimum of two verbs interacting together in a grammatical sequence with a subordinator thus subscribing to the definition of hypotaxis. This study used the MR R2 to establish a theoretical viaduct between target languages, CEFR proficiency levels and SFL to create a quantifiable meaning-based framework (MBF) that can be used to analyze large data sets. The use-case scenario for the MBF was generated to incorporate fine-grained meaning-based indices as an alternative to traditional indices to gauge meaning-based complexity in L2 writing (Ryshina-Pankova, 2015; Zhang & Lu, 2022). The following section details the proposed research questions followed by a synopsis of the remaining contents in this thesis.

1.2 The Present Study

The research objectives of this study are multifaceted. Within the greater concept of syntactic complexity, studies have called for more research into meaning in complexity through SFL (Bulté & Housen, 2012; Norris & Ortega, 2009; Ortega, 2003, 2012; Pallotti, 2009; Ryshina-Pankova, 2015). The first goal of this study is to triangulate a measurement system using fine-grain measures taken from *An Introduction to Functional Grammar* (2014) and *Método de los Relojes* (2018) R2 coupled with CEFR proficiency levels to accurately

calculate meaning-based complexity levels and draw conclusions based on the frequency of targeted forms in taxis, expansion and logico-semantic relations as well as meaning-based units (MBUs), a term used by this study for individual indicators of subordination and coordination. The first research question is designed to achieve this goal.

RQ1: To what extent can taxis, expansion, logico-semantic relations and MBUs be operationalized through the *Método de los Relojes* R2 and CEFR levels A1-C2 to create an effective and accurate framework to measure meaning-based complexity?

Given the amalgamation of theoretical and cross-linguistic influences, this study expects a meaning-based framework can be formed to measure subordination and coordination frequencies in taxis, expansion, logico-semantic relations and MBUs in the large data set taken from the second version of the *Education First-Cambridge Open Language Database*, *Corpus escrito del español L2* and *Corpus de aprendices de español*. We hypothesize the crux of RQ1 will entail sorting through MBUs to systematically compile compatible measures which will gauge the reality of language use in L2 written texts.

The second research question investigates the DSSICH (Norris & Ortega, 2009) through the viewpoint of the meaning-based framework. Testing this hypothesis would indicate parataxis would be highest at beginner levels (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012). Hypotaxis through subordination would then show peak frequency levels at intermediate levels and then decrease at advanced levels (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). The second research question is proposed to see if findings are consistent with results in previous studies.

RQ2: Does the meaning-based framework validate the developmental sequence of syntactic interlanguage complexification in English and Spanish L2 texts?

On the basis of the previously detailed studies which have looked into the DSSICH (Ai & Lu, 2013; Asención-Delaney & Collentine, 2011; Bulté & Housen, 2014; Kuiken & Vedder, 2019; Lahuerta Martínez, 2018a; Lei et al., 2023; Lorenzo & Rodríguez, 2014; Lu, 2011; Norris & Ortega, 2009) in different L1s, we do not expect to find significant correlations with the hypothesis proposed by Norris and Ortega (2009) in which syntactic

complexity measures show a higher use of coordination at beginner levels and increased frequency of subordination at intermediate levels with a regression at advanced levels. This expectation is based on the alternative approach involving the application of a meaning-based framework in combination with a large data set. The main goal of research question two is to identify and clarify the way a meaning-based framework can indicate frequency patterns specific to L2 English and Spanish written texts to identify learner trends.

The third research question takes an in-depth look into how SFL principles are used in English and Spanish L2 texts. The justification lies in the perceived lack of research in this area with only five known studies looking into the matter (Rasool & Mahmood, 2023; Sulistyaningrum & Rasyid, 2015; Wenhui Xuan, 2019; Yang et al., 2017; Zarco-Tejada et al., 2016). Given that any known research does not comprehensively explore the cross-linguistic interrelations in English and Spanish L2 texts through taxis, expansion, logico-semantic relations and MBUs in the scaffolding of CEFR proficiency levels A1-C2, research question three tackles this matter head on.

RQ3: To what extent does the meaning-based framework identify cross-linguistic meaning-based complexity patterns through taxis, expansion, logico-semantic relations and MBUs in English and Spanish L2 texts in relation to CEFR proficiency levels?

The primary focus of research question three is to distinguish how L2 learners use taxis, expansion, logico-semantic relations and MBUs across CEFR proficiency levels. Previous research (Rasool & Mahmood, 2023; Wenhui Xuan, 2019; Yang et al., 2017; Zarco-Tejada et al., 2016) leads this study to make a generalized hypothesis that users will use hypotaxis through enhancement with causal-conditional and temporal indicators at a high frequency. Additionally, parataxis through extension is expected to be used at overall higher frequencies throughout the CEFR proficiency scale. Research question three digs deep into how learners use each previously noted category in fine-grained detail to add the body of knowledge in this line of research, explore new possibilities for investigation and develop new hypotheses based on observed frequency patterns.

The focus of research question four concerns the extent to which the MBF can identify trends that indicate L2 acquisition and L1 to L2 transfer. Fine-grained measures will allow for a granular analysis into use patterns detailing what types of taxis, expansion, logico-

semantic relations and MBUs are being used by learners and at what levels. The following research question is formulated to investigate a matter in which no other known study using SFL has delved into.

RQ4: To what extent can the meaning-based framework highlight frequency trends which indicate L2 acquisition and L1 to L2 transfer?

1.3 Thesis Structure

The structure of this investigation starts with the literature review in Chapter 2 which outlines the theoretical foundation of SFL by first going into a brief overview in 2.1 followed by 2.1.1 which delves into the basic definition of parataxis and hypotaxis. Elaboration, extension and enhancement are reviewed to explain the conceptual fundamentals of expansion in 2.1.2 with a detailed explanation of logico-semantic categories and MBUs.

Complexity measures are then critiqued in 2.2. A definition of syntactic complexity is provided in 2.2.1 followed by a recap of research using the vast set of extensive measures employed by relevant studies as well as calls for a meaning-based approach. As an integral part of this thesis, 2.2.2 dives into the CEFR followed by a summary of studies using proficiency levels to measure syntactic complexity in L2 writing. 2.2.3 recounts all known studies involving SFL meaning-based indices to explore what research has been carried out, underline pertinent findings and pinpoint areas of opportunity to further knowledge in this field. Acquisition of subordinate clauses and how L1 transfer plays a role in this process for English and Spanish L2s is covered on 2.3.

Crucial to the underpinning of this thesis, the *Método de Los Relojos* (2018) was used as theoretical cairn in connecting SFL with English and Spanish languages through subordination. 2.4.1 to 2.4.3 broadly describes various aspects of said method before going into R2 in 2.4.4. Although R2 contains a total of twelve hours which largely describes how subordination through a primary and secondary verb is used in Spanish, this section details One o'clock through Ten o'clock as the final two hours are not applicable to the research at hand. Finally, 2.5 provides a summary of the section.

The basis for Chapter 3 is the creation of a MBF using measures grounded in SFL. 3.1 provides a brief introduction followed by a description of the methodology used to establish the framework in 3.2. Section 3.3 comprehensively details R2 hours One o'clock through Ten o'clock to ascertain a cross-linguistic and theoretical analysis of English and Spanish subordination as well as establish a basis for comparability. 3.4 is the final section in this chapter and presents the MBF which will be used as an instrumental tool in the analysis of the corpus data outlined in the following chapter.

Chapter 4 applies the MBF to analyze L2 English and Spanish texts. The first part of the chapter provides an introduction in 4.1. The methodology employed in this investigation is outlined in 4.2 which includes descriptions of the three corpora used to provide a representative sample in 4.2.1. A summary of data from each corpus can be found in 4.2.2 with a breakdown by CEFR level, words, learners, scripts, gender and collection dates. 4.2.3 chronicles the technical details and evaluation employed in analyzing data which serves to clarify how results were achieved.

The final four sections form the empirical analysis of this thesis, reporting on results of four separate experiments culminating in the proposal of a novel meaning-based approach to complexity. Each experiment contains subsections with separate results, discussions and conclusions to divide information into manageable pieces. Experiment 1 in 4.3 explores the broad categories of hypotaxis and parataxis. Experiment 2 in 4.4 investigates elaboration, extension and enhancement in parataxis and hypotaxis. Experiment 3 further subdivides expansion onto logico-semantics meaning-based categories (MBCs) in 4.5. Finally, 4.6 deals with each singular coordinator and subordinator to provide a fine-grained analysis of how each MBU trends and sequences throughout proficiency levels. This level of analysis delivers an explanation for how patterns found in each of the three previous experiments were established. As a result of the findings in hypotaxis and enhancement, 4.6.3.1 introduces the developmental sequence of meaning-based interlanguage complexification hypothesis (DSMBISH) as an alternative to the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Chapter 5 summarizes the key findings for research questions one through four. Additionally, the need for more research into meaning-based complexity is addressed through

a variety of recommendations for future research. Paramount to increasing the body of knowledge in this category, a significant area for investigation would involve developing new or larger meaning-based frameworks through natural language processing and applying them to different L2s to test the developmental sequence of meaning-based interlanguage complexification hypothesis.

2. Literature Review

This section examines the theoretical models of Structural Functional Linguistics, *Método de los Relojes* (2018) R2 and the *Common European Framework of Reference for Languages* (2001) to establish a meaning-based framework for measuring complexity. The MBF incorporates taxis, expansion, logico-semantic relations and MBUs to offer an alternative perspective to analyze syntactic complexity when measuring L2 cognitive development, L2 acquisition and L1 to L2 transfer through CEFR proficiency levels.

Section 2.1 begins by reviewing M.A.K. Halliday's approach to Structural Functional Linguistics as found in an *Introduction to Functional Grammar* (2014). The characteristics of parataxis and hypotaxis are described in 2.1.1 followed by an in-depth account of the three categories of expansion in 2.1.2 including elaboration, extension and enhancement. The purpose of beginning the review with an explanation of SFL is to lay out a theoretical foundation which will need to be understood in later chapters.

Section 2.2 first dives into a description complexity measures before confronting the broad definition of syntactic complexity and the extensive measures used in a variety of studies (Section 2.2.1). A summary of studies investigating the relationship between syntactic complexity measures and L2 writing using the CEFR is included in 2.2.2. The definition of the CEFR relating to syntactic complexity is detailed in 2.2.2.1. A review of studies involving syntactic complexity and CEFR proficiency levels is provided in 2.2.2.2. Section 2.2.3 assimilates Structural Functional Linguistics with the concept of syntactic complexity by outlining relevant theory and reviews current research in the matter.

L1 acquisition of L2 subordination in Spanish and English is outlined in 2.3. To give a more accurate comparison of findings relevant to this study, 2.3.1 looks into how the Spanish L1 has an effect on English L2 writing and likewise 2.3.2 delves the ramifications of English L1 on Spanish L2 writing. Section 2.3 was included to give an empirical backdrop related to L1 acquisition of L2 subordination in Spanish and English with studies that do not incorporate CEFR levels or a meaning-based approach.

The MR is outlined in 2.4 with a general summary of the subject matter followed by summarized descriptions of the properties of the four verb types (Section 2.4.1), Versions 1:-5: (Section 2.4.2) and the five basic verb forms (Section 2.4.3). The majority of the content in this section is devoted to the organic properties of R2 (Section 2.4.4). Sections 2.4.4.1 through 2.4.4.10 are devoted to explaining the details of hours One o'clock until Ten o'clock. Finally, 2.5 summarizes information included in the entirety of this section.

2.1 Structural Functional Grammar: A Clause-Based Standpoint

M.A.K. Halliday and Christian Matthiessen wrote and edited *Introduction to Functional Grammar* (2014) as the first presentation of the functional framework for explaining English grammar. The authors set themselves apart by interpreting English grammar from clause-based standpoint rather than word-based. Halliday makes a significant contribution to functional grammar with his interpretation of grammatical function as the subject of the sentence. The subject is represented in dialogic terms and its functional contribution is the exchanges between speaker and listener. In this manner Halliday puts forward a functionally different notion of the clause by illuminating how the traditional notion of subject is an amalgamation of textual clause function (Matthiessen, 1989, p. 863).

This study focuses on what Halliday (2014) explains as “Above the clause” which describes the clause complex and investigates how clauses are linked together through logico-semantic relations to form clause complexes representing sequences of figures that are presented as textually related messages. The notion of “Above the clause” takes the point of view on how the flow of events are construed in the development of texts at the level of semantics.

All clauses linked through logico-semantic relations are considered interdependent. This means the relational structure is built on one unit that is mutually dependent on another structure. The structures in a clause complex may be of equal status (parataxis) or unequal status (hypotaxis). The following section delves into parataxis and hypotaxis to lay out a detailed foundation as to how these processes are construed.

2.1.1 Taxis

The clause complex has a degree of interdependency which is technically known as taxis. There are two degrees of interdependency. Parataxis denotes an equal status which is the relation of two like elements with one initiating and the other continuing. Hypotaxis represents an unequal status denoting the relation between a dependent element and a dominant element. A pair of clauses related by interdependency, or taxis, is referred to as a clause nexus. Within the clause nexus, parataxis is assigned numbers to express each unit with 1 initiating clause followed by 2 indicating a continuing clause. Hypotactic clauses, symbolized by Greek letters α and β , are made up of a dominant and a dependent clause as seen below (Halliday, 2004, pp. 374-375).

	Primary	Secondary
Parataxis	1 (initiating)	2 (continuing)
Hypotaxis	α (dominant)	β (dependent)

Within the logico-semantic relations of expansion and projection, there are a wide number of relations which may come in between the primary and the secondary clause. The two general types are expansion and projection. When considering expansion (1), the secondary clause expands the primary clause by elaborating, extending or enhancing it. Expansion relates the phenomena as being of the same order of experience. Projection (2) takes place when the secondary clause is projected through the primary clause which instates it as a locution or an idea (Halliday, 2014, pp. 442-443). Since the focus of this thesis is expansion, no further reference will be made to projection.

- (1) *If we get enough time, nobody in the audience will be able to see through the disguises.*
- (2) *Gandhi next asked her if she knew what a spinning wheel was.*

Expansion encompasses three different subtypes. Elaborating clauses contain one clause that expands on another by elaborating on it in terms of restating word order, specifying greater detail, commenting or exemplifying. Extension happens when one clause expands another by extending beyond it through adding a new element, giving an exception to it or offering an alternative. Enhancement occurs when one clause expands another through

embellishing around it, thus qualify it with some circumstantial feature of time, place, cause or condition. Table 1 provides a summary of parataxis and hypotaxis under the elaboration, extension and enhancement through expansion (Halliday, 2014, p. 447).

Parataxis is the linking of elements of equal status in which the initiating and continuing elements are free since they can stand as a functioning whole. Hypotaxis, the binding of unequal elements in a clause, takes place when the dominant element is free, and the dependent element is not. Moreover, the hypotactic relation is logically non-symmetrical and non-transitive (Halliday, 2014, p. 452).

		Paratactic	Hypotactic
(1) expansion	(a) elaboration	1 <i>John didn't wait;</i> =2 <i>he ran away.</i> (apposition)	α <i>John ran away</i> = β <i>which surprised everyone.</i> (non-defining relative clause)
	(b) extension	1 <i>John ran away,</i> =2 <i>and Fred stayed behind.</i> (coordination)	α <i>John ran away,</i> + β <i>whereas Fred stayed behind.</i>
	(c) enhancement	1 <i>John was scared,</i> =2 <i>so he ran away.</i>	α <i>John ran away,</i> + β <i>he was scared.</i> (adverbial clause)

Table 1 Summary of Expansion Clause Complexes (Halliday, 2014, p. 447).

Conjunctions may be used in the secondary clause as binders in parataxis. Linkers such as *and*, *or* and *but* are used in parataxis when the logico-semantic relation is expansion. Linkers can additionally serve a cohesive function in clauses. Hypotactic binders are used with expansion (*when*, *while*, *because*, *since*, *if*, *although*). The significance of the binding conjunctions in expansion is that many of them, specifically *wh*-words, evoke post non-finite verb forms. In either parataxis or hypotaxis, correlative conjunctions may be used when the second conjunction marks the primary clause. The hypotactic clause nexus may include a finite or non-finite verb form (3) (Halliday, 2014, pp. 452-454).

(3) ***On coming to the thicket, he heard the faint rustling of leaves.***

Parataxis incorporating the linker *and* takes place in (4). Hypotaxis using binders and a conjunctive preposition appears in (5). Example (6) shows parataxis with a correlative conjunction. Hypotaxis with correlatives can be seen in (7) with the use of *if* and *then* (Halliday, 2014, p. 454).

- (4) *Make your way back towards the Pump House **and** walk under Pier Street to the southern end of Darling Harbour.*
- (5) *Follow the pathways around the landscaped gardens and over the bridges **before** resting at the tea house.*
- (6) *He is **either** holidaying **or** he's on another job.*
- (7) ***If** the majority say we go **then** we're prepared to go with it.*

Table 2 shows that within the hypotactic clause nexus the dependent clause can be finite or non-finite. Ellipsis can occur between the subject and non-finite dependent clauses as seen with *to* in (b). The non-finite dependent clause is generally co-referential with the subject of a dominant, although examples can be found where it is not co-referential with the subject of a dominant clause (Halliday, 2014, p. 454).

a	finite	<i>As he came to the thicket,</i>	<i>he heard the faint rustling of leaves.</i>
	non-finite	<i>On coming to the thicket,</i>	
		$x\beta$	α
b	finite	<i>He headed for the river</i>	<i>so that he could wash his wound.</i>
	non-Finite		<i>to wash his wound.</i>
		α	$x\beta$
c	non-finite	<i>I told him</i>	<i>to send it off.</i>
	finite		<i>that he should send it off.</i>
		α	" β "

Table 2 Finite and Non-finite Dependent Clauses (Halliday, 2014, p. 454).

2.1.2 Expansion: Elaborating, Extending and Enhancing

The three ways for expanding a clause are by elaborating it, extending it or enhancing it. The systems defined as “Types of Expansion” delineate regions within a continuous semantic space. At certain points in the system the line between distinguishing expansion and projection may appear may become very close (Halliday, 2014, p. 460).

2.1.2.1 Elaboration

Elaboration is realized when one clause elaborates on another by further specifying or describing it. The secondary clause does not introduce a new element into the sentence but rather provides a further characterization of one that is already there by restating it, clarifying it, refining it, or adding a descriptive attribute or comment (Halliday, 2014, pp. 461-464).

Paratactic elaboration (1=2) gives way to three categories with which the first two are considered forms of apposition:

- (i) exposition ‘*in other words*’
- (ii) exemplification ‘*for example*’
- (iii) clarification ‘*to be precise*’

Exposition (i) occurs when the secondary restates the thesis of the primary clause in different words to present it from another point of view or to reinforce the message. The connection between clauses can be made explicit by conjunctive expressions such as *or* (rather), *in other words*, *that is to say* or *I mean*. This is frequently realized by *i.e.* in writing. It is common for a lexico-semantic link to accompany elaboration as exemplified with *needed-healthy* (8). When limited, the second clause tends to repeat the first through lexical repletion or synonymy thus reinforcing the message (9) (Halliday, 2014, pp. 461-463).

(8) III *I probably needed that*; II *it was very healthy*. III

(9) III *Where’s our cake?* II *It’s coming, it’s coming*. III

Exemplification (ii) in parataxis comes about when the secondary clause develops the thesis of the primary clause by becoming more specific about it. This frequently happens

when it cites the actual example with explicit conjunctives like *for example, for instance* and *in particular*. This is normally realized by *e.g.* in writing (Halliday, 2014, p. 463).

Clarification (iii) in parataxis clarifies the thesis of the primary clause in which it supports with some kind of explanation or explanatory comment (10). Clarification in the clause often involves a shift in polarity, from positive to negative or vice versa (11). Clarification may come as an evaluative statement. Expressions found in connection with clarification are *in fact, actually, indeed* and *at least* with the closest written abbreviation being *i.e.* or *viz.* Conjunctives are cohesive rather than structural markers of a paratactic relationship with the two clauses often being juxtaposed. In written language, apposition may be signaled by a special punctuation mark like a colon (Halliday, 2014, p. 464).

(10) III *They used to work over here*; II *that's* [[***how they met***]]. III

(11) III *I wasn't surprised* II – *it was* [[***what I had expected***]]. III

Hypotactic elaboration, the notion that $\alpha = \beta$, gives way to non-defining relative clauses which serve to describe a primary clause (12) in finite or non-finite forms. At this point it should be noted that this section details finite non-defining relative clauses since non-finite forms fall outside of the scope of this study. Elaboration under hypotaxis is a strategy for introducing background information into the discourse, characterization, an interpretation of some act of the dominant clause or a form of evaluation. Moreover, it can be a sense of an explanatory comment (13) like clarification under paratactic elaboration. There are special cases in which the dominant clause can be elaborated more than once (Halliday, 2014, pp. 464-465).

(12) III *Yu, << who has been visiting Taiwan this week, >> did not elaborate.* III

(13) III *So we picked Iowa* II *because that was closer to Wyoming,*
II ***where he was from.*** III

The first hypotactic elaboration grouping pertains to clauses with *which* whose domain is either the whole of the primary clause or some part of it that is a nominal group (14). The act of *talking down to people automatically puts their backs up* provides an evaluation of the primary clause introduced by *which* (Halliday, 2014, pp. 465-466).

- (14) III *He talks down to people*, II **which** *automatically puts people's backs up*. III

The second grouping concerns *which* (15), sometimes *that*, *who* (16) or *whose* in which their domain is a nominal group. When the nominal group is non-final in the primary clause, the secondary clause is often enclosed so that it follows it (17). The structure in the previous example is $\alpha \ll \beta \gg$. The angle brackets denote enclosure which are doubled when there is a delimited element in a clause (Halliday, 2014, p. 466).

- (15) III *This was the first English Department class at the University of Ibadan*, II **which** *had just been founded*. III
- (16) III *People had trouble* II *working with Doc Humes*, II *so I got hold of George Plimpton*, II **who** *was at Cambridge then*. III
- (17) II *Inflation*, \ll **which** *was necessary for the system*, \gg *became also lethal*. III

The third category deals with clauses with *when* (18) or *where* (19) as having domain over some expression of time or place. Clauses with *where* often refer to abstract space (20). Furthermore, the elaborated clause may be enclosed in this group as with *that* in (21). The previous example illustrates a marked theme of time or place. Additionally, *when* (22) and *where* can be found in elaborations of temporal expressions (Halliday, 2014, p. 466).

- (18) III *The first few days are a time for adjustment*, II **when** *the kitten needs all the love and attention* $[[$ you can give it $]]$. III
- (19) III *Go up three flights of escalators to the Podium Level*, **where** *lifts leave for Sydney Tower Observation Deck (\$5,00 adults, \$3.00 children)*. III
- (20) III *Now consider the opposite situation*, II **where** *the velocity decreases*. III
- (21) III *One evening*, \ll *when the boy was going to bed*, \gg *he couldn't find the china dog* $[[$ that always slept with him $]]$. III
- (22) *That night*, \ll **when** *Kukul slept on his straw mat*, \gg *Chirumá came upon him*. III

2.1.2.2 Extension

Extension materializes when one clause extends the meaning of another by adding something new to it. The element can be an addition, replacement or an alternative. A single

system of categories can be used for both kinds of expansion relations, however there are certain gaps in the paradigm which include negative additive relations which are only paratactic and not hypotactic (Halliday, 2014, p. 471).

Hypotactic extension markers (Table 3) are mixed in origin and take the form of enhancing binders like *while*, *if* in *if ... not (... then)* and linkers followed by *that* which appear as *except that* and *but (for the fact) that*. Conjunctive prepositions and preposition groups appear as hypotactic extension markers and include *besides*, *without*, *apart from*, *instead of*, *other than*, etc. The two additive markers of finite clauses, being *while* and *whereas*, are used in the sense of *and* with positive addition and *but* in adversative (Halliday, 2014, p. 471).

Extension	Category	Meaning	Parataxis Marker	Hypotaxis Finite Marker	Hypotaxis Non-Finite Marker
addition	‘and’, additive: positive	X and Y	(both...) <i>and; not only...but also</i>	<i>while, whereas</i>	<i>besides, apart from, as well as</i>
	‘nor’, additive: negative	not X and not Y	(neither...) <i>nor</i>	–	–
	‘but’, adversative	X and conversely Y	<i>but, (and) yet</i>	<i>while, whereas</i>	<i>without</i>
variation	‘instead’, replacive	not X but Y	<i>but not; ... but</i>	–	<i>instead of, rather than</i>
	‘except’, subtractive	X but not all X	<i>only, but except</i>	<i>except that, but (for the fact) that</i>	<i>except for, other than</i>
alternation	‘or’	X or Y	(either...) <i>or (else)</i>	<i>if ... not (... then)</i>	–

Table 3 Categories of Extension and Hypotactic Non-finite Markers (Halliday, 2014, p. 471).

Paratactic extension, the idea clause 1 plus clause 2, gives ways to coordination between clauses. This notion is typically expressed by *and*, *nor*, *or* and *but*. There are three subtypes which include addition, variation and alternation. Addition takes place when one process is adjoined to another and there is no implication of any casual or temporal

relationship between them. There are three subtypes which include the additive positive *and*, the additive negative *nor*, the adversative *but* and *and yet*. Paratactic additions are often accompanied by cohesive expressions serving as conjunctive adjuncts such as *too*, *in addition*, *also*, *moreover* and *on the other hand* (Halliday, 2014, p. 472).

Additive parataxis through positive extension takes place in (23) and (24). The referents of the two processes can have a relation in the world of experience, if they share the same semiotic plane. Nevertheless, they must share simultaneity. Paratactically related clauses introduced by *and* are often additive extensions, yet other possibilities exist. When the meaning is <<and then>> or <<and so>>, the hypotactic version is enhancing the dependent clause, and the paratactic nexus is considered as one of enhancement instead of extension. If the clause starts with *and that* or *and this*, with *that/this* referring back to some part of the previous clause, the meaning might be one of elaboration, especially if the continuing clause is a relational one (25) (Halliday, 2014, p. 473).

- (23) III *He'd been a medieval history student in college* II ***and*** *I was interested in medieval literature, too.* III
- (24) III *There was much sickness in the corps,* II ***and*** *the men were, in addition, without clothing, shoes, and blankets needed for the winter weather.* III
- (25) III [1] *But we've got to find those* II [=2] ***and that*** *is the hard part.* III

An example of paratactic extension linked by an additive negative relation appears in (26). Since *nor* realizes negative clausal polarity, it includes a finite verb. Therefore, the sequence, appearing with a tonal break (^), is *nor* ^ Finite ^ Subject. Clauses linked with an adversative relation appear in (27) with the linker *but* containing the same semantic feature of *and* (Halliday, 2014, p. 473).

- (26) III *He could neither explain the whole situation to the editor* II ***nor*** *could he accept his rebuke.* III
- (27) III *We liked that breed of dog,* II ***but*** *we felt* II *we weren't in a position* [[*to own one at the time*]]. III

Variation under paratactic extension is presented as being the total or partial replacement of a clause. There are two subtypes which include the replacive *but* (28) and

instead (29) or the subtractive *except*. The clauses related in this way often differ in polarity value in which one can be negative and the other positive. Therefore, *but* is not adversative so it cannot be replaced with *yet*. Neither is *but* concessive since it does not correspond with the hypotactic *although*. Expressions with total replacement include *instead* and *on the contrary*. A clause linked by a subtractive relation can be seen with *only* in (30) (Halliday, 2014, p. 473).

- (28) III *Witnesses said* II *the sand dredger seemed to go past the Marhioness* II ***but*** *suddenly smashed into the side* II *and went right over it.* III
- (29) III *They should not be broad statements* [[*saying* II *where we hope to be*]], II ***instead*** *plans* [[*specifying* II *what we want to do next* II *and exactly how we are going to do it*]].
- (30) III *Nelly looked rather put out* II *and replied* II *that he was quite all right,* II ***only*** *the poor little chap was highly strung.* III

Alternation in paratactic extension is realized when one clause is presented in an alternative order as in (31) and (32). The correlative pairing takes the form of *either - or* and the associated cohesive conjunctions include *conversely*, *alternatively* and *on the other hand* (Halliday, 2014, p. 474).

- (31) III ***Either*** *you go ahead and take the plunge* II ***or*** *you wait* II *till you think* II *you can afford it,* II *which you never will.* III
- (32) III *The melt is then cooled at a few degrees per hour* II *until crystals start to form,* II ***or alternatively*** *the flux is evaporated at a constant rate.* III

Hypotactic extension, $\alpha + \beta$, incorporates addition, variation and alternation with the extending clause as the dependent one. The dependent clause may be finite or non-finite. Finite addition in combination with hypotactic clauses are introduced by the conjunctions *whereas* (33) and *while* (34). There is no clear line between the positive additive and the adversative in which these clauses sometimes have an adversative component and other times not (Halliday, 2014, p. 474).

- (33) III *Whereas* most children's fathers worked at an office, II my father worked at a studio, II so I went on the set. III
- (34) III And yet Frank grows up, II *while* Huck never grew up. III

Variation in hypotactic extension transpires when there is no finite for replacement. Subtraction comes about when the finite clause is introduced by *except that* (35) and *but that* (36). Finite clauses with *whereas*, *while* and *except that* take on a paratactic appearance when they follow a primary clause. The line between parataxis and hypotaxis is not very sharp as a working rule. If the extending clause could precede, the relationship is hypotactic. (37) is an example of a clause where the extending clause could not precede and would be interpreted as paratactic since in such instances the conjunction is unaccented (Halliday, 2014, pp. 474-475).

- (35) III Camara pulls back to show Kane and Susan in much the same position as before, II *except that* they are older. III
- (36) III Language began II when interjections ended II *but that* man still utters cries and used interjections II and that their significance is merely affective, i.e., expressing fear, surprise, etc. III
- (37) III He pretended to know all about it II – *whereas* in fact he had no idea of what was happening. III

Hypotactic extension involving alternation materializes with *if... not* with the dependent clause usually coming first (38) (Halliday, 2014, p. 475).

- (38) III *If* they're **not** in their usual place II they could have fallen through the – III

Non-finite extending clauses cover addition and variation. Moreover, the two subtypes which are not present in the non-finite system are 'negative additive' addition and 'alternative' variation. The non-finite form of an extending clause is imperfective in which $\alpha + \beta$ (39). The non-finite clause is often introduced by a preposition or a prepositional group functioning conjunctively. *Apart from* (40), *besides* (41) and *with* (42) appear in additive non-finite hypotactic extending clauses (Halliday, 2014, p. 475).

- (39) III *We used to go away at the weekend*, II **taking** all our gear with us. III
- (40) III **Apart from** being amusing II *what else does The Nun's Priest's Tale do?* III
- (41) III **Besides** being gifted with literary talent, II *Amir Khusrau was a musician too*. III
- (42) III *Most families are dependent on two salaries coming into the home*, II **with** women now constituting half of the work force. III

Non-finite clauses in this category are often introduced by a preposition or prepositional group functioning conjunctively. Addition non-finite hypotactic extending clauses can take the form of *apart from* (43). Additive and adversative clauses can take shape using the preposition *without* (44). Variation can be replacive using *instead of* (45) or subtractive using *other than* (46) (Halliday, 2014, p. 475).

- (43) III **Apart from** being amusing III *what else does The Nin's Priest's Tale do?* III
- (44) III *The arrow changed its course* II *and fell to the ground* II **without** harming anyone. III
- (45) III **Instead of** finding the perpetrators, II *they criminally charged the Earth First activist*, II *who was left crippled for life*. III
- (46) *We call him a murderer*, II *but for him there is no way out* II **other than** doing the deed. III

2.1.2.3 Enhancement

Hypotactic Enhancement $\alpha \times \beta$ happens in Table 4 when one clause enhances the meaning of a another by qualifying it by including reference to time, place, manner, cause, condition, etc. Just like extension, the parallel between parataxis and hypotaxis is very close, although there are certain gaps in principle categories in the paradigm (Halliday, 2014, p. 476).

Enhancement	Category	Meaning	Parataxis	Hypotaxis		
				Finite	Non-finite: Conjunction	Non-finite: Preposition
temporal	same time	A meanwhile B	(and) <i>meanwhile</i> ; (when)	[extent] <i>as, while</i>	<i>while</i>	<i>in</i> (the course/ process of)
			-	[point] <i>when, as soon as, the moment</i>	<i>when</i>	<i>on</i>
			-	[spread] <i>whenever, every time</i>	-	-
	different time: later	A subsequently B	(and) <i>then</i> ; <i>and</i> + <i>afterwards</i>	<i>after, since</i>	<i>since</i>	<i>after</i>
	different time: earlier	A previously B	<i>and/but</i> + <i>before that/first</i>	<i>before, until/till</i>	<i>until</i>	<i>before</i>
spatial	same place	C there D	<i>and there</i>	[extent] <i>as far as</i>	-	-
			-	[point] <i>where</i>	-	-
			-	[spread] <i>wherever, everywhere</i>	-	-
manner	means	N is via/by means of M	<i>and</i> + <i>in that way</i> ; (and) <i>thus</i>	-	-	<i>by means of</i>
	comparison	N is like M	<i>and</i> + <i>similarity</i> ; (and) <i>so, thus</i>	<i>as, as if, like, the way</i>	<i>like</i>	-
causal-conditional	cause: reason	because P so result Q	[cause ^ effect] (and) <i>so</i> ; <i>and</i> + <i>therefore</i>	-	-	-
			[effect ^ cause] <i>for</i> ; (because)	<i>because, as, since, in case, seeing that, considering</i>	-	<i>with, though, by, at, as a result, because of, in case of</i>
	cause: purpose	because intention Q so action P	-	<i>in order that, so that</i>	-	(in order/so as) <i>to</i> ; <i>for</i> (the sake of), <i>with the aim of</i> , <i>for fear of</i>

	-	-	<i>so that</i>	-	<i>to</i>
cause: result condition: positive	if P then Q	(and) <i>then; and</i> + <i>in that case</i>	<i>if, provided</i> <i>that, as long</i> <i>as</i>	<i>if</i>	<i>in the event</i> <i>of</i>
condition: negative	if not P then Q	<i>or else; (or)</i> <i>otherwise</i>	<i>unless</i>	<i>unless</i>	<i>but for,</i> <i>without</i>
condition: concessive	if P then contrary to expectation Q	[concession ^ consequence] <i>but; (and) yet,</i> <i>still; but +</i> <i>nevertheless</i>	<i>even if, even</i> <i>though,</i> <i>although,</i> <i>while</i>	<i>even if, even</i> <i>though,</i> <i>although</i>	<i>despite, in</i> <i>spite of,</i> <i>without</i>
		----- [consequence ^ concession] (<i>though</i>)			

Table 4 Enhancement through Parataxis and Hypotaxis (Halliday, 2014, pp. 477-478).

Long sequences are more probable to be construed in paratactic constructions than hypotactically. Paratactic temporal sequences play a significant role of the construction of event lines in stories, recounts, procedures and other passages in texts where chronology is an important organizing principle. While the enhancing paratactical subtype is often the same throughout the entire series, extended hypotactic chains also occur and may maintain the same logico-semantic subtype throughout as with (47) displaying causal-conditional purpose and (48) where there is a switch from concession to time (Halliday, 2014, p. 476).

- (47) III [α :] *Everyone at VES is working hard* II [$x\beta$:] *to change the law* II [$x\gamma$:] ***so that*** *we will have voluntary euthanasia legalized in England within the next five years.* III
- (48) III [1α :] *Two men were killed by lethal injection in Texas this year,* II [$1x\beta$:] ***even though*** *they were 17* II [$1x\gamma$:] ***when*** *they committed their offences,* II [$+2$:] *and another 65 juveniles are on death row across the country.* III

Paratactic enhancement, 1 x 2, gives way to a kind of coordination but with a circumstantial feature built into it. This often happens in the subtypes of time and cause. The circumstantial feature is normally expressed by three different groups of conjunctions. The first group uses conjunctions such as *then, so, for, but, yet* and *still*. Secondly, the circumstantial feature can be expressed by a conjunction group with combinations like *and then, and there, and thus, and so* and *and yet*. Lastly, *and* in combination with a conjunctive expression tends to be cohesive with *at that time, soon afterwards, till then, in that case, in*

that way, etc. Table 5 lists examples of the different types of paratactic enhancement (Halliday, 2014, p. 478).

Enhancement	Category	Example
temporal	same time	<i>It's the Cheshire Cat: now I shall have somebody to talk to.</i>
	later time	<i>The three soldiers wandered about for a minute or two, and then quietly marched off after the others.</i>
spatial	same place	<i>I ran downstairs and there he was nearly fully dressed, all back to front.</i>
manner	means	<i>Keep on subtracting the difference, and in that way you will arrive at the correct figure.</i>
	comparison	<i>Your body goes on changing every instant and so does your mind.</i>
causal-conditional	(a) cause: reason – effect ^ cause	<i>In her books, Tove Jansson spoke initially to children, so the hero is himself quite young.</i>
	(b) cause: reason: effect ^ cause	<i>It is amazing [how effective this system is], for the tower stays as stiff as a ram-rod even in the most blustery conditions.</i>
	condition: positive	<i>That would save a fortune and then we'd have the cash [that we need to, you know, go on to the next step].</i>
	condition: negative	<i>This is very much essential, otherwise a lot of time is usually wasted for sighting the staff.</i>
	condition: concessive – concession ^ consequence	<i>Through mounting irritation I kept telling him that I needed a cure for my son and nothing for myself; still I answered his questions with all the politeness I could muster.</i>
	condition: concessive – consequence ^ concession	<i>I was an English major, but I took courses in biology and ornithology.</i>

Table 5 Paratactic Enhancement Categories (Halliday, 2014, pp. 478-480).

Hypotactic enhancement is the notion of $\alpha \times \beta$ and appears in traditional adverbial clauses. In the same way as parataxis, these forms take shape in clauses relating to time, place, manner, cause and condition. The normal construction of hypotactically enhancing chains is limited to two clauses that take place with one clause qualifying another clause (49). Hypotactic chains of more than two clauses are not uncommon. The subtype of hypotactic enhancing clauses typically changes as the clause chain is developed. Different from paratactic chains, hypotactic ones move further away from the place in the discourse where the dominant chain is located (Halliday, 2014, p. 481).

- (49) ||| [1α:] *I'd parted with the Zen master [I was working with originally], || [1¥β:] as had most of his senior students, || [¥2:] so I was without a teacher. |||*

Hypotactically enhancing clauses may take form as being finite or non-finite. Finite clauses are typically introduced by a binder that is a subordinating conjunction (Table 6). Non-finite clauses are introduced by prepositions like *on*, *with* and *by* that function conjunctively. It is important to note that the same word can be both a conjunction and a conjunctive preposition as with *before* and *after*. Non-finite hypotactically enhancing clauses can function with a subset of binders as with *when* (Halliday, 2014, p. 482).

	Category	Examples
time	-	<i>Moomintroll, that chubby, cheerful being, came into existence as a family joke when Tove Jansson was a young girl</i> .
place	concrete place	<i>The lbo never accept anything [which is rigid and final and absolute]: wherever one thing stands, another thing will stand beside it.</i>
	abstract place	<i>As a result, disagreement is carried out in the absence of an audience, where ideological and performance changes may be made without the threat of damage to the goals of the team, as well as the character of the individual.</i>
	abstract place shading into matter	<i>As far as it can, the Zoo tries to be self-supporting, and you will notice the names of companies and individuals on many of the cages [who sponsor the animals].</i>
manner	quality	<i>As it happens, Margo was an extremely rich woman.</i> <<by chance>>
	comparison	<i>He just shakes his head and shoves it at her again and says 'Give Massin,' as if he knew there'd be no problem at all.</i>
	means	<i>These theories include the solar theory, whereby periodically the amount of nitrogen compounds is enhanced.</i>
cause-condition	cause: reason	<i>Gradually, they outgrow their baby shoes – if the expression is pardoned, as Snufkin is in fact the only one of them [who uses footwear at all].</i>
	cause: purpose	<i>Everyone at VES is working hard to change the law so that we will have voluntary euthanasia legalized in England within the next five years.</i>
	cause: result	<i>After that, the ozone hole developed rapidly, especially after September 5, so that by October 5, the ozone over the middle of Antarctica had dropped from 320 Dobson units (DU) to 120 DU.</i>
	concession	<i>Even though it was a somewhat silly book about the grand passions of college students, it really was a novel.</i>
	condition: positive	<i>If I had a different view, then perhaps I would write more novels.</i>
	condition: negative	<i>You will cherish them on your bookshelves for a long time – unless, of course, someone borrows them and somehow 'forgets' to return them.</i>

Table 6 Finite Hypotactic Enhancing Clauses (Halliday, 2014, pp. 481-485).

In many hypotactically enhancing clauses, it should be noted the enhancing relation can be internal instead of external. This takes form in the β -clause relating to the enactment of a proposition or proposal realized by the α -clause rather than the figure it represents (50). (Halliday, 2014, p. 484).

(50) *If it is not too personal an inquiry, what limits do you set.*

In finite clauses the conjunction serves both the purpose being the hypotactic status and the enhancing relationship (Table 6). Additionally with simple conjunctions (*because*, *when* and *if*) as well as conjunctive groups (*if*, *even if*, *soon after* and *so that*), there are three kinds of complex conjunctions. Complex conjunctions can be derived from verbs, nouns and adverbs (Halliday, 2014, p. 484).

Verbal constructions are acquired from the imperative, present/active participle or past/passive participle + (optional element) as with *provided (that)*, *seeing (that/how)*, *suppose/supposing (that)*, *granted (that)* or *say (that)*. In their origin, these are projections in which their function as expanding conjunctions reflects the semantic overlap between expansion and projection in the realm of irreal in the sense that *let's say/think that* equals *if* in (51) (Halliday, 2014, p. 484).

(51) *Say that they can't mend it, shall I just throw it away?*

Nominal conjunctions take the form of *in case*, *in the event that*, *to the extent that* and *the* + various nouns of time and manner such as *the day*, *the moment*, *the way*, etc. Nominal conjunctions have evolved from prepositional phrases with the enhancing clause embedded in them like in *on the day we arrived*. However, they now function to introduce a hypotactic clause just like other conjunctions as in (52). Adverbial conjunctions are *as/so long as*, *as/so far as* and *as much as*. These clauses express limitation up to a particular point to which a certain circumstance is valid (Halliday, 2014, p. 484).

(52) *Their daughter was born the day we arrived, **the way** they're working now the job'll be finished in a week.*

Non-finite hypotactic enhancing clauses can be realized in time, concession, condition and manner through means when it is explicitly marked by a structural conjunction such as *when, while, if, although* and *though* or a conjunctive preposition such as *before, after, since, because of, without* or *by*. The enhancing relationship can also be left implicit in time, cause: reason, cause: purpose, cause: result. Table 7 shows examples of the above categories of non-finite hypotactic enhancing clauses (Halliday, 2014, pp. 484-485).

Logico-semantic Category	Example
time (explicit)	<i>Follow the pathways around the landscaped gardens and over bridges before resting at the Tea House where the scent of lotus flowers mingles with that of freshly brewed tea and traditional cakes.</i>
concession (explicit)	<i>Similarly Mr. G. S Sawhney, largely due to the recommendation of Mr. K. K. Shah, then Governor of Tamil Nadu, was transferred from Collector of Customs, Bombay, to become Director of Revenue Intelligence, despite having himself been under investigation by the CBI and having been listed as a suspect in the Directorate of Revenue Intelligence.</i>
manner: means (explicit)	<i>Bacteria can also aid chemical precipitation of calcite by making the water more alkaline.</i>
time (implicit)	<i>Catch a ride on the monorail to the ritzy shopping centre of Sydney, taking in the Queen Victoria Building and Centrepont on the way.</i>
cause: reason (implicit)	<i>This view was not empirically based, having arisen from an a priori philosophy.</i>
cause: purpose (implicit)	<i>To jazz up the title, use the mouse to click on the text and type something new.</i>
cause: result (implicit)	<i>He was taken away from the city, never to be seen again.</i>

Table 7 Non-finite Hypotactic Enhancing Clauses (Halliday, 2014, pp. 484-486).

In enhancing clauses, the non-finite dependent clause without a subject is interpreted by reference to the subject of the dominant clause. Examples can be found where the subject of the dependent clause is not co-referential with the subject of the dominant clause as in cases when the dependent subject typically refers to the speaker (53). The subject can additionally refer to the agent in a receptive clause whether it is structurally present or not (54) (Halliday, 2014, p. 486).

(53) ||| *But, of course, **having said that**, the hope is that at least now we know.* |||

(54) ||| *If this occurs in limestone, || beautifully preserved fossils with delicate features intact can be recovered || **by** dissolving the limestone with acid.* |||

If the dependent clause appears as non-finite, the resulting circumstantial relationship is made explicit by the structural conjunction or conjunctive preposition. These conjunctions are a subset of the ones occurring in finite clauses. Therefore, the meaning is essentially the same. In this case, the prepositions tend to be less specific as with *in turning the corner* and *on thinking it over* and the meaning of the clause introduced by the preposition may vary according to the sense of the primary clause. Hypotactic enhancement takes place through cause: reason in (55) in the sense of the clause being *because I wasn't there* (Halliday, 2014, p. 486).

(55) ||| **Without having been there** || I can't say what happened. |||

Expansion clauses that are not explicitly marked for any logico-semantic relation contain certain markers of expansion that are multivalent. These markers can distinguish between elaboration and extension, or extension and enhancement. There are three different examples using *but* incorporating adversative (56), replacive (57) and concessive (58) meanings. The concessive example embodies a logical opposition between two terms thus being an agnate hypotactic nexus. Table 8 gives examples of conjunctive markers with two or more meanings (Halliday, 2014, p. 487).

(56) *They're pretty, **but** I can't grow them.* (on the other hand)

(57) *Don't drown them, **but** give them just enough.* (instead)

(58) *I don't look after them, **but** they still grow.* (nevertheless)

Marker	Elaboration	Extension	Enhancement
<i>and</i>		additive: ' <i>and also</i> '	temporal: ' <i>and then</i> ' causal: ' <i>and so</i> '
<i>but</i>		adversative: ' <i>on the other hand</i> ' replacive: ' <i>instead</i> '	concessive: ' <i>nevertheless</i> '
<i>yet</i>		adversative: ' <i>on the other hand</i> '	concessive: ' <i>nevertheless</i> '
<i>or</i>	exposition: ' <i>or rather</i> '	alternative: ' <i>or instead</i> '	

<i>while</i>	additive: 'and also` adversative: 'and yet`	temporal: same time: spread: 'and meanwhile` concessive: 'nevertheless`
<i>as</i>		temporal: same time: spread: 'when` causal reason: 'because`
<i>since</i>		temporal: different time: later: 'after` causal: reason: 'because`
<i>if</i>	alternative (<i>if ... not</i> [<i>then</i>]) 'or`	conditional: positive: 'in case`

Table 8 Conjunctive markers used for more than one type of expansion (Halliday, 2014, p. 487).

2.2 Complexity Measures

Research in second language acquisition has been concerned with finding objective ways to describe language development and assess linguistic ability (Larsen-Freeman, 2006). Linguistic complexity research came into existence in the 1970s when L2 researchers started using metrics of grammatical complexity and accuracy developed by L1 acquisition research. The mid-90s brought about a proficiency model introducing three separate dimensions which included complexity, accuracy and fluency (CAF) as the three main constructs. CAF addresses two general questions at the heart of many studies in second language acquisition (SLA) and applied linguistics: What makes a second language learner a proficient language user? And how can L2 proficiency be more adequately (i.e. validly, reliably and feasibly) measured? (Housen et al., 2012).

The purpose of using CAF measures is to establish an empirical framework on how L2 develops by documenting what parts of the interlanguage system change as acquisition takes place, in what ways anticipated change proceeds and why at times not much change seems to take place. Scholars and researchers who work in instructed SLA use CAF for measuring how and why language competencies develop for certain learners and target languages. Moreover, they use measures to gauge the response to tasks, teaching and other stimuli which are then compared to the details of developmental rate, route and ultimate outcome (Norris & Ortega, 2009).

The three-fold combination of complexity, fluency and accuracy has for some time been recognized by researchers as a principal dimension for gauging L2 writing quality. Accuracy tends to be the oldest, most transparent and consistent construct of CAF. Accuracy refers to deviances from a particular norm as the ability to produce target-like and error free language. Fluency is described as the ability to produce the L2 with native-like rapidity, pausing, hesitation, or reformulation (Bulté & Housen, 2012; Housen & Kuiken, 2009; Housen et al., 2012; Norris & Ortega, 2009; Pallotti, 2009; Wolf-Quintero et al., 1998).

Linguistic complexity has been the measure which has clearly received the most attention in L2 writing research (Lahuerta Martínez, 2018a). Housen and Kuiken (2009) denote the term complexity as the most complex, ambiguous and least understood dimension of the CAF triad (p. 463). Linguistic complexity is explained as the ability to use a wide and varied range of sophisticated structures and vocabulary in L2 (Housen et al., 2012, p. 2). De Clercq and Housen (2017) describe linguistic complexity in general terms as an absolute, essentially quantitative property of a linguistic unit or system, determined by the number of its components and types of connections between those components.

Pallotti (2015) separates the many definitions of linguistic complexity into three specific groups including morphological, syntactic and lexical. Morphological complexity is defined as measuring the number of exponents. Syntactic complexity measures the number of phrases per clause, number of clauses per unit or number of word-order patterns. Lexical complexity is measured by lexical diversity which looks at type/token ratios. It has commonly been pointed out by many studies that syntactic complexity is arguably the most frequent and intensively used component in linguistic complexity in SLA research (Kuiken et al., 2019).

Complexity has been used as a term to refer to task, cognitive or linguistic complexity in SLA literature. Linguistic complexity may be considered as a dynamic property of a learners' interlanguage system overall and as a more stable property of the individual linguistic elements that make up the interlanguage system (Housen & Kuiken, 2009). Among the subconstructs of linguistic complexity, syntactic complexity has been characterized as the range of syntactic structures that are produced and the degree of sophistication of those structures (Lu & Ai, 2015; Ortega, 2003; Pallotti, 2015).

Linguistic complexity can be investigated at the level of the language system as a whole or taking into accounts its major subsystems including individual linguistic features such as forms, structures, patterns and rules. The complexity of these structures can in turn be studied from formal and functional dimensions. All these different components and subdimensions of complexity can be studied across various domains of language such as lexicon, syntax, and morphology. The importance of syntactic complexity in second language writing research and pedagogy has long been recognized in a large number of studies examining the relationship of syntactic complexity of L2 writing to L2 proficiency (Ai & Lu, 2013; Lu, 2011; Norrby & Håkansson, 2007; Ortega, 2000, 2003; Wolfe-Quintero et al., 1998).

2.2.1 Syntactic Complexity

Syntactic complexity is understood broadly as the range and sophistication of grammatical resources exhibited in language production (Ortega, 2015). It has meant sorting through data and calculating the average length of a syntactic unit of production, computing the density of subordination and counting the frequency of occurrence of selected forms that are linguistically sophisticated. A reason often cited for measuring complexity in SLA is to be able to benchmark developmental levels (Ortega, 2012).

A wide range of studies have used a variety of ways to approach subordination as an acquisitional target which has its own developmental characteristics and sequence. Subordination has been approached as a descriptor of learner language to gauge proficiency, describe performance and benchmark development within CAF. CAF is often applied in task-based SLA research relying on subordination ratios to measure L2 complexity as valid indicators of L2 proficiency to show linear developmental trajectory (Chen et al., 2021, p. 811).

Syntactic complexity has been measured using a wide range of indices since its initial conception (Bulté & Housen, 2012; Lu, 2011; Norris & Ortega, 2009; Ortega, 2003; Wolfe-Quintero et al., 1998). Many studies have applied a small number of measures to relatively small amounts of data which is problematic for two reasons. First, the relationship of different syntactic complexity measures to L2 proficiency vary substantially thus making it difficult to generalize results pertaining to any specific measure to the general construct of

syntactic complexity. Second, syntactic complexity is thought of as multi-dimensional, with each dimensional requiring one or more different measures appropriate for that dimension (Lu & Ai, 2015, p.19).

A number of cross-sectional studies have delved into the extent of which different syntactic complexity measures reliably index a L2 writers' global proficiency (Ai & Lu, 2013; Bardovi-Harlig & Bofman, 1989; Ferris, 1994; Lu, 2011; Sotillo, 2000). There have been longitudinal studies that tracked and compared learner development over time (Casanave, 1994; Ishikawa, 1995; Mazgutova & Kormos, 2015; Norrby & Håkansson, 2007; Ortega, 2000). The relationship between syntactic complexity to L2 proficiency or L2 writing quality may vary for different measures or dimensions, and development of syntactic complexity may also vary among learners (Lu & Ai, 2015, p. 18).

The most common syntactic complexity measures are: (a) length which is calculated by dividing words, such as multiclausal units, by a chosen production unit, (b) subordination which is computed by counting all clauses and dividing them over a given production unit of choice thus yielding mean number of clauses per T-unit. T-units are defined as main clause plus any subordinate or embedded clauses that may occur in it, and the amount of subordination as measured by mean number of finite clauses produced. On average, typical values found across 40 studies suggest that intermediate second language writers can produce T-units that average 12 words and contain a subordinate clause in every other main clause in written texts (Wolfe-Quintero et al., 1998; Norris & Ortega, 2009; Ortega, 2003).

The mean length of a clause ought to be most predictive measure at an advanced level in development, when processes of grammatical metaphor begin to unfold, and more synoptic styles emerge in the repertoires of high-proficiency L2 learners and users. It is suggested that mean length be used as a measure for fluency and not complexity. However, in L2 writing, T-unit mean length has seemed to be the most used complexity measure. A salient observation made across 16 studies pointed to a trend in which researchers use frequency counts of selected forms to measure complexity as structural variety and sophistication (Norris & Ortega, 2009).

Bulté and Housen (2012) and Norris and Ortega (2009) compiled a summary of complexity measures for subordination by various studies (Table 9). The authors highlight that various measures are not created equal and some of them redundantly measure the same

thing. The subordination indices listed all feature clauses (subordinate or dependent) in numerator. Norris and Ortega (2009) explain:

Subordination metrics are all equivalent, regardless of the denominator choice, in that they all use complexification as a phenomenon of subordination exclusively. The only way values for this group of measures would increase is when more subordinate or independent clauses are produced. Mathematically speaking, the measure of mean number of dependent or subordinate clauses per total clauses is a faithful replication of mean number of total clauses per multi-clausal unit, only that the ratio is calculated at a level of analysis down the syntactic continuum from dependent/subordinate clause unit to the clause and then the multi-clause unit. (p. 560)

Central Focus of Calculation	Measures
Length (in words, morphemes, characters, etc.)	Mean length of utterance
	Mean length of T-units
	Mean length of C-units
	Mean length of clause
Amount of subordination	Mean number of clauses per T-unit
	Mean number of clauses per C-unit
	Mean number of clauses per AS-unit
	Mean number of dependent or subordinate clauses per total clauses

Table 9 Bulté and Housen (2012) and Norris and Ortega (2009) Summary Syntactic Complexity Measures.

A subordination measure used for a study should depend on which unit of discourse segmentation is more appropriate for the data at hand. Therefore, the utterance or the analysis of speech unit (AS-unit) may be more appropriate for dialogic oral data since it may contain many non-syntactic segments. In contrast, the T-unit may be ideal for intermediate to advanced written data. Independent clauses with modifiers (C-unit) might be considered more appropriate for data which may include many non-syntactic segments like that produced by low-proficiency learners (Norris & Ortega, 2009, p. 260).

Empirical CAF research has taken a rather narrow, reductionist, perhaps even simplistic view on and approach to what constitutes L2 complexity (Bulté & Housen, 2012). A call for attention has been directed at devising measures that include a wide range of

developmentally ranked structures, regardless of their status as target-like or nontarget-like, to help researchers characterize L2 production that ranges along the full developmental continuum. At the early stages of L2 development, syntactic complexity is first established through coordination. Only at later intermediate stages is when subordination becomes the dominant means of syntactic complexity. At even more advanced stages of L2 development, syntactic complexity would be mainly achieved through increasing complexity at the phrasal level. In addition to measures of subordination, research should also include measures of coordination and phrasal complexity (Norris & Ortega, 2009, p. 567).

Bulté and Housen (2012) and Norris and Ortega (2009) provide an adequate initial reference for complexity measures in Table 9. However, Lu and Ai (2015) recognized the importance of measuring syntactic complexity as a multidimensional construct and introduced a more comprehensive set of indices in Table 10 which include 14 measures used by previous studies (Ai & Lu, 2013; Lu, 2010, 2011; Ortega, 2003, 2009; Wolfe-Quintero et al., 1998).

Measure	Code	Definition
<u>Length of production unit</u>		
Mean length of clause	MLC	# of words/# of clauses
Mean length of sentence	MLS	# of words/# of sentences
Mean length of T-unit	MLT	# of words/# of T-units
<u>Amount of subordination</u>		
Clauses per T-unit	C/T	# of clauses/# of T-unit
Complex T-units per T-unit	CT/T	# of complex T-units/# of T-unit
Dependent clauses per T-unit	DC/C	# of dependent clauses/# of T-clauses
<u>Amount of Coordination</u>		
Coordinate phrase per clause	CP/C	# of coordinate phrases/# of clauses
Coordinate phrase per T-unit	CP/T	# of coordinate phrases/# of T-units
T-units per sentence	T/S	# of T-units/# of sentences

Degree of Phrasal sophistication

Complex nominal per clause	CN/C	# of complex nominals/# of clauses
Complex nominal per T-unit	CN/T	# of complex nominals/# of T-unit
Verb phrases per T-unit	VP/T	# of verb phrases/# of T-units

Overall sentence complexity

Clauses per sentence	C/S	# of clauses/# of sentences
----------------------	-----	-----------------------------

Table 10 Syntactic Complexity Measures Lu and Ai (2015).

Table 10 uses a large set of measures for several considerations which include recommendations from the study conducted by Wolfe-Quintero et al. (1998) in its early review of L2 writing measures. Table 10 was compiled with the awareness that some measure sets are stronger indicators of L2 proficiency while other measures might be partially redundant of each other (Lu, 2011; Ortega, 2003; Norris & Ortega, 2009). Lu and Ai (2015) acknowledge the field as a whole is still in search of the best set of non-overlapping measures that are consistently indicative of L2 proficiency or L2 writing quality and capture all major dimensions of syntactic complexity (p. 19).

Bulté and Housen (2018) coined the developmental sequence of syntactic interlanguage complexification hypothesis (DSSICH) which refers to the proposal given by Norris and Ortega (2009) as theoretical justification to measure multidimensional syntactic complexity. Based on Halliday and Matthiessen (2006), the DSSICH proposes that language development proceeds from: (i) the expression of ideas through mostly parataxis or sequencing self-standing words, sentences and clauses; through (ii) expansion by hypotaxis which expresses the logical expression of ideas through grammatically intricate texts; lastly (iii) the emergence of and reliance on grammatical metaphor which leads to advanced language that exhibits lower levels of subordination but higher levels of lexical density with more complex phrases instead of clauses.

With the interest in complexity measures, dissatisfaction has appeared with measurement approaches which typically assess L2 writing through quantitative measures of lexical complexity that include lexical density, syntactic complexity and subordination. Findings have been inconsistent in several studies using these factors (Ryshina-Pankova, 2015). Norris and Ortega (2009) as well as Ortega (2012) indicate that an increase in subordination and length of T-units highlight some levels of language development, yet they

do not capture the entire span. Bulté and Housen (2014) shed doubt on if subordination is adequate to gauge L2 complexity in all circumstances. Pallotti (2009) incidentally underlines that complexity is looked at as the ability to produce certain kinds of language forms which lack the consideration of meaningful production content and realization of communicative goals.

Bulté and Housen (2018) conducted an exploratory study that investigated longitudinal development by analyzing a substantive number of English L2 writing samples. Their analysis involved students around 11 years old who formed a group of 10 Dutch English as a foreign language learners (EFL). The study took place over a 19-month period using a judicious selection of quantitative syntactic complexity measures. Their results pointed out that selective units of the complexity used by beginner L2 writers increased over time (i.e. length of T-unit and clausal subordination). An increase of L2 complexity at a group level was linear, but at the level of an individual learner, there was a factor of high variability. The study also showed that a complex interplay exists between different complexity measures, and possibly between complexity components that are meant to quantify.

In one of the more recent studies taking a different approach to analyzing syntactic complexity was used by Zhang and Lu (2022) in which they investigated to what extent traditional versus fine-grained syntactic complexity indices could predict second language raters' quality ratings. The study analyzed two genres of writing including 581 application letters and 595 argumentative essays produced by college-level Chinese EFL students. Table 11 lists the fine-grained indices used in the study in addition to traditional indices. Their results were consistent with previous findings in which fine-grained indices had a stronger predictive power than traditional indices. Findings also highlighted useful insights into the effect of genre on the explanatory power of traditional and fine-grained indices as well as which indices significantly predict L2 writing quality.

A need has been expressed for an alternative approach to L2 writing that acknowledges the connection between language complexity and its discourse-semantic function of constructing certain types of meaning in particular texts. The notion that has been missing from traditional complexity measures and from existing taxonomies is the dimension of meaning in complexity. SFL characterizes linguistic complexity in development and

communicative appropriateness. The central idea is that language use and development are motivated by meaning-based, contextual and communicative demands of the tasks in which learners actively respond. Arguments have been made that contextual aspects are at the heart of a meaning-based approach and should be added to the measures of complexity in L2 writing. Furthermore, contextual aspects in a meaning-based approach might explain variations found in syntactical and lexical complexity found in L2 production (Ryshina-Pankova, 2015, p. 52).

Index	Label	Description
<u>Clausal complexity</u>	acomp	An adjective functioning as a complement in a copular clause.
Adjective complements		
Adverbial clauses	advcl	A clause modifying a verb phrase.
<u>Phrasal complexity</u>		
Nominal subject	nsubj	A subject of a (nonpassive) clause that is a noun phrase.
Nominal complement	ncomp	A noun or a noun phrase that functions as a complement in a copular clause.
<u>Syntactic sophistication</u>		
Average main verb lemma frequency of a VAC	acad_av_lemma_freq	Average frequency score of all main verb lemmas of a VAC in COCA.
Average delta p score verb (cue) – construction (outcome)	fic_av_delta_p_verb_cue	Average delta P score of all verb-VAC combinations in COCA.

Table 11 Sample Fine-grained Syntactic Complexity Indices Zhang and Lu (2022).

The advancement of automated complexity and natural language processing tools has led to the growth of research in fine-grained complexity measures. The Coh-Metrix Analyzer (Grasser et al., 2004), L2 Syntactical Complexity Analyzer (Lu, 2010), the Lexical Complexity Analyzer (Lu, 2012), Tool for Automatic Analysis of Syntactic Sophistication and Complexity (Kyle, 2016) and the Tool for Automatic Analysis of Lexical Sophistication (Kyle et al., 2018) are a number of automated resources used to analyze syntactic complexity in texts. These tools incorporate a wide variety of overall and fine-grained measures which target specific and linguistically sophisticated measures. However, the overall issue facing

these automated tools is to what extent and which measures are redundant, how valid and reliable they are, and until what point they can function as an index in L2 development (Kuiken, 2023, p. 86).

2.2.2 Syntactic Complexity Viewed Through the CEFR

This section will first cover a brief description of The *Common European Framework of Reference for Languages* (2001) and the six different proficiency levels which are crucial for this study. Studies using CEFR levels as part of their research are then reviewed to provide a landscape of the current relevant research being conducted in syntactic complexity through levels A1 to C2.

2.2.2.1 What is the CEFR?

Established in 2001, the CEFR created a common basis for the elaboration of language syllabuses, curriculum guidelines, examination and textbooks across the European continent (Council of Europe, 2001; Deygers et al., 2018a). The design of the framework is interwoven with the idea of encouraging free movement of people and ideas by increasing transparency across educational systems through the common use of the same proficiency levels (Deygers et al., 2018b).

With companion volume updates in 2018, the documents provide solutions for theorists, researchers and language professionals to describe what it means to learn a language. The framework addresses the need to operationalize and sequence languages for learning, teaching and assessment purposes to give them a real-life approach that finds common ground in levels of attainment (Figueras, 2012, pp. 477-478).

The overall language proficiency descriptive scheme considers general competences, communicative language competences, communicative language activities and communicative language strategies (Council of Europe, 2001). The foundation is a reference tool rather than an instrument to be applied. Institutions that use the CEFR should merge the activities, competences and proficiency scales that are appropriate for their level on a local basis (North, 2007, p. 656). The fact that each member state needs to specify content on a

local basis has been seen as one limitation of the CEFR since the proficiency scheme and level descriptors are not language-specific (Little, 2007, p. 649).

The CEFR defines reception, production, interaction and mediation as the four language and learning activities that replace the traditional model containing the four traditional skills which are listening, speaking, reading and writing. The CEFR pioneered mediation as the fourth mode of communication alongside the other three (North & Piccardo, 2016, p. 455). The goal of defining four new groups is to capture the complex reality of communication (Council of Europe, 2018; North, 2007, p. 658).

The *Common European Framework of Reference for Languages: Learning, Teaching, Assessment Companion Volume with New Descriptors* (2018) is the second publication in the series and carries on with the original message of the CEFR as promotion of positive formulation of educational aims and outcomes at all levels through an action-oriented approach in which language is learned for a social purpose. An aspect brought out in the original document and the companion guides is the insistence of “Can do” descriptors in the proficiency scale focusing on what learners have not yet acquired rather than linguistic deficiencies thus putting forth a positive outlook for the learner (Alderson, 2007, p. 660; Figueras, 2007, p. 673; Little, 2007, p. 646; North, 2007, p. 656). The significance of descriptors is that they show how the quantitative element of what a student can do is interwoven with the qualitative dimension of how well the learner is able to do it (Hulstijn, 2007, p. 663).

One of the most significant contributions of the CEFR is the proficiency scale which includes six main levels (Figure 1) describing in general terms how language is used through examples of activities and tasks (Figueras, 2012, p. 478). An important aspect and often criticized characteristic of the proficiency scale is the fact that it is not language specific (Little, 2007, p. 646).

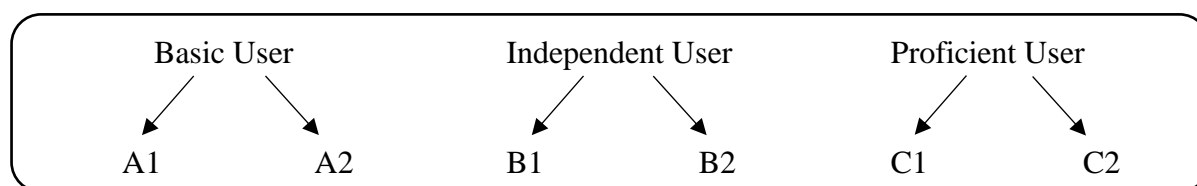


Figure 1 CEFR levels (Figueras, 2012, p. 478).

The CEFR profile proficiency level illustrative descriptors are another significant achievement described by North (2007) as “An item bank of empirically calibrated descriptors with mathematical values on a common scale and known statistical values” (p. 657).

The most recent update *Common European Framework of Reference for Languages: Learning, Teaching, Assessment Companion Volume* (2018) explains the necessity of including profile descriptors for each level as a necessary simplification. Levels are used to organize learning, track progress and answer questions like “How good is your French?” or “What proficiency should we require from candidates?” However, any simple answer like B2 or even B2 receptive, B1 productive – hides a complex profile. The reason the CEFR includes so many descriptor scales is to encourage users to develop differentiated profiles. Descriptor scales can be used firstly to identify which language activities are relevant for a particular group of learners and, secondly, to establish which level those learners need to achieve in those activities to accomplish their goals (p. 38).

2.2.2.2 Syntactic Complexity Studies Factoring in EFL Writing Across CEFR Levels

An area of research that appears to have received less attention from researchers is the investigation of syntactic complexity using CEFR proficiency levels in EFL. The framework has become increasingly important in Europe since it describes scales defining what language learners can do in an L2 at different stages of proficiency (Dygers et al., 2017; Hulstijn et al., 2010). These levels are general descriptions of learning stages in L2 development, yet it does not factor in syntactic complexity (Khushik & Huhta, 2020).

Kim (2004) conducted an initial study with CEFR-rated scripts from 33 Chinese EFL students which used clauses and T-units as the basic unit of analysis. The investigation investigated three aspects of the subordinate clauses: (i) the variety of structures (adverbial, adjective and nominal clause per clause), (ii) number of subordinate clauses (clauses and dependent clauses per T-unit, dependent clause per clause), and (iii) shift from clauses to phrase (prepositional, participial, gerund, participial and infinitive phrases per clause). The results showed a clear difference between A2 and B2 levels in all the measures except for nominal and gerund phrases per clause. The differences between the A2 and B1 levels were

not striking, yet more pronounced than the differences between B1 and B2. The drawback of this study was that it investigated a small group of learners.

Sentence length increased significantly between adjacent levels from A2 to C2 as reported from a study that used the *English Grammar Profile* to gauge learners' performances in language test tasks taken from the large-scale Cambridge learner corpus (Hawkins & Filipovic, 2012). Green (2012) also used the *English Grammar Profile* which detailed a significant difference between noun phrase incidents and the number of modifiers per noun between B2 and C1.

Verspoor et al. (2012) undertook a large-scale study of 489 Dutch EFL learners aged 12-15 comprised of 40,000 words which used 64 variables. Learners wrote one descriptive text on topics varying on grade level. The texts were graded on a 5-point scale corresponding to CEFR levels A1.1, A1.2, A2, B1.1 and B1.2. The results indicated that T-unit mean length increased across levels and drastically differentiates in A1.2 versus B1.1 and A2 versus B2.2. It reported the proportion of the ratio of simple versus complex sentences to be a good separator for levels with the clearest connection taking place between A1.2 and A2. The study further sheds light on how the proportion of dependent clauses can distinguish proficiency levels. Furthermore, finite relative clauses saw a steady increase across levels and most significantly between A2 and B1.1.

Gyllstad et al. (2014) used CEFR levels A1 to C2 to examine emails and stories written by 120 Swedish participants writing in L2 English, L3 French and L4 Italian. Three complexity indices, which included T-units, mean length of clauses and clauses per T-unit, were used to analyze 235 texts. Their data set ranged from A1 to B2, with not many texts at the B2 level and zero at C1 and C2. Results showed proficiency levels can be characterized in linguistic profiles based on syntactic complexity. Beginner levels in all three languages did not exhibit significant differences thus suggesting their linguistic profile at A1 and A2 is rather homogeneous. On the other hand, both English and French showed differences between languages in mean length of T-unit and mean number of subordinate clauses. This was not seen in Italian due to a lower representative sample. Data from this study suggests, that from intermediate levels and above, learners have developed a rich enough linguistic repertoire in order for differences to emerge and for language-specific properties to take shape. The small representative sample at the upper intermediate level and the lack of data at

advanced level together with few subordinate clause indices over very broad CEFR scale categories were recognized as drawbacks in this study.

Alexopoulou et al. (2015) incorporated a large and diverse database containing 1,180,543 scripts from 174,771 learners across CEFR levels ranging from A1-C2. Their study investigated subordinate clause indices in EFL texts using the first version of the *Education First-Cambridge Open Language Database* open access corpus. Their results found sentence length to increase across all CEFR levels. They additionally reported a distinct increase in subclausal density from A2 to B2. However, it was not clear if these changes were statistically significant. This study examined every CEFR level by using a large data set, yet a drawback is that it only included three subordinate clause indices (Khushik & Huhta, 2020).

Lahuerta Martínez (2018a) analyzed opinion essays which were graded using CEFR levels A1 and C2 from 188 intermediate secondary education Spanish EFL learners. The study highlighted a significant link between syntactic complexity and writing quality. The results underscored that sentence length, compound/complex sentences ratios, coordinate/dependent clause ratios and noun phrase per clause differed between levels. The study found longer units at a clausal and sentential level as well as coordination and subordination to be an indicator of higher writing quality. There was a significant increase in coordination and subordination at higher grades which lined up with results obtained by Lorenzo and Rodríguez (2014). An important aspect of this study was that it was reasonably large scale, and all the participants completed the same task under the same conditions.

Chen et al. (2021) used the large data set in the second version of the *Education First-Cambridge Open Language Database* (EFCAMDAT2) to draw from 31,040 scripts across 16 proficiency levels equivalent to the six CEFR levels. The AutoSubClause dependency parser was used to extract subordinate clauses from the texts from L1 Brazilian, Chinese, Russian and Japanese learners writing in L2 English. The parser was able to extract adverbial, complement and relative clauses as well as the level of embeddedness of the clause.

The breakdown of clause numbers in the Chen et al. (2021) study for all four L1s is shown in Table 12 and normalized by 1k. What was additionally included by the current study is the overall CEFR clause total as well as separate totals for adverbial, complement and relative clauses per each CEFR level. The total clauses in each CEFR level appeared to

increase from A1 to C1 with a dramatic decrease at the C2 level. Complement and relative clauses followed the same trend while adverbial clauses increased from A1 (601) to B2 (5,712) then saw a minor decrease at C1 (5,557). Their findings appear to indicate that overall subordination in the four L1s increases as proficiency increases through the C1 level.

EF Teaching Levels	CEFR Levels	Adverbial Clauses	AC Total	Complement Clauses	CC Total	Relative Clauses	RC Total	Total Clauses	CEFR Total Clauses
1		99		328		364		791	
2	A1	270	601	639	1,774	532	1,350	1,441	3,725
3		232		807		454		1,493	
4		978		964		635		2,577	
5	A2	875	2,790	1,386	4,564	789	2,482	3,050	9,836
6		937		2,214		1,058		4,209	
7		1,057		2,053		1,071		4,181	
8	B1	1,416	4,328	2,159	6,405	1,103	3,307	4,678	14,040
9		1,855		2,193		1,133		5,181	
10		1,521		2,560		1,294		5,375	
11	B2	2,009	5,712	3,113	8,817	1,812	4,617	6,934	19,146
12		2,182		3,144		1,511		6,837	
13		1,625		2,843		1,764		6,232	
14	C1	2,130	5,557	3,064	8,908	1,839	5,480	7,033	19,945
15		1,802		3,001		1,877		6,680	
16	C2	1,458	1,458	2,205	2,205	1,867	1,867	5,530	5,530
Total		20,446		32,676		19,103		72,222	

Table 12 EFCAMDAT2 Clause Type Per CEFR Proficiency Level (Chen et al., 2021, p. 812).

Chen et al. (2021) considered the mean number of subordinate clauses and compared the developmental trajectory of each L1. The overall trend indicated adverbial subordination reached a plateau around the B2 level which were consistent with findings from Alexopoulou et al. (2015) and Bardovi-Harlig and Bofman (1989). However, the current study noted there was only a difference of 155 adverbial clauses between B2 (5,712) and C1 (5,557). The L1 use of different types of subordinate clauses showed uneven spread across all L1s with the most frequently used clause type being adverbial and complement. The general finding pointed out there are L1 and clause type effects on the L2. The shortfall of this study was the few indices used in the analysis and the unclear end result of whether the study was used as a test for the AutoSubClass parser or further research on subordination. Another limitation of

the Chen et al. (2021) analysis was how formulaic sequencing might have affected learning tasks or contribute to subordinate clause production.

Chen et al. (2021) additionally noted that few studies have analyzed the fine-grained aspects of subordinate clauses like the use of subordinators, the level of embeddedness, the semantic function of adverbial clause, the role of the head noun in the relative clause in main as well as subordinate clauses. There appears to be an area of research in which more information is needed in general, and particular to this study, using fine-grained measures for Spanish L2 and English L2. Furthermore, a future line of future research would be to investigate what factors contributed to what Chen et al. (2021) described as formulaic sequencing since the study lacked detail on this matter.

In a study carried out by Khushik and Huhta (2020), two groups of EFL learners with different L1s (Sindhi and Finnish) were examined to see whether CEFR levels A1, A2 and B1 could be distinguished by 28 syntactic complexity measures and if results differ between the two control groups in a written argumentative essay task. Their study recognized the lack of consensus on the definition of complexity apart from acknowledging that it is complex and comprises many levels and dimensions. Kaushik and Huhta's (2020) findings were consistent with several studies in the sense they encountered sentence length to separate all CEFR levels between A2 and C2 (Hawkins & Filipovic, 2012), subordination was seen to increase with proficiency in Finnish (Wolfe-Quintero et al., 1998) and subordination distinguishes the lower proficiency level of A2 between the higher B2 level (Kim, 2004). Coordination indices failed to separate CEFR levels even when their values increased between A1 and A2. However, T-units per sentence were a good indicator in Finnish between A1-A2.

In relation to the similarities and differences between L1 groups, regarding syntax complexity and CEFR level descriptors, Khushik and Huhta (2020) determined A1 to be the most comparable in EFL writing. Beyond A1, syntax complexity measures differed substantially between L1 Sindhi and Finnish learners. Sindhi learners used more coordination while Finnish learners used more subordination. Sindhi students wrote longer yet simpler sentences, yet T-units were equally proficient as their Finnish peers. The main conclusion reached was the three lowest CEFR levels were not comparable to syntactic complexity in EFL writing in L1 Finnish and Sindhi. Due to this finding, they proposed the need for

development of descriptors, teaching materials and assessment for syntax complexity based not only on the target language, but the learners L1.

Study	Indices	CEFR Levels
Hawkins and Filipović (2012)	Sentence length	A2 vs B1, B1 vs B2
		B2 vs C1, C1 vs C2
Green (2012)	Noun phrase incidence; number of modifiers per noun; sentence syntax similarity	B2 vs C1
		C1 vs C2
Gyllstad et al. (2014)	T-unit length; clause length; clauses per T-unit	Analysis of levels
		A1 through C2
Verspoor et al. (2012)	T-unit length	A1 vs A2
		A2 vs B1
Kim (2004)	Adverbial, adjective, and nominal clauses per clause; clauses and dependent clauses per T-unit; dependent clauses per clause; prepositional, participial, gerund, and infinitive phrases per clause	A2 vs B2 (more clearly between B1/B2 than between A2/B1)
Alexopoulou et al. (2015)	Sentence length; mean length of clause; subordinate clauses per T-unit	A1/A2 to B2
Lahuerta Martínez (2018a)	Sentence length, compound, and complex sentence ratios; coordinate and dependent clause ratios; noun phrases per clause	A2 vs B1

Kaushik and Huhta (2020)	<p>Sentence length; standard deviation of sentence length; length of clause; length of T-unit; dependent clause per unit; dependent clause per clause; complex T-units per T-unit; clauses per T-unit; coordinate phrases per clause; coordinate phrases per T-unit; clauses per sentence; T-unit per sentence; complex nominal per clause; complex nominal per T-unit; verb phrases per T-unit; syntactic simplicity (z-score & percentile); left embeddedness; modifiers per noun; minimal edit distance per part of speech; sentence syntax similarity; noun phrase density; verb phrase density; adverbial phrase density; preposition phrase density; negation density; gerund density; infinitive density</p>	A1 vs A2 vs B1
Kuiken and Vedder (2019)	<p>Clauses per T-unit; dependent clause per clause; coordination between T-units; coordination within a T-unit; coordination between constituents; complement clause density; adverbial clause density; relative clause density; number and mean-length of post-modifying noun phrases</p>	A2 vs B1 vs B2
Chen et.al., 2021	<p>Subordinate clause per word count; number of adverbial clauses; number of relative clauses; number of coordinate clauses</p>	<p>Analysis of levels A1 through C2</p>

Table 13 Previous syntactic complexity studies in EFL writing across CEFR levels (Khushik & Huhta,2020).

Table 13 provides of general summary of indices used in recent studies focusing on syntactic complexity using CEFR proficiency levels in their research. The vast amount of information was compiled by Khushik and Huhta (2020) with several newer data points being included by this study.

Kuiken and Vedder (2019) carried out an exploratory study looking into the variation of syntactic complexity measures across A2-B2 levels in L1 and L2 in argumentative essays from 89 writers in Dutch, Italian and Spanish. Their measures took into account overall complexity, subordination, coordination and phrasal complexity. They hypothesized that L2 learners would score higher on coordinate structures, yet lower on subordinate structures and post-modifying phrases.

Significant correlations were found for Italian L2 learners in coordination within T-units, use of relative clauses and length of post-modifying noun phrases. Findings mostly confirm the developmental sequence of syntactic interlanguage complexification hypothesis that with an increase in L2 proficiency, learners will use more subordinate and coordinate structures. One caveat is that their results were largely based on Italian L2 learners which was the largest group in the study and significant correlations for L2 Dutch and Spanish were not found (Kuiken & Vedder, 2019).

Kuiken and Vedder (2019) hypothesized that L2 groups would score higher on coordinate structures but lower on subordinate and post-modifying phrases. This hypothesis was confirmed both in Italian with post-modifying phrases and in Spanish with relative clauses. Spanish L1 writers used more coordinators within T-units than other L2 learners. In contrast, no significant differences were found in between L1 and L2 Dutch. Bearing in mind the findings in the three languages, the authors call for further research into L1 and L2 cross-linguistic differences.

The Kuiken and Vedder (2019) study indicated there was a variation of syntactic complexity in written L2 production across levels A2-B1, across the three target languages (Dutch, Italian, Spanish) as well as between L2 and L1. Findings underline the importance of distinguishing the different types of complexity. The reason for this is that fine-grained measures appear to lead to specific findings that cannot be demonstrated by use of general measures. Results highlight the importance of employing general and specific measures for researching syntactic complexity across both proficiency levels and languages. The authors highlighted that further research with more participants and a higher range of proficiency levels may unveil more types of variation. Additional issues to be researched could be whether and to what extent there are interchanges between complexity by coordination, subordination and relative clauses.

2.2.3 Applying Structural Functional Grammar to Syntactic Complexity

Systemic Functional Grammar views the gateway to semantics as being clause-based rather than word-based. This approach interprets grammar categories within the context of a clause or group of words. The concept of expansion within the clause complex is traditionally a form of subordination within the clause which is reinterpreted as a distinction between hypotaxis in the clause complex (Halliday & Matthiessen, 2006). Norris and Ortega (2009) assert that theoretical and empirical justifications exist for measurement of syntactic complexity through SFL. Individuals learning a first or second language are expected to change from a dynamic style to a synoptic style of expression as they become more sophisticated in their linguistic competencies and as they grow more capable of dealing with written and academic formal registers (Ortega, 2003).

Bulté and Housen (2018) designated the developmental sequence of syntactic interlanguage complexification hypothesis in relation to Norris and Ortega's (2009) premise that expansion from dynamic to synoptic styles would mean parataxis would be a marker indicating coordination at beginning level (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012). Hypotaxis would indicate complex language use at an intermediate level. Hypotaxis would then decrease at advanced levels in lieu of learners increasing use grammatical metaphor (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Wolf-Quitero et al., 1998). Subordination subsides at advanced levels of development in favor of greater use of phrasal-level complexification, which becomes the persuasive means by which syntactic complexity is achieved (Norris & Ortega, 2009; Ortega, 2009, p. 563). Biber (2006) provided empirical evidence for such patterns in mature L1 English academic registers working with corpus linguistics which was agnostic to Hallidayan grammar.

On a granular level Halliday (2006) explains that dynamic styles serve for oral, low-formality levels in non-technical communication. Synoptic style evolves out of communicative purpose of written, high formality levels of communication. Therefore, dynamic styles emerge with naturalistic acquisition of an L1 while synoptic styles come later in life as cognitive abilities increase. In reference to L2 acquisition, Ortega (2003) applies Halliday's perspective:

And indeed, from an L2 acquisition perspective, the empirical evidence has begun to corroborate that, with increasing competence in a target language, and if there is a concomitant need to meet increasingly more formal and technical communication demands, L2 users develop synoptic styles that are built out of nominalization and other forms of grammatical metaphor. Importantly for a theory of interlanguage complexification, the reliance on nominalization occurs at the expense of subordination, whose importance recedes in the language used by advanced learners in formal and academic genres. (p. 145)

Applying SFL to subordination and linguistic complexity allows researchers wide ranging possibilities to measure interlanguage complexity for a variety of purposes. When learners move towards advanced levels of proficiency, subordination is likely to peak or decrease and its predictive power greatly diminishes. Mean length of clause is a good index of complexity typical of synoptic styles which should characterize advanced levels under formal (academic or technical) communicative demands (Norris & Ortega, 2009; Ortega, 2003, 2012).

Subordination is considered an appropriate choice for measuring language complexity at the point in development when higher density and range of subordinating devices may begin to be deployed in production. Evidence in SLA literature suggests subordination occurs in low-intermediate to intermediate levels of proficiency, which takes place in most curriculums during the first two years of study for L2 learners (Byrnes et al., 2010; Colombi, 2002). On the other hand, subordination appears uninformative to measure complexity in the early years of L2 development and at the point of very advanced use (Ortega, 2012, p. 146).

While subordination is one factor in measuring linguistic complexity, findings using traditional complexity measures have been inconsistent (Ryshina-Pankova, 2015). Norris and Ortega (2009) and Ortega (2012) detail that an increase in subordination and length of T-unit are indicative of only some levels of language development and do not capture the entire span. Bulté and Housen (2014) infer that subordination may not be adequate to gauge complexity in every circumstance (p. 56). The issue at stake is that traditional complexity measures employing subordination are separated from the idea of meaningful content and communicative goals (Pallotti, 2009).

Linguistic complexity that focuses on the meaning dimension of complexity has been absent from research in traditional complexity measures. SFL can be used to characterize linguistic complexity in development and communicative appropriateness. The initiative considers that language use and development are motivated by the meaning-based, contextual and communicative tasks in which learners actively respond (Ryshina-Pankova, 2015).

SFL as a theory of language takes educational development seriously by suggesting a language-based “idealized knowledge path” (Veel, 1997, p. 189). SFL appears to be well suited for a collegiate foreign language program whose approach to writing development is aspired to make such connections to facilitate advanced levels of L2 literacy. It would do so by inextricably linking knowledge in various content levels and language, language and L2 learning, and language and gaining a competent non-native voice over an extended instructional period (Byrnes, 2009, p. 51).

Byrnes et al. (2010) provided evidence that SFL takes educational development seriously in their robust longitudinal and cross-sectional study of college-level German L2 learners. They demonstrate that reliable and noticeable increases in interlanguage complexity occur over the span of four years in instructed development. The results of their findings indicate overall length was reflective of complexification yet could not shed light on different types of complexifications involved that affected length at different time junctures. Subordination density was a reliable and valid indicator only for changes that occurred during the first two years of study as intermediate learners worked out subordination. Clausal length was a reliable and valid indicator only for changes that occurred after learners finished their third year (Ortega, 2012, p. 146).

Support for synoptic styles outlined by SFL spans several target languages. The interesting question of not only cross-linguistic validity but additionally cross-linguistic influence appears (Ortega, 2012, p. 147). Neff et al. (2004) argued that formal written styles in L1 Spanish favor a much greater use of subordination than English, and that this linguistic-rhetorical preference is transferred by L2 writers to the new language. Therefore, an improvised cross-linguistic understanding of dynamic and synoptic styles is imperative if SLA researchers are to be able to generate precise appraisals of developmental patterns of complexification vis-a-vis cross-linguistic influences (Ortega, 2012).

Ortega (2012) summarizes how SFL can improve a cross-linguistic understanding of dynamic and synoptic styles:

It will be important to pay closer attention to cross-linguistic comparisons and to understand grammatical metaphor from the combined perspective of the two languages of L2 users: How pervasive is grammatical metaphor in either language, in what styles does it occur, and through what lexico-grammatical resources is it instantiated? Second, it is worthwhile to investigate both subordination and grammatical metaphor not as unitary phenomena, but as comprising different dimensions and types of complexity. (p. 148)

It would appear meaning-based approaches have received a lack of attention. However, meaning-based theorization of complexity in L2 writing has raised significant questions in research. Ryshina-Pankova (2015) and Bulté and Housen (2012) propose a meaning-based SFL approach to complexity which would entail rethinking the concept of the theoretical taxonomy of complexity constructs by foregrounding the links among formal complexity (lexical, syntactic), semantic complexity (concrete vs. abstract) and contextual complexity (a written account for a friend vs written argumentation for a professional audience).

A SFL meaning-based approach would use a particular linguistic resource, related to the communicative demand of the task thus initiating an investigation which would start with the analysis of genre and the situational context of the tasks an L2 writer must complete. This type of analysis involves not just the organizational structure (i.e., narrative, argumentation) but other variables like subject matter (field) and the type of audience (tenor) which indicate the level of linguistic complexity needed for a given task such as subordination versus phrasal elaboration (Ryshina-Pankova, 2015, p. 58).

A multi-year longitudinal case study involving three L2 German learners who completed three consecutive instructional levels in an articulated undergraduate program at Emory University in the United States involved traditional syntactic complexity variables in addition to SFL-based measures of learners' emergent meaning-making (e.g., transitivity, taxis, thematization patterns) and learner interviews. Analysis of learners' emerging and

textual meaning-making abilities used four forms of analysis. Firstly, coordination and subordination from the linguistic complexity analysis served to indicate the relationship between parataxis and hypotaxis. Secondly, relationships were further examined to indicate logico-semantic relations between expansion and projection to underline the meaning-based relationship between clauses. Lastly, a transitivity analysis was conducted of nominal groups, processes with verbal groups and circumstances (Maxim, 2021).

Learners in the Maxim (2021) study reported in interviews that there were times when there was a close connection between explicit instruction and written production. There were also occurrences when learners indicate they were guided more by the L1 thought process. Learners cited the influence of SFL instruction at different points in the curriculum and on different language points. It is suggested that individual variability in language development would benefit from constant and consistent assistance that a meaning-based curriculum could provide.

Study	SFL Indices	Participants	Task
Byrnes (2009)	grammatical metaphor	14 L1 English learners of L2 German	Writing tasks: 1 narrative 1 journalistic report 1 public speech
Sulistiyani- ngrum & Rasyid (2015)	parataxis/hypotaxis; extension/enhancement/elaboration	8 L1 Indonesian learners of English L2	Oral presentations
Zarco- Tejada et al. (2016)	parataxis; extension/enhancement/elaboration; exposition/clarification/addition/adversative/ variation/matter/manner/spatial-temporal/causal- conditional	L1 Spanish learners of English L2	A2/B1/B2 written English & B1/B2 oral English
Yang et al. (2017)	parataxis; extension/enhancement; addition/adversative/causative/temporal	64 L1 Mandarin learners of L2 English	Equivalent to A2 3 narrative essays
Wenhui Xuan (2019)	parataxis/hypotaxis; extension/enhancement/elaboration; temporal/result/reason/purpose/ manner/condition/concession	50 L1 Chinese learners of English L2	10 writing tasks collected over 1 year

Rasool & Mahmood (2023)	parataxis/hypotaxis; extension/enhancement/elaboration; addition/variation/alternation/temporal/ spatial/result/reason/purpose/ manner/condition/concession	L1 Pakistani learners of L2 English	100 high and low graded argumentative essays in ICLE
-------------------------	---	-------------------------------------	--

Table 14 Studies using SFL Indices to Research Syntactic Complexity.

Table 14 gives a summary of studies using meaning-based indices derived from SFL. Although not dealing with parataxis nor hypotaxis, Byrnes (2009) reported on 14 English L1 emergent adult L2 German learners using SFL as a structural framework to research grammatical metaphor in a longitudinal study. In a basic sense, using grammatical metaphor would mean nominalizing a verbs and nouns. The study found that at the intermediate and advanced level grammatical metaphor in writing tasks remained stable. A significant increase appeared between advance and high advanced levels with occurrences of grammatical metaphor nearly tripling. An important take away from this study is that it confirms DSSICH which claims use of grammatical metaphor increases advanced levels.

Although not written texts, the Sulistyaningrum and Rasyid (2015) study was included since it is one of few studies which investigated the kinds of taxis and logico-semantic relations. Oral presentations given by eight L1 Indonesian students in an accelerated program were analyzed according to the previously mentioned SFL indices. The students were divided into two groups by gender with four girls composing the first group and four boys in the second group. Either group was required to speak about two different subjects. Results of their study can be seen in Table 15.

Taxis	Group 1 (girls)	Group 2 (boys)	Total
Paratactic Elaboration	58	18	76
Hypotactic Elaboration	7	10	17
Paratactic Extension	34	13	47
Hypotactic Extension	0	3	3
Paratactic Enhancement	10	6	16
Hypotactic Enhancement	37	16	53
Total	146	66	212

Table 15 Sulistyaningrum and Rasyid (2015) Number of kinds of Taxis Group 1 and Group 2.

Paratactic elaboration (76) had the highest number of occurrences when compared to hypotactic elaboration (17). Moreover, paratactic elaboration was the category with the highest number of uses in the entire study. Results for extension showed parataxis (47) with significantly higher use by learners than hypotaxis (3). In contrast, the findings involving enhancement showed a different pattern. Enhancement through hypotaxis (53) was used much more often than parataxis (16). A drawback to this study is that it is unclear if the primary focus was results relating to SFL or the accelerated language program.

The CEFR-levelled English Corpus in combination with the Coh-Metric tool for computing cohesion and coherence was used by Zarco-Tejada et al. (2016) to examine L2 English written texts at proficiency levels A2, B1 and B2 as well as L2 English spoken texts at the B1 and B2 levels. Both written and spoken texts were analyzed using measures taken from SFL. The authors chose the most representative texts of written, exercises and the same number of grammatical exercises totaling 10,000 words per level. The data is assumed to be taken from L1 Spanish students at the University of Cádiz in Spain, however it is not explicitly stated in the study. What is noteworthy about this study is that it only focuses on parataxis within the categories of elaboration, extension and enhancement then through separate conjunctions under the makers for apposition, clarification, addition, adversative, variation, matter, manner, spatial-temporal and causal-conditional. Summaries for the above markers will be given for the relevant categories found in the current study.

Zarco-Tejada et al. (2016) found that parataxis through extension (Table 16) showed addition as the category with the highest number of uses by learners and is driven by the conjunction *and*. Adversative conjunctions are the second most often used linguistic resource with *but* having the most occurrences. The CEFR level B1 showed peak use in both addition and adversative conjunctions. Variation was the least used category in extension with *or* most often being used by L2 English learners. In the case of variation, A2 and B1 showed the highest number of hits with learners using these conjunctions less at B2.

Marker	Conjunction	A2	Total A2	B1	Total B1	B2	B2 Total	Marker total
Addition	<i>and</i>	245		273		265		
	<i>also</i>	8	253	7	280	5	272	805
	<i>nor</i>	0		0		3		
Adversative	<i>but</i>	55		87		71		
	<i>yet</i>	4	60	8	98	3	77	235
	<i>however</i>	1		3		3		
Variation	<i>instead</i>	0		2		1		
	<i>apart from that</i>	0	21	0	20	1	10	51
	<i>or</i>	21		18		8		
		A2 Total 334		B1 Total 398		B2 Total 359		Marker Total 1091

Table 16 Zarco-Tejada et al. (2016) Extending Conjunctions Summary.

Markers for enhancement (Table 17) included a much larger number than extension yet the overall occurrences were lower thus showing learners used these conjunctions less often. Zarco-Tejada et al. (2016) explained that manner conjunctions were not used at all by learners since they are not found in learning material at indicated CEFR levels (p. 230).

Marker	Conjunction	A1	Total A1	B1	Total B1	B2	B2 Total	Marker total
Manner	<i>likewise</i>	0		0		0		
	<i>similarly</i>	0	0	0	0	0	0	0
	<i>in a different</i>	0		0		0		
	<i>way</i>	0		0		0		
Spatial- temporal	<i>afterwards</i>	0		0		1		
	<i>then</i>	1		17		19		
	<i>next</i>	1		21		7		
	<i>first</i>	3		8		14		
	<i>just</i>	3		17		32		
	<i>now</i>	5	32	13	103	23	151	286
	<i>finally</i>	0		2		9		
	<i>in the end</i>	1		2		0		
	<i>so</i>	16		22		27		
	<i>next time</i>	2		0		0		
	<i>at that time</i>	0		1		0		

Causal- conditional	<i>so</i>	0		0		0		
	<i>then</i>	17		19		26		
	<i>as a result</i>	0		0		1		
	<i>otherwise</i>	0		0		1		
	<i>if not</i>	0	18	0	38	1	50	106
	<i>yet</i>	0		8		3		
	<i>still</i>	0		5		11		
	<i>though</i>	0		3		4		
	<i>however</i>	1		3		3		
		A2 Total 50		B1 Total 141		B2 Total 201		Marker Total 392

Table 17 Zarco-Tejada et al. (2016) Enhancing Conjunctions Summary.

The findings to highlight in this category are that as proficiency increases, use of spatial-temporary conjunctions also increases. Additionally, causal-conditional enhancement was mainly driven by the conjunction *then* with a scattering of occurrences with different conjunctions showing peak use at various proficiency levels. The overall trend for spatial-temporal and causal-conditional conjunctions and enhancement was as proficiency increases, use also increases.

Table 18 shows the findings for elaborating conjunctions in the Zarco-Tejada et al. (2016) study with no single conjunction standing out with higher occurrences. L2 English learners used clarifying conjunctions at a higher rate than appositive ones. Learners used relatively higher amounts of elaboration as proficiency level increases. However, the results for elaboration demonstrated the lowest occurrences in the entire study thus insinuating learners do not use these resources very often at the A2, B1 or B2 level.

Marker	Conjunction	A2	Total A2	B1	Total B1	B2	B2 Total	Marker total
Appositive	<i>I mean</i>	0		0		1		
	<i>thus</i>	0	0	0	1	1	3	4
	<i>for example</i>	0		1		1		
Clarifying	<i>at least</i>	0		0		2		
	<i>anyway</i>	0	2	2	4	3	6	12
	<i>actually</i>	1		2		0		
	<i>in fact</i>	1		0		1		
		A2 Total 2		B1 Total 5		B2 Total 9		Marker Total 16

Table 18 Zarco-Tejada et al. (2016) Elaborating Conjunctions Summary.

Yang et al. (2017) carried out a study looking into parataxis through extension and enhancement using additive, adversative, causal and temporal markers (Table 19). The study took written data from 64 L1 Mandarin Chinese learners writing in a series of three narrative essays in the equivalent level coinciding with A2. The category with the most frequent occurrences was adversative followed by causal then additive and temporal. Each marker had one or two conjunction that stood out as having a much higher frequency than other categories. *And* had the highest frequency in additive, *but* in adversative, *because/so* in causal and *then* in temporal.

Marker	Total	Conjunction	Frequency
Additive	79 times	<i>and</i>	91.14%
		<i>and also</i>	2.53%
		<i>or</i>	2.53%
		<i>that is</i>	2.53%
		<i>besides</i>	1.26%
Adversative	155 times	<i>but</i>	86.45%
		<i>though</i>	12.26%
		<i>however</i>	0.65%
		<i>in fact</i>	0.65%
Casual	125 times	<i>because</i>	56%
		<i>so</i>	42.2%
		<i>therefore</i>	0.8%

Temporal	48 times	<i>then</i>	58.33%
		<i>before that</i>	10.42%
		<i>finally</i>	10.42%
		<i>at this time</i>	10.42%
		others (<i>after that, at this</i>	10.41%
		<i>moment, at first)</i>	

Table 19 Yang et al. (2017) Frequency of Conjunction Application by Participants.

Wenhui Xuan (2019) conducted a year-long study using ten varied writing tasks from 50 Chinese high school students to investigate parataxis and hypotaxis in extension and enhancement through temporal, result, reason, purpose, manner, condition and concession indicators. Table 20 references the results for logico-semantic relations in student writing. Learners used hypotactic elaboration (2.99%) only slightly more than paratactic elaboration (1.99%). Findings for extension and enhancement are more impactful. Extension through parataxis (35.32%) was used at a significantly higher frequency than hypotaxis (1.00%). In contrast, enhancement through hypotaxis (41.29%) was the linguist tool most often used by Chinese L1 writers in L2 English. Parataxis (1.49%) only accounted for a small percentage of use in enhancement.

Logico-semantic relations	Parataxis	Hypotaxis
Elaboration	1.99%	2.99%
Extension	35.32%	1.00%
Enhancement	1.49%	41.29%

Table 20 Wenhui Xuan (2019) Logico-semantic Relations in Student Writing.

What stands out about the Wenhui Xuan (2019) study is how it details enhancement markers (Table 21), yet one drawback is it never goes into more detail in terms of the fine-grained coordinators and subordinators that influence results. Nevertheless, temporal markers occurred more often in hypotaxis (24.43%) with no occurrences happening in parataxis. Result markers saw learners using parataxis (1.16%) slightly less than hypotaxis (3.49%). In contrast, hypotaxis was used in reason (27.91%) and purpose (8.14%) while there was no recorded use in parataxis for either enhancement marker.

Enhancement Marker	Parataxis	Hypotaxis
Temporal	0.00%	24.42%
Result	1.16%	3.49%
Reason	0.00%	27.91%
Purpose	0.00%	8.14%
Manner	1.16%	4.65%
Condition	1.16%	26.74%
Concession	0.00%	1.16%

Table 21 Wenhui Xuan (2019) Enhancement Markers in L1 Chinese Writing in L2 English.

Markers for manner saw parataxis (1.16%) being used less than hypotaxis (4.65%). Conditional hypotaxis (26.74%) is used at a much higher frequency than parataxis (1.16%). Markers for concession did not make up a significant percentage in enhancement with parataxis showing zero use and hypotaxis (1.16%) very little use. It is important to note that Wenhui Xuan (2019) is one of the very few studies found that went in-depth into L2 use of hypotaxis and extension as well as the principal markers of enhancement.

In one of the more recent studies, Rasool and Mahmood (2023) looked at 100 high-graded essays (HGAEs) and low graded argumentative essays (LGAEs) in the International Corpus of Learner English (ICLE). The notable aspect about this study is it looked into the types of taxis and logico-semantic relations using a large number of measures. Table 22 shows how extension is the most often used resources in parataxis followed by enhancement then elaboration. Enhancement is used at a higher frequency in hypotaxis than either extension or elaboration.

Expansion Relation		Occurrences (HGAEs)	Percentage	Occurrences (LGAEs)	Percentage
Parataxis	Extension	470	17.4%	358	20%
	Elaboration	43	1.6%	22	17.8%
	Enhancement	52	1.9%	32	1.10%
	Total	569	21%	412	20%
Hypotaxis	Extension	26	0.9%	19	0.95%
	Elaboration	15	0.56%	11	0.50%
	Enhancement	407	15.1%	316	15.7%
	Total	448	16.6%	345	17.1%

Table 22 Rasool & Mahmood (2023) Occurrences and Percentages of Expansion Relations.

One of the key characteristics of the Rasool and Mahmood (2023) study is the amount of detail that is applied to taxis, expansion and principal markers. Table 23 describes the different attributes of extension through parataxis and hypotaxis. In both HGAEs and LGAEs the total for additive extension (771) is by far the most often used linguistic resource in parataxis and the entire study. If HAGEs and LGAEs are added together, parataxis through alternation (30) is used more often than variation (27). Extension in hypotaxis displayed a similar pattern with the total for addition (32) being the most often used resource. Hypotactic variation (11) was used more by learners than alternation (2).

Extension Relation		Occurrences (HGAEs)	Percentage	Occurrences (LGAEs)	Percentage
Parataxis	Additive	438	16.2%	333	16.5%
	Variation	15	0.56%	12	0.6%
	Alternation	17	0.63%	13	0.65%
Hypotaxis	Additive	19	0.74%	13	0.65%
	Variation	6	0.19%	5	0.25%
	Alternation	1	0.04%	1	0.05%

Table 23 Rasool & Mahmood (2023) Occurrences and Percentages of Extension Relations.

When we look at the totals for HGAEs and LGAEs in elaboration (Table 24), paratactic clarification (42) is used more than exemplification (23) while learners did not resort to exposition in their argumentative essays. Hypotaxis incorporated two categories with clarification (18) being used more than description (7) (Rasool & Mahmood, 2023).

Extension Relation		Occurrences (HGAEs)	Percentage	Occurrences (LGAEs)	Percentage
Parataxis	Clarification	29	1.08%	13	0.65%
	Exposition	0	0%	0	0%
	Exemplification	14	0.53%	9	0.45%
Hypotaxis	Clarification	11	0.41%	7	0.35%
	Description	4	0.15%	3	0.15%

Table 24 Rasool & Mahmood (2023) Occurrences and Percentages of Elaboration Relations.

Enhancement was the largest grouping of measures in the Rasool and Mahmood (2023) study and demonstrated a variety of results (Table 25). Totals for both HGAEs and LGAEs temporal indicators showed they were used in hypotaxis (163) with a much higher frequency than parataxis (16). Spatial markers saw no use in parataxis while in hypotaxis (3)

there were not many occurrences. Manner was not included in the parataxis study but recorded 68 total hits in hypotaxis. Conditional markers showed a rather large disparity in use case with learners using hypotaxis (171) at a much higher frequency than parataxis (7). The study differentiates paratactic causal markers (57) from hypotaxis cause indicators (318). While it is not certain if this is an error in the categorical markers, it can be assumed by the data that hypotaxis is more often used than parataxis.

Enhancement Relations			Occurrences (HGAEs)	Percentage	Occurrences (LGAEs)	Percentage
Parataxis	Temporal	Same Time	4	0.15%	2	0.10%
		Later Time	6	0.22%	4	0.20%
	Total		10	0.37%	6	0.30%
	Spatial		0	0%	0	0%
	Manner	Means	0	0%	0	0%
		Comparison	1	0.04%	3	0.15%
		Total	1	0.04%	3	0.15%
	Conditional	Positive	1	0.04%	1	0.05%
		Negative	0	0%	0	0%
		Concession	3	0.11%	2	0.10%
		Total	4	0.15%	3	0.15%
	Causal	Cause	09	0.33%	05	0.25%
		Effect	28	1.04%	15	0.75%
		Total	37	1.37%	20	1%
Hypotaxis	Temporal		86	3.19%	77	3.84%
	Spatial		1	0.04%	2	0.10%
	Manner	Means	34	1.26%	13	0.65%
		Comparison	07	0.26%	5	0.25%
		Quality	07	0.26%	2	0.10%
		Total	48	1.78%	20	1%
	Conditional	Positive	61	2.26%	67	3.34%
		Negative	05	0.19%	17	0.85%
		Concessive	16	0.59%	5	0.25%
		Total	82	3.04%	89	4.43%
	Cause	Result	63	2.34%	36	1.79%
		Purpose	33	1.22%	19	0.95%
		Reason	94	3.49%	73	3.64%
		Total	190	7.05%	128	6.38%

Table 25 Rasool & Mahmood (2023) Occurrences and Percentages of Enhancement Relations.

2.3 Acquisition of L2 Subordinate Clauses: Overview of L1 Spanish and English

There is evidence that indicates bilinguals learn to integrate cues from both languages thus applying rules from their L1 to the L2 in certain occasions. On other occasions, they use a specific set of rules for each language or use an amalgamation of strategies (Hernández et al., 1994; Hernández et al., 2000; Kilborn, 1989; Wulfeck et al., 1986). It is important to consider the nonhomogeneous nature of the native target, and we should not expect complete convergence in L1 and L2 development (Larsen-Freeman, 2006).

The following overview summarizes a series of relevant research within L1 English and Spanish acquisition of L2 subordination. Most of the studies include a broad range of syntactic complexity indices with subordination measures composing only a fraction of total indices. Therefore, the summary of results is intermingled with other measures, yet it tries to show how L2 subordination is acquired within the broad picture of syntactic complexity.

2.3.1 Spanish L1 Acquisition of English L2 Subordinate Clauses

Native and L2 spoken language data serve a dual purpose to assess whether a particular linguistic phenomenon is available in a language and to allow native/L2 comparisons in comparable contexts. Studies relating to subordination call for more research into how Spanish L1 affects English L2 learning (Basterrechea & Weinert, 2017; Neff et al., 2004; Reyes & Hernández, 2006).

Neff et al. (2004) conducted a three-year study comparing Spanish and U.S. professional and developing writers in both L1s and English L2. The study found a significantly higher finite subordination and less non-finite subordination in Spanish L1 than English L1. Fourth year Spanish L2 university students produced considerably more T-units than first year students thus suggesting the likely developmental aspect of fluency on L2 writing.

A crucial finding in Neff et al. (2004) was that subordination complexity in finite clauses per T-unit in fourth year Spanish university students showed a significantly higher level of clausal embedding than native English students. This occurrence appeared due to greater ratios of production of nominal and adverbial clauses per T-unit than the U.S. study

abroad group. When non-finite and finite clauses were included in T-unit production in subordination ratios, fourth year students were more complex than first year English L2 learners. There appeared regular, however not significant, progression of total subordinate clause ratios from first year to fourth year in Spanish professional writers which may be evidence of L1 transfer on to the L2. A preference for finite subordination in the L1 might encourage the use of these types of clauses in L2 English.

Information related variables showed finite *that*-clauses were overwhelmingly used by EFL Spanish students in the Neff et al. (2004) study. This pattern doesn't seem to indicate a transfer in L1 discourse patterns rather a developmental pattern as advanced university students tend to adopt an impersonal stance. Non-finite clause production by Spanish students and professional writers seemed to offer evidence of L1 transfer regarding the type of subordinate clauses produced. The fourth-year group tended to produce different non-finite clauses and are likely influenced by a dominant preference for infinitival clauses. The significance of the results suggests that once EFL students become proficient enough in English, they can more readily transfer their Spanish L1 preferences to the L2.

Research conducted by Basterrechea and Weinert (2017) explores Spanish L1 subordination transfer in spoken L2 English. This study investigates the concept of L2 learners' ability to integrate subordinate clauses as an indication of higher proficiency (Ellis & Barkhuizen, 2005; Tarone & Swierzbis, 2009). Results of their study showed that English L2 speakers exhibited a strong preference for pre-posed *if*-clauses consisting of one *if*-clause. English L2 speakers used fewer multiclausal *if*-clause complexes as well as postscript and semiformal post-posed clauses in conversation. Their study found English L2 speakers did not use single *if*-clause directives in the map task analysis. The findings seem to indicate that L2 speakers were constrained in their structural and hence functional repertoire (Basterrechea & Weinert, 2017, p. 897).

The Lahuerta Martínez (2018a) study examined written texts from 188 lower and intermediate level students in their third and fourth year of secondary education. The analysis investigated quantitative syntactic complexity measures including a gender analysis with a holistic rating of learners' overall writing ability at sequential, clausal and phrasal levels of syntactic organization. Measures included mean length of sentence, simple sentence ratio, compound sentence ratio, complex sentence ratio, compound-complex sentence ratio, coordinate clause ratio, dependent clause ratio and noun phrase per clause.

Results from the Lahuerta Martínez (2018a) study revealed a strong correlation between holistic ratings and in all but one of the complexity measures with a general increase from grade 3 to grade 4. It was observed that girls obtained a higher score in the general quality of compositions and in all examined measures. There was a significant increase in sentence coordination and subordination in higher grades and a decrease in simple sentences.

2.3.2 English L1 Acquisition of Spanish L2 Subordinate Clauses

A multitude of studies have assessed L2 Spanish writing proficiency in universities across the United States ranging from beginner to advanced level courses with research focusing on language development using written learner corpora, but few studies have focused on the degree of longitudinal growth across proficiency levels (Asención-Delaney & Collentine, 2011; Collentine & Collentine, 2020; Lu, 2011; Polio & Shea, 2014). Restrepo-Remos (2021) conducted a study involving twenty-two learners enrolled in two sections of a third-year Spanish composition class who were assigned nine compositions of 150-200 words each. The results indicated learners' proficiency levels were marked by different syntactic predictors, such as verbs, auxiliary verbs, adverbial condition, coordination, relative clauses, adverbial condition and T-units.

Beginners increased their production of all syntactic indices including adverbial and relative clauses, yet T-units and coordination saw a decrease. Intermediate use of relative clauses decreased while adverbial clauses slightly increased at the end of the semester. T-units and coordination measures also decreased for intermediates at the end of the semester. Results suggest beginners greatly benefited as their rate of production significantly increased in number of words, relative and adverbial clauses. In contrast, this effect was absent at the intermediate level which might suggest that this group needs more time to develop more complex syntactic production. Enrollment for Spanish classes is the highest for foreign languages at U.S. universities, however there are few studies of Spanish L2 writing in college classes (Restrepo-Remos, 2021, pp. 678-680).

Collentine and Collentine (2020) assert commonly used approaches to linguistic complexity oversimplify the phenomenon and disregard theoretical and corpus linguistic conceptualizations. These approaches are thought to be incompatible with cognitive and

psycholinguistic conceptualizations of linguistic complexity as they do not acknowledge a learners' developmental stages. More importantly, common measures are developed in the context of English as a L2 and do not consider the morphological complexities of a highly inflected L2 like Spanish.

The corpus-based approach used by Collentine and Collentine (2020) incorporated the *Corpus del español* involving 130 interlocutors with 55,640 words at the developmental stages of beginner, intermediate and advanced. The syntactic structures used in this study assimilated adverbial clauses of cause, contingency, purpose, causal-conditional and time as well as nominal and relative clauses. Findings indicated at the beginning and intermediate levels learners produced complexity in the form of causal adverbial phrases such as *porque/because* and *puesto que/since*. At the advanced level, learners tend to complexify relative clauses (Collentine & Collentine, 2020).

Asención-Delaney and Collentine (2011) carried out one of the first-known multidimensional studies of L2 Spanish looking into how speakers used lexical and grammatical phenomena in writing within different types of interlanguage discourse. The corpus study contained 202,241 words in L2 Spanish texts written by second-year and third-year university native English speakers focusing on 78 measures that extrapolated information from narratives, argumentative essays, summaries, mini-essay questions, descriptions and expository communicative tasks. The findings did not show any significant signs of an increase in relative clauses or subordinate clause use which tend to be indicators of syntactic complexity.

2.3.3 Interim Summary

The following summarizes the importance of the previous research outlined in Section 2.3 with reference to L1 English and Spanish acquisition of L2 subordination. A series of studies highlight the need for more research in how the L1 in both of the respective languages involved in this study affects L2 acquisition of subordinate clauses (Asención-Delaney & Collentine, 2011; Basterrechea & Weinert, 2017; Collentine & Collentine, 2020; Lu, 2011; Neff et al., 2004; Reyes & Hernández, 2006). Nevertheless, variables including the number of participants, length of study, educational context, learner proficiency level, subordination measures and genres of written tasks greatly differ. The goal of this section is to demonstrate

the ways in which English and Spanish L2 subordination is acquired within the broad picture of syntactic complexity.

The Neff et al. (2004), Asención-Delaney and Collentine (2011), and Restrepo-Remos (2021) studies shared the commonality of conducting research on L2 writing in the context of higher education. In the case of the Neff et al. (2004) study, it included a broad spectrum of L2 Spanish and English learners ranging from professional writers to university students at various stages in their academic trajectory with an assortment of subordination measures which extended from T-units production, clausal embedding, finite/non-finite subordination, and production of nominal/adverbial clauses. On the other hand, Restrepo-Remos (2021) analyzed data from twenty-two learners enrolled in two sections of a third-year university L2 Spanish composition class incorporating syntactic complexity measures such as verbs, auxiliary verbs, adverbial condition, coordination, relative clauses, adverbial condition and T-units. While the variables in both studies were greatly different, results in either study showed an increase in adverbial subordination as learners advanced. In contrast, Asención-Delaney and Collentine (2011) found no indication of an increase in subordinate or relative clauses as indicators of syntactic complexity in L2 Spanish learners in second and third-year university students.

Lahuerta Martínez (2018a) examined L1 Spanish secondary education students writing in L2 English using measures including gender analysis, mean length of sentence, simple sentence ratio, compound sentence ratio, complex sentence ratio, compound-complex sentence ratio, coordinate clause ratio, dependent clause ratio and noun phrase per clause. Overall results found an increase in subordination at higher grades. Basterrechea and Weinert (2017) studied subordination transfer in L1 Spanish to L2 English with learners' ability to produce pre-posed, post-posed and two-clause complex *if*-clauses. Their findings highlight that L2 English speakers use few non-subordinate, multiclausal postscript, and semiformulaic *if*-clauses. Moreover, L1 Spanish learners show a predisposition for one pre-posed *if*-clause plus a main clause. One take-away from the Basterrechea and Weinert (2017) study is that subordination as a measure of complexity, proficiency and development is not straightforward since it assumes a wide-range of phenomena (p. 915).

Collentine and Collentine (2020) used the *Corpus del español* to study L1 English use of Spanish L2 subordination at beginning, intermediate and advanced developmental stages.

The novel aspect of their study is that they investigated adverbial subordination through clauses of cause, contingency, purpose, causal-conditional and time in addition to nominal and relative clauses. Their findings suggested beginning and intermediate students used causal-conditional reason clauses while advanced students used more complex relative clauses.

The objective of Section 2.3 is multifaceted in which it firstly summarizes research specific to L2 English and Spanish acquisition that does not fit the parameters of syntactic complexity viewed through CEFR measures or meaning-based complexity viewed through SFL. Moreover, as previous studies focus on a variety of L1s and L2s, this section centers on Spanish and English. Secondly, this section briefly details the numerous measures used to gauge L2 English and Spanish acquisition ranging from traditional indices such as T-units (Restrepo-Remos, 2021) down to subordination through clause type (Collentine & Collentine, 2020) and individual subordinators (Basterrechea & Weinert, 2017). Neff et al. (2004) was the only study that conducted research on a cross-linguistic data set comparing acquisition of L2 English and Spanish. Thirdly, the data used by the studies outlined in this section is not representative of a large data set as proposed in the current study. Fourthly, different proficiency indicators were used such as beginner, intermediate and advanced (Collentine & Collentine, 2020) as well as different academic years in higher education (Asención-Delaney & Collentine, 2011; Neff et al., 2004; Restrepo-Remos, 2021) and secondary school (Lahuerta Martínez, 2018a). The loophole opened by the summary provided in this section serves to indicate a gap in cross-linguistic L2 English and Spanish research in which the current study aims to fill through the application of the proposed rankshifted meaning-based framework using the CEFR proficiency scale to measure meaning-based development in a large data set.

2.4 Delimiting the Método de los Relojes

The *Método de los Relojes* (2018) is a descriptive grammar approach used for teaching Spanish as a foreign language. The grammar method is a metaphorical map of the Spanish language represented by three different “relojes”, meaning “clocks” in English, that each detail specific grammar structures. The system of three “relojes” are subdivided into a series of hours in which each hour is interpreted by different verb types, forms and

subordinators. The focus of this study is on Reloj 2 since it portrays how the majority of messages are organized in the Spanish language through subordination.

The MR R2 serves as a theoretical backdrop to construct a linguistic viaduct thus bridging hypotaxis from Spanish to English with the purpose of creating a fine-grained meaning-based framework. To provide an overall picture of how the MR functions, a summary is given of its most pertinent features including verb types, forms, versions and finally an in-depth listing of R2 hours from One o'clock to Ten o'clock.

2.4.1 The Four Verb Types

MR Type 1 verbs are not very numerous, but very frequent and necessary since they express opinion and emotions. Type 1 verbs are unique because they are conjugated with personal pronouns that are located before the verb and take the *a mí me* form. It is obligatory that Type 1 verbs be accompanied by an atonic pronoun while the optional tonic dative pronoun is an expressive element that highlights the indirect compliment. The last elemental attribute of Type 1 verbs is that the direct object complement is the lexical subject of the sentence (Pérez, 2018).

Type 2 consists of two verbs: *ser* and *estar*. This classification does not have prepositions following them except for several specific conditions involving location. Type 2 verbs connect a subject to adjectives or present adjectives as nouns. Type 3 verbs are recognizable by the complements and pronouns that appear with them. Type 3 verbs must use the preposition *a* when a connection to a person is made. If a reference to a person is not made, an object appears after the Type 3. Type 4 verbs are a special verb class in which the main characteristic of these verbs is that they possess a complement which is introduced by a preposition or an adverb (Pérez, 2018).

2.4.2 Versions 1:-5:

Versions 1:-5: detail how a message can be seen from different points of view according to its possibility, probability and relation in time. Messages which are produced through versions 1: and 2: are primarily categorized as being real experiences while versions 3:, 4: and 5: are unreal experiences. The versions provide five different points of view that can

be found separately in different verb forms. It should be understood that verb versions and forms work on a parallel basis and not on separate accounts. Versions 1:-5: provide a tool to analyze, organize and clarify the messages that are articulated through verb forms. This tool locates the existence of a verb in time and reality and makes it possible to understand different characteristics of the overall verb tense depending on the message (Pérez, 2018).

2.4.3 The Five Basic Verb Forms

The five basic verb forms are divided into subcategories based on their inherent characteristics including normal forms, subjunctive forms, *amar* forms, *amando* forms, the *amado* form and the special case *ama (tu)*. A summary of verb versions and forms is given detailing their relation in time, probability, possibility, meaning and use (Pérez, 2018).

The normal form *amo* is classified in 1: and is described as representing a present or habitual action. Version 2: describes the realm of the past and contains four forms. *He amado* defines a past real action while *amé* embodies a real action located in that past that has no relation to the present. The imperfect preterit *amaba* form indicates an action that happened in the real past that takes place over a long period of time. *Había amado* describes a real action that happened in the past before an explicit past action which usually takes the *amé* form (RAE, 2010; Pérez, 2018).

Amaré and *ir a amar* in 3: describe the future as something that is probable. *Habré amado* is the compound future form and signifies a probable future action that happens explicitly before another future action that takes the form of *amaré*. Version 4: *amaría* chronicles a future action as hypothetically possible and is formally labelled as the simple conditional form. *Habría amado*, the compound conditional form, denotes a hypothetically possible action taking place before a second action occurring in *amaría*. *Habría amado* 5: renders a past action as an impossible hypothesis which is seen as something in the past that did not happen, although it was envisioned (RAE, 2010; Pérez, 2018).

Pérez (2018) delimits the subjunctive as having no unique or permanent semantic value, although there are cases in which the subjunctive form expresses future, possibility, doubt, misinformation, desire and indifference. The subjunctive can manifest through combinations of different versions. The present subjunctive is configured by *Ame* 1:/3:.. The

compound perfect subjunctive category is made up of 1:/2:/3: and is represented by the form *haya amado*. The imperfect preterit subjunctive 2:/4: *amara* (*amase*) form contains the meaning that depicts actions located in the 2: past and another 4: action posed as something possible. The pluperfect preterit subjunctive form is 5:/4:/2: *hubiera amado* or *hubiese amado* (Pérez, 2018; RAE, 2010).

Amado, the participle form, is usually part of a verbal expression that indicates there is another conjugated verb in which it depends on (RAE, 2010). It is commonly used as part of a verbal expression involving the Spanish verbs *estar*, *haber* or *ser*. The three constructions can be used in 1:-5: when the first verb is conjugated within the parameters of the version it takes place in (Pérez, 2018; RAE, 2010).

The *amar* forms are comprised of two separate structures with the first one being the simple infinitive *amar* and the second structure covered is the compound infinitive and *haber amado*. The *amar* form is neutral with only semantic properties but it does not contain grammatical attributes. *Haber amado* is present within 1:-5: and describes a neutral action posed explicitly prior to another action. In this case, the only verb that can be present is *amado* (Pérez, 2018; RAE, 2010).

The *amando* form encompasses two different constructions which are the gerund *amando* and the compound gerund *habiendo amado*. The *amando* form does not have key words and always depends on a conjugated verb. This form does not have any variations and is never combined with prepositions. *Habiendo amado* is present in 1:-5: and denotes an action that has already finished and posed as happening explicitly prior to another action. This classification is one of the verb forms that is not normal, isn't conjugated, doesn't depend on prepositions yet always relies on another conjugated verb that expresses a prior action (Pérez, 2018; RAE, 2010).

Lastly, *ama* (*tú*) 3: is used to express desires and orders. It is one of the options which is not a normal form verb in Spanish and is considered the most direct form of communication. This verb form this is expressed by five pronouns: *tú*, *vosotros*, *usted*, *ustedes* and *nosotros* (Pérez, 2018).

2.4.4 Reloj 2

R2 portrays the organizational schema in which the larger percentage of messages in Spanish are organized. R2 is the most profound and representative example of the three clocks and is the focus of this study. The messages are organized starting with the initial verb (V1) which is then followed by a subordinator that comes immediately before the subsequent verb (V2) in clausal sequence. There are also instances in which a V1 from a special verb group can directly introduce a V2 (Pérez, 2018, p. 116). However, the special verb groups are beyond the scope of this study as the focus is primarily on subordinator categories.

R2 contains twelve different grammar tendencies that arise from specific interactions between the V1 and V2 in the case of each of the twelve hours. Considering that each hour focuses on a specific grammar construction, this study will focus on the first ten hours of the clock as they form a vital base for a student learning the Spanish language. Structures belonging to one hour on the clock cannot appear in other hours. Once a structure appears in its given hour, its predetermined grammatical behavior is decided by the verb it introduces (Pérez, 2018, p. 116).

Subordinators which fall under R2 can be made up of one word like *cuando/when*. A preposition can adopt the role of a subordinator as in *para que/so that* or they can be formed of various words such as *a pesar de que/despite*. The MR classifies R2 subordinators as being either rigid, flexible, or derived from a noun. Flexible subordinators are represented by words like *para que/so that* and have a priority sequence (59) in which the subordinator appears in the middle of the sentence and a secondary order in which the subordinator comes at the beginning of the sentence (60) (Pérez, 2018, pp. 117-118).

(59) *Estoy aquí **para que** tú aprendas.*

*I am here **so that** you learn.*

(60) ***Para que** tú aprendas, estoy aquí.*

***So that** you learn, I am here.*

Normal tense verbs and *ama (tú)* forms, which are subject to many restrictions on their use, are the only categories that can work as a V1 and V2 in a R2 sentence. All other

verb forms can assume the role of a V2. One important difference is 4: *amaría* can also be included in 2:. The reason is that it can express the future from a past point of view (61) (Pérez, 2018, pp. 121-122).

- (61) *Ella me dijo el mes pasado que **llegaría** la semana pasada.*
*She told me last month that she **would arrive** last week.*

The subjunctive tense appears as the V2 in several hours of R2 under specific circumstances. A problem is evident in the fact that the entire message can be delivered through 1:-5:, but there are only three subjunctive verb tenses. The solution comes into view by repeating *ame* and *amara*. In the case of *ame*, it would be used in 1: when a subjunctive form is needed but it can also be used in 3: as long as a subjunctive form is required. *Amara* is directly related to 2: when a subjunctive form is needed as well as 4: when it falls under the same set of circumstances. This strategy allows all five versions to be covered by three subjunctive verb tenses (Pérez, 2018, p. 122).

2.4.4.1 One o'clock

The MR describes One o'clock as having many of the popular structures found in the Spanish language. The actors in the sentence are subjects which assume the actions of V1 or V2. When the V1 and V2 have the same subject (62), *amar* is used as the V2. If a different subject (63) is used the V2 changes to the subjunctive tense according to the version and it is always accompanied by the particle *que* (Pérez, 2018, p. 124).

- (62) *Yo estudio para aprender.*
I study in order to learn.
- (63) *Yo estudio para que todos podamos vivir bien el futuro.*
I study so that we can all live well in the future.

There are two representative groups of subordinators that function at One o'clock and signify two different meanings with both appearing with grammatical pauses. The subordinators listed below represent examples of this hour yet there are many other ones that are at One o'clock. Table 26 shows the first group of subordinators which have the meaning of <<para/for>> and include *para que*, *a fin de que*, *con vistas a que* and *con intención de*

que. The second group is represented by the meaning <<cuando/when>> and are formed by *antes de que* and *sin que* (Pérez, 2018, p. 124).

Meaning	Subordinator
<<para/for>>	<i>para (que)/so that</i>
	<i>a fin de (que)/so that</i>
	<i>con vistas a (que)/so that</i>
	<i>intención de (que)/with the intention of</i>
<<cuando/when>>	<i>antes de (que)/before</i>
	<i>sin (que)/without</i>

Table 26 Subordinators at R2 One o'clock (Pérez, 2018, p. 124).

2.4.4.2 Two o'clock

Only two types of structures can be located at Two o'clock. Although there are not as many structures in this hour as One o'clock, they are often more frequent and necessary. The V2 appears in normal forms in affirmative sentences (64) when there is no negation before the subordinator. The V2 appears in normal or subjunctive forms when the point of negation comes before the subordinator. With a normal V2, when negation takes place, it relates to the action of V1 (65). In contrast, if the V2 is subjunctive, the negation refers to the action presented by V2 (66) (Pérez, 2018, p. 126).

(64) *Yo como rápido porque tengo mucho trabajo.*

I eat fast because I have a lot of work to do.

(65) *Yo no como mal porque tengo mucho trabajo.*

I don't eat badly because I have a lot of work to do.

(66) *Yo no como rápido porque tenga mucho trabajo.*

I don't eat fast because I have a lot of work to do.

The *amar* form is an option for the V2 in R2 at Two o'clock that can be used with or without the same subject, or not at all. In (67), there is not a second subject. However, in (68), there are two subjects which are identical. When there are two different subjects, it is better to specify the second subject immediately before *amar* (69) (Pérez, 2018, p. 127).

- (67) *Yo creo tener razón.*
I believe I am right.
- (68) *Yo creo que (yo) tengo razón.*
I believe I am right.
- (69) *Yo estoy aquí por tener tú problemas.*
I am here because I have your problems.

When *ama (tú)* is present as the V1, the rules are organized differently. *Ama (tú)* appears as the subject of the sentence and the subjunctive tense comes after the subordinator (70). In the case of negation, *ama (tú)* comes at the head of the sentence and a normal form verb comes after the subordinator (71) (Pérez, 2018, p. 127). Table 27 shows subordinators at Two o'clock appearing with and without grammatical pauses as well as their meanings.

- (70) *Estate tan preparado que puedas pasar el examen.*
Be so prepared that you can pass the exam.
- (71) *No creas (tú) que tu primo viene hoy a la ciudad.*
Don't think your cousin is coming to town today.

Meaning	Subordinator
<<porque/because>>	<i>por (que)/for, because</i>
	<i>debido a que/due to</i>
	<i>gracias a (que)/thanks to</i>
<<consecuencias/consequence>>	<i>tan/so</i>
	<i>que/because</i>

Table 27 Subordinators at R2 Two o'clock and Meaning (Pérez, 2018, p. 128).

2.4.4.3 Three o'clock

Special attention must be paid to R2 at Three o'clock since there are two perspectives, real and unreal, that stipulate the criterion for managing the V2. Versions 1: and 2: present the message as something real and always require a V2 in the normal tense. The remaining versions 3:, 4: and 5: offer the same message but only as something hypothetical thus requiring the V2 to be subjunctive (Pérez, 2018, p. 129).

The subordinators at Three o'clock are formed of various words that encompass different meanings. The three main meanings are found in Table 28. Special attention should be paid to <<cuando/when>> as there are a variety of subordinators which compose this category (Pérez, 2018, p. 129).

Meaning	Subordinator
<<cuando/when>>	<i>cuando/when</i> ¹
	<i>siempre que/whenever</i> <<long action>>
	<i>apenas/as soon as</i> <<just after>>
	<i>según/as soon as</i> <<just after>>
	<i>tan pronto como/as soon as</i> <<just after>>
	<i>una vez que/once that</i> <<after>>
	<i>hasta que/until</i> <<until>>
<<así/like>>	<i>como/like</i> ²
<<donde/where>>	<i>donde/where</i> ³

Table 28 Particles in R2 at Three o'clock (Pérez, 2018, p. 129).

2.4.4.4 Four o'clock

Four o'clock is the only hour in which C-groups can function. The basic scheme of C-Groups (CI-CVI) is what MR calls the structures formed with a noun + *que*. These subordinators are not complements of the verbs, but nouns within the basic context V1 + V2. They have only one location which is immediately after a noun. The structures in which they are located have great freedom of movement since they can complement any noun in a sentence (Pérez, 2018, p. 131).

In general, R2 at Four o'clock presents two phrases with a single verb combined into one to avoid the repetition of the noun that is common to both (72) (Pérez, 2018, p. 131).

- (72) *El chico es rubio. + El chico habla bien. = El chico que es rubio habla bien.*
The boy is blond. + The boy speaks well. = The boy who is blond speaks well.

¹ *Cuando* appears at Three o'clock, Four o'clock and Seven o'clock in different <<meaning>> categories.

² *Como* appears at Three o'clock, Eight o'clock and Nine o'clock in different <<meaning>> categories.

³ *Donde* appears at Three o'clock and Four o'clock in different <<meaning>> categories.

Subordinators play an important role in R2 at Four o'clock (Table 29). The first subordinator in the series is *lo que* and is described as a neutral word. It is the only noun that cannot fuse with another noun given that in Spanish there are not neutral nouns (73). The subordinator *donde* represents an important difference between Three and Four o'clock. When *donde* comes after the verb, the sentence squarely falls into the Three o'clock category (74). In contrast, when *donde* comes between the noun and the verb, it is located is at Four o'clock (75) (Pérez, 2018, p. 132).

(73) *No está claro **lo que** quieres.*

*It is not clear **what** you want.*

(74) *La casa está **donde** tú vives.*

*The house is **where** you live.*

(75) *La casa **donde** tú vives es bonita.*

*The house **where** you live is beautiful.*

Meaning	Subordinator	Substitution Combination	Examples
<<donde/where>>	<i>donde/where</i> ⁴	<i>el el/la/los/las que</i>	<ul style="list-style-type: none"> • <i>La casa está donde tú vives.</i> • <i>The house is where you live.</i> • <i>La casa en la que vives...</i> • <i>The house where you live...</i>
<<cuando/when>> ⁵	n/a		<ul style="list-style-type: none"> • <i>El momento en el que ...</i> • <i>The moment when...</i>
<<así/which>>			<ul style="list-style-type: none"> • <i>La forma en la que ...</i> • <i>The way in which...</i>

Table 29 R2 Four o'clock Subordinators, Meaning and Substitution Combinations (Pérez, 2018, p. 132).

At this point Pérez introduces the concept of different groups labelled C-Groups. While there are five separate C-Groups at Four o'clock, the most pertinent to this study is CI. The origin of CI is that a noun is presented in a unit or a group of units that can be something

⁴ *Donde* appears at Three o'clock and Four o'clock in different <<meaning>> categories.

⁵ *Cuando* appears at Three o'clock, Four o'clock and Seven o'clock in different <<meaning>> categories.

unique. Therefore, it is marginalized from any other reference and in this context the two grammatical pauses, in the form of commas, are obligatory. The first comma is after the said noun and the second after V2. It can be understood the mentioned pauses work as a parenthesis and are characterized by containing extra information about noun (Pérez, 2018, p. 133).

When the pronouns *yo/tú/él/ella, mi/tu/su, este/esta/estos/estas* or a proper noun appear in a sentence, the R2 at Four o'clock phrase is always CI (76). Since the origin of the structure is known to a certain degree, the normal tense is used. The subjunctive is an option in CI when it appears in 3:-5: (77) (Pérez, 2018, pp. 133-134). In summary, Table 30 provides an example of R2 at Four o'clock CI structures.

- (76) *Aurelia, quien toma café siempre, hoy bebe coca-cola.*
Aurelia, who always drinks coffee, today drinks coca-cola.
- (77) *El hombre, el que venga el miércoles, será un poco tonto.*
The man, who comes on Wednesday, is a bit of a fool.

	CI: Noun +, que + V2, + V1	Meaning
Nouns	<i>que</i>	<<person>> or <<thing>>
	<i>el/la/los/las/lo que</i>	<<person>> or <<thing>>
	<i>el/la/los/las/lo cual es</i>	<<person>> or <<thing>>
	<i>quien/es</i>	<<person>>
	<i>cuyo/a/os/as</i>	<<person>> or <<thing>>
	<i>donde en el/la/los/las que</i>	<<cuando>> and <<así>>
Examples	<ul style="list-style-type: none"> <i>Mi amigo Juan, que/el que/el cual/quien no es tímido, declaró su amor.</i> <i>My friend John, who is not shy, declared his love.</i> <i>La casa, donde vive Paco, es pequeña.</i> <i>The house, where Paco lives, is small.</i> <i>Ese instante, cuando suena el teléfono, no me gusta.</i> <i>That moment, when the phone rings, I don't like it.</i> 	

Table 30 R2 Four o'clock CI Examples (Pérez, 2018, pp. 197-198).

2.4.4.5 Five o'clock

Five o'clock is treated as a special hour because grammatical structures introduce hypothetical versions in V1. The particle *si/if* is the only one that functions within this group

and is only compatible with 3:-5: (Table 31). The action of the verb is oriented from version 3: future as a probable option. Version 4: presents the action as a possibility. Lastly, *si* in 5: proposes a condition from the present going to the past that is impossible (Pérez, 2018, pp. 137 & 359).

Meaning	Subordinator
<<si/if>>	<i>si/if</i> ⁶

Table 31 R2 at Five o'clock Subordinator and Meaning (Pérez, 2018, p. 137).

2.4.4.6 Six o'clock

R2 at Six o'clock, together with Four o'clock, presents quadruple choice structures that allow the selection between a normal or subjunctive V2 depending on what one wants to express. The conditions go beyond the grammatical realm. Therefore, the use of a normal tense infers the security, reality, or evidence of the action of the V2 while the subjunctive form conveys insecurity, remoteness, or unimportance of that action. Structures in R2 at Six o'clock are interesting because the speaker uses them to show that extra information, in many other occasions, is not interesting at all and opts for structures of similar meaning but at a different hour (Pérez, 2018, p. 138).

Amar as a V2 is the only option for subordinators such as *después de que* and *a pesar de que* when it does not relate to the same subject. This type of structure can always be substituted by a conjugated form (78) (Pérez, 2018, p. 138). Table 32 shows the particles that are included at Six o'clock along with their meanings.

- (78) *Ella come después de venir él. = Ella come después de que él venga.*
She eats after he comes. = She eats after he comes.

⁶ *Si* appears at Five o'clock, Seven o'clock and Nine o'clock in different <<meaning>> categories.

Meaning	Subordinator
<<cuando/when>>	<i>después de (que)/after</i>
<<cuánto/how much ⁷ >>	<i>tan...como/as</i> <i>más que/more than</i> <i>menos que/less than</i>
<<obstáculo/obstacle>>	<i>aun (que)/even though</i> <i>a pesar de (que)/even though</i> <i>pese a (que)/despite</i>
<<y/and>>	<i>además de (que)/besides</i>

Table 32 R2 at Six o'clock Subordinators and Meanings (Pérez, 2018, p. 139).

2.4.4.7 Seven o'clock

R2 at Seven o'clock contains the trait in which the role of the V1 introduces substantive subordinate propositions in combination of a V2 that can only take shape in the normal tense or the *amar* form with a limited set of predefined verbs. The use of *amar* is a secondary option at this hour and is only compatible with a small grouping of verbs in this special group (Pérez, 2018, pp. 142-143). A description of said predefined verb group is beyond the scope of this study since the scope of R2 is limited to indicators of subordination. Subordinators which are present at Seven o'clock are seen in Table 33.

Meanings	Subordinator
<<o/or>>	<i>si/whether⁸</i>
(various)	<i>qué/what</i> <i>cuándo/when⁹</i> <i>cómo/how</i> <i>dónde/where</i> <i>por qué/why</i> <i>quién/-es/who</i>

Table 33 Seven o'clock Subordinators and Meanings (Pérez, 2018, p. 143).

⁷ *Cuánto* is abbreviated in the English translation in the sense it can mean *how much* or *how many* depending on if the noun it refers to is countable or uncountable.

⁸ *Si* appears at Five o'clock, Seven o'clock and Nine o'clock in different <<meaning>> categories.

⁹ *Cuando* appears at Three o'clock, Four o'clock and Seven o'clock in different meaning categories.

2.4.4.8 Eight o'clock

The most significant characteristic of R2 at Eight o'clock is that it strictly functions with normal tense verbs. Furthermore, Table 34 illustrates the subordinators that make up this hour which all highlight the meaning <<porque/because>> (Pérez, 2018, p. 149).

Meanings	Subordinator
<<porque/because>>	<i>puesto que/since</i>
	<i>ya que/now that</i>
	<i>como/as¹⁰</i>
	<i>dado que/given that</i>

Table 34 R2 at Eight o'clock Subordinators and Meanings (Pérez, 2018, p. 149).

2.4.4.9 Nine o'clock

R2 at Nine o'clock has the V1 in the normal tense while the V2 is in the subjunctive tense. The series of subordinators that are enchainned to this grammatical sequence signify <<si/if>>, <<consecuencia/consequence>>, <<cuánto/how many¹¹>> and <<para/for>> (Pérez, 2018, p. 150). The subordinators and their meanings can be seen in Table 35.

Meanings	Subordinator
<<si/if>>	<i>como/if¹²</i>
	<i>mientras/while</i>
	<i>salvo que/unless</i>
	<i>siempre que/provided that¹³</i>
	<i>siempre y cuando/as long as</i>
<<consecuencia/consequence>>	<i>de ahí que/so that</i>
<<cuánto/how many>>	<i>como si/as if</i>
<<para/for>>	<i>de forma que/in the way that</i>
	<i>de manera que/in the way that</i>
	<i>de modo que/in the way that</i>

Table 35 R2 at Nine o'clock Subordinators and Meanings (Pérez, 2018, p. 149)

¹⁰ *Como* appears at Three o'clock, Eight o'clock and Nine o'clock in different <<meaning>> categories.

¹¹ *Cuánto* is abbreviated in the English translation in the sense it can mean *how much* or *how many* depending on if the noun it refers to is countable or uncountable.

¹² *Como* appears at Three o'clock, Eight o'clock and Nine o'clock in different <<meaning>> categories.

¹³ *Siempre que* appears at Three o'clock and Nine o'clock in different <<meaning>> categories.

2.4.4.10 Ten o'clock

At Ten o'clock in R2 the V1 is in the normal tense while the only possibility for the V2 is *amar*. The majority of verbs at this hour are classified in a special verb group and are too extensive to cover. There are few subordinators at this hour and they are used for writing in an educational setting more than for speaking (Pérez, 2018, p. 152). Table 36 gives an outline of subordinators found in R2 at Ten o'clock.

Meanings	Subordinator
<<cuando/when>>	<i>al¹⁴/upon</i>
<<después/after>>	<i>tras/after</i>
<<porque/because>>	<i>al¹⁵/upon</i>
<<si/if>>	<i>de/if</i>

Table 36 Ten o'clock Subordinators and Meanings (Pérez, 2018, p. 158).

2.5 Summary

Chapter 2 has laid the theoretical foundation for the following study. Section 2.1 firstly provided a detailed summary of Structural Functional Grammar theory in order to understand the concepts of parataxis and hypotaxis as their inherent definitions are crucial to the SFL based developmental sequence of syntactic interlanguage complexification hypothesis later explained in this chapter. Elaborating, extending and enhancing expansion was reviewed to introduce the fundamental categories followed by logico-semantic relations and MBUs to provide a foundation for the meaning-based framework.

The basic definition of CAF is described in Section 2.2. including its origins, development and significance in L2 research. Linguistic complexity measures have received a great deal of attention in L2 writing research, yet they are considered to be the most complex, ambiguous and least understood (Housen & Kuiken, 2009; Lahuerta Martínez, 2018a). Within the constructs of linguistic complexity, syntactic complexity is broadly defined as the range of syntactic structures that are produced and the degree of sophistication of those structures (Lu & Ai, 2015; Ortega, 2003; Pallotti, 2015) and has been the focus of

¹⁴ *Al* appears at Ten o'clock in different <<meaning>> categories.

¹⁵ *Al* appears at Ten o'clock in different <<meaning>> categories.

many studies investigating L2 writing and proficiency (Lu & Ai, 2013; Lu, 2011; Norrby & Håkansson, 2007; Ortega, 2000, 2003; Wolfe-Quintero et al., 1998).

Section 2.2.1 delves into the basic definition of syntactic complexity and explores the manner in which subordination is used as a descriptor of learner language to gauge proficiency, describe performance and benchmark development (Chen et al., 2021). A point of contention in research is the wide range of indices used to measure syntactic complexity. Within the field syntactic complexity research, it is recognized that various measures are not created equal and some redundantly measure the same thing (Bulté & Housen, 2012; Lu, 2011; Ortega, 2003; Wolfe-Quintero et al., 1998).

A significant justification for including SFL to measure linguistic complexity is the idea is that language use and development are motivated by meaning-based, contextual and communicative demands of the tasks in which learners actively respond. Arguments have been made that contextual aspects at the heart of a meaning-based approach should be added to the measures of L2 writing complexity (Palloti, 2009; Ryshina-Pankova, 2015).

Section 2.2.2.1 firstly describes the essence as well as the origin of the CEFR and provides a basic outline of proficiency levels A1-C2. Section 2.2.2.2 describes the body of research using syntactic complexity measures across CEFR levels and notes there is a need for more studies in this line of investigation (Dygers et al., 2017; Hulstijn et al., 2010). The majority of studies which have looked into this topic have not included the entire range of proficiency levels or used a set of general measures which have resulted in varied findings. There is a need for fine-grained measures which lead to more specific findings across different proficiency levels and languages (Kuiken & Vedder, 2019). A detailed summary of studies and their findings using CEFR levels to look into syntactic complexity is referenced in Table 13.

Studies applying SFL to syntactic complexity are detailed in Section 2.2.3. A fundamental element for this study is that syntactic measures should be multidimensional with SFL as theoretical justification using parataxis and hypotaxis to measure L2 learner development from beginning to advanced proficiency levels. The developmental sequence of syntactic interlanguage complexification hypothesis in which expansion goes from dynamic to synoptic styles would mean parataxis would be a marker indicating coordination at

beginning level (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012). Hypotaxis would indicate complex language use at intermediate levels while grammatical metaphor emerges at advanced levels as subordination subsides (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Section 2.3 details L1 acquisition of L2 subordinate clauses. Chen et al. (2021) summarizes the extensive body of L2 subordination research by explaining that it tends to view subordination as a unitary construct of global syntactic complexity. There have been a considerable number of studies researching L2 acquisition using a variety of languages in which subordination is an acquisitional target (Baten & Håkansson, 2015; Ehret & Szmrecsanyi, 2019; Harrington, 1987; Müller & Penner, 1996; Sasaki, 1991). There appears to be a void when researching the different types of subordinate clauses. Furthermore, L2 researchers have been requesting a more detailed analysis of linguistic complexity and subordination complexity (Biber et al., 2004; Biber et al., 2011; Bulté, & Housen, 2012; Kyle & Crossley, 2018).

Section 2.3.1 looks at studies researching L1 Spanish acquisition of L2 English subordinate clauses and notes there is a need for more research into how Spanish L1 affects English L2 learning of subordinate clauses (Basterrechea & Weinert, 2017; Neff et al., 2004; Reyes & Hernández, 2006). Neff et al. (2004) found a significantly higher rate of finite subordination and less non-finite subordination by Spanish L1 students and professional writers than English L1. There appeared a progressive increase of total subordinate clause use from first to fourth year students possibly indicating Spanish L1 transfer to the L2. Spanish L1 writers produced more non-finite clauses which possibly demonstrated L1 transfer. Results from fourth year students suggested that once EFL students become proficient enough in English, they can more readily transfer their Spanish L1 preferences to the L2.

Findings from Basterrechea and Weinert (2017) suggested that Spanish L1 speakers of L2 English exhibited a strong preference for pre-posed *if*-clauses consisting of one *if*-clause. English L2 speakers used few multiclausal *if*-clause complexes and postscript and semi-formulaic post-posed clauses in the conversation. Additionally, English L2 speakers did not use single *if*-clause directives, frequently used by Spanish L1 speakers. The findings suggested L2 speakers were constrained in their structural and hence functional repertoire.

Section 2.3.2 details English L1 acquisition of Spanish L2 subordination. In the Restrepo-Remos (2021), beginner and intermediate English L1 students were assigned nine compositions dealing with different themes in two sections of a third-year composition course. Results suggest beginners greatly benefited as their rate of production significantly increased in number of words, relative and adverbial clauses. In contrast, this effect was absent at the intermediate level which might suggest that this group needs more time to develop more complex syntactic production.

The Collentine and Collentine (2020) corpus study investigated the developmental stages of beginner, intermediate and advanced English natives using Spanish L2 clauses of cause, contingency, purpose, causal-conditional and time as well as nominal and relative clauses. Results showed at beginning and intermediate levels learners produced complexity in the form of causal adverbial phrases while at the advanced level learners tend to complexify relative clauses.

Asención-Delaney and Collentine (2011) incorporated 78 measures in their corpus study examining information from narratives, argumentative essays, summaries, mini essay questions and expository communicative tasks. The results did not show signs of relative clauses or subordinate clauses which tend to be indicators of syntactic complexity.

Section 2.4 provides an overall summary of the *Método de los Relojes* (2018) R2 starting with a brief introduction to the four verb types, versions 1:-5: and the five basic verb forms. One o'clock through Ten o'clock are described to provide an understanding of the interactions between V1 and V2 as well as the subordination categories outlined in each hour to set a theoretical backdrop for the creation on a meaning-based framework.

3. Operationalizing Parataxis and Hypotaxis Through MR R2 to Form a Meaning-Based Framework

3.1 Introduction

Clause subordination was seen in 2.2.1 Syntactic Complexity as an important linguistic phenomenon that is relevant to psycholinguistics, cognitive sciences, behavioral sciences and language acquisition (Chen et al., 2021). Numerous studies have used a wide range of syntactic complexity indices in different ways to measure, among other variables, subordination (Bulté & Housen, 2012; Lu, 2011; Norris & Ortega, 2009; Ortega, 2003; Wolfe-Quintero et al., 1998). Many studies have incorporated relatively small amounts of data as well as few indices (Lu & Ai, 2015, p.19). Evidence has demonstrated that fine-grained indices have a stronger predictive power than traditional indices in L2 writing (Zhang & Lu, 2022).

The need for an alternative to traditional syntactic complexity measures which recognize the discourse-semantic function and the dimensions of complexity through meaning in L2 writing (Ryshina-Pankova, 2015) was reviewed in 2.2.3. SFL has been proposed as a multidimensional way to measure complexity through the production of meaningful content and realization of communicative goals (Bulté & Housen, 2012; Norris & Ortega, 2009; Pallotti, 2009). Contextual aspects at the heart of a meaning-based approach should be added to the measures of L2 writing complexity (Ryshina-Pankova, 2015).

This thesis presents the MBF as a tool to measure not just subordination, but coordination through the focal lens of SFL while using MR R2 as a cross-linguistic connection to analyze L2 writing in Spanish and English. The concept of the MBF is to provide an alternative systematic approach using meaning-based fine-grained indices to gauge syntactic complexity in L2 writing. The measures established by the MBF are taken from SFG and MR R2 to provide a roadmap of common and comparable indices to measure L2 writing in Spanish and English through CEFR proficiency levels. The measures in the following sections are based upon their comparability for use in this cross-linguistic study

using data from L2 learner corpora. The following section will describe the organic composition of the MBF adhering to the notions of hypotaxis and parataxis.

Section 3.2 describes the methodology in creating the MBF. Section 3.3 covers the results and discussion as well as profoundly compares and contrasts English and Spanish within the spectrums of the MR R2 and SFG. This is done to find a basis of theoretical and pragmatic adaptability to form the foundation for the MBF. Section 3.4 titled Mean-Based Framework is the conclusion of this chapter and culmination of the theoretical concepts of this study which brings forth a set of cross-linguistic measurements to be applied to the corpus study in this thesis. The Spanish sentences used as examples in this and following sections are all extracted from the MR and translated to English. When an English sentence appears without a Spanish translation, it should be noted that it is provided as an example of the various English constructions that can be derived from the MR example.

3.2 Methodology

Research has been described as a systematic process of inquiry to address a problem in which applied linguists may use a theoretical framework to guide their research inquiry as well as a methodological approach inherent to theory (Phakiti et al., 2018, p. 10). In the case of this thesis, the theoretical framework connects hypotaxis and parataxis as found in *An Introduction to Functional Grammar* (2014) through CEFR proficiency levels to *Método de los Relojes* (2018) R2 hours One o'clock to Ten o'clock.

Contrastive linguistics encounters the problem of comparability of incommensurable systems. The solution to the comparability lies with data which deals with pairs of languages that are socio-culturally linked and can rely on bilingual output such as translations (Gast, 2012). Micro-level data is at the level of an individual, word or text (Saldanha & O'Brien, 2014). The first stage of micro-level data collection involves noting, extracting and translating subordinators and their meaning category from MR R2 hours One o'clock to Ten o'clock in order to establish separate measures. In the second step, theoretical triangulation (Litosseliti, 2010) was employed to identify accurate cross-linguistic translations of micro-level data using various resources such as wordreference.com, deepl.com in addition to theoretical references found in an *Introduction to Functional Grammar* (2014), *Método de los*

Relojes (2018) and *Systemic-Functional Grammar of Spanish: A contrastive study with English* (2010). A myriad of language dictionaries and studies were referenced throughout this thesis to ensure the accurate categorization of micro-level data. Therefore, when multiple English translations for a single Spanish subordinator or coordinator were encountered, or vice versa, the most common word was chosen in line with the previously mentioned resources and relevant cross-linguistic grammar parameters.

Table 37 illustrates the methodological approach using *mientras* at Three o'clock which results in *while* as the straightforward translation. However, the translation of *después de (que)* at Six o'clock has various possibilities. The resulting translation deemed by this study to be the most accurate for research purposes is *after*. This study consulted the previously cited resources to triangulate what was thought to be the most often applied and generalizable interpretation of any given subordinator or coordinator when more than one option existed. It must be acknowledged that a variable in this study is translation and that any number of pragmatic interpretations of a given subordinator or coordinator might be influenced by individual preference and circumstance at the time of utterance.

MR Hour	MR Meaning Category	MR Subordinator	Word Reference & Deepl Results	Final Results English Subordinator
Three o'clock	<<cuando>> <<acción simultánea>>	<i>mientras</i>	<i>while</i>	<i>while</i>
Six o'clock	<<cuando>> <<después>>	<i>después de (que)</i>	<i>after, then, afterward, afterwards</i>	<i>after</i>

Table 37 Example of translation methodology.

A mix of purposive selecting, a sample based on pre-defined critical parameters and stratified sampling, was used to obtain a sufficient number of equal measures to provide a valid representative sample ensuring that specific groups are represented. Once an equivalent English subordinator was encountered from the MR, SFG was consulted to construct a unit of analysis. The triangulation process was reversed in the same manner when including additional subordinate and coordinate units from SFG with the purpose of increasing the total number of units of analysis. The working definition used for a unit of analysis considers the

total sum of measurable concepts such as lexical items, sentences, clauses, phrases and collocations (Saldanha & O'Brien, 2014). The meaning-based framework was created to highlight a series of rank shifted measurement levels.

The first rank shifted measurement level encompasses taxis. The second rank shifted layer is expansion through extension, enhancement and elaboration. The third layer is defined as the meaning-based category (MBC) which incorporates the principle logico-semantic markers of expansion like temporal, causal-conditional, etc. Within most MBCs there is a smaller construct which will be defined as the meaning-based subcategory (MBSC). The final measurement is the micro-level data, the subordinator or coordinator in question, which is termed the meaning-based unit (MBU). When relevant, the meaning and hour in relation to R2 is included to add an additional layer of analysis. Table 38 provides a concrete example of how units of analysis are constructed.

Level 1 Taxis	Level 2 Expansion	Level 3 MBC	MBSC	Level 4 MBU	MR Hour/ Meaning
Hypotaxis	Extension	Addition	Additive: positive	<i>además de (que)/besides</i>	Six o'clock <<y>> <<incluso>>
	Enhancement	Temporal	Different time: later	<i>después de (que)/after</i>	Six o'clock <<cuando>> <<después>>
	Elaboration	Non-defining relative clause	-	<i>quien/who</i>	Four o'clock C1
Parataxis	Extension	Addition	Additive: positive	<i>y/and</i>	-
	Enhancement	Temporal	Same time	<i>y mientras tanto/and meanwhile</i>	-
	Elaboration	Exposition	-	<i>en otras palabras/in other words</i>	-

Table 38 Units of Analysis within Meaning-based Framework (Halliday, 2014; Pérez, 2018).

The unit of analysis is connected to the important concept of operationalization. The meaning of operationalization in the context of this thesis refers to the operations involved in measuring the dependent variable. The dependent variable(s) in this study is the meaning-

based framework. Operationalization refers to the tools selected to accomplish the goal of the research at hand which is to use the MBF to measure the frequency of occurrence in data extracted from corpus databases.

An important element that has yet to be mentioned is the operationalization of CEFR proficiency levels A1 through C2. Within each unit of analysis, and when applied in the corpus study, proficiency levels are included as a gauge to measure L2 acquisition and L1 to L2 transfer. The below example shows CEFR proficiency levels which are firstly incorporated into the Level 1 hypotaxis then Level 2 expansion through enhancement. The measure number is presented then the name of the corpus is listed. The SFG format lists the Level 3 temporal MBC then the same time MBSC. Finally, Level 4 MBUs *while/mientras* are listed. The MR R2 hour is referenced to show which category it is classified under. Lastly, CEFR levels are included to measure the frequency throughout the proficiency scale. This brief explanation has been supplied to give an overview of the methodology. This aspect of the study will be outlined in greater detail in the following chapters dealing with the corpus study.

	Corpus	MBC	MBSC	MBU	MR R2	A1	A2	B1	B2	C1	C2
1	EFCAMDAT2	temporal	same time	<i>while</i>	Three o'clock	-	-	-	-	-	-
	CEDEL2 & CAES			<i>mientras</i>		-	-	-	-	-	-

3.3 Results and Discussion

The following section covers the results and discussion of the cross-linguistic and theoretical analysis of MR R2 and SFG to form the MBF. This section starts with MR at One o'clock and finishes at Ten o'clock. As mentioned earlier, Eleven and Twelve o'clock are not included since they do not cover a concept of subordination which is applicable to this study in the same way as the previous ten hours.

3.3.1 One o'clock

The main characterization for R2 at One o'clock is the unique use of the Spanish subordinator groups with the connotation <<para/for>> and <<cuando/when>> (Pérez, 2018,

p. 124). The first subordinator group detailed in R2 at One o'clock involves the category <<para/for>>. The structures signifying *a que* (79), *a fin de que* (80) and *con vistas a que* (81) are translated into the English as *so that* which directly correlates to Halliday's notion of hypotaxis and expansion through enhancement in the casual-conditional subcategory under condition through cause and purpose introducing a finite verb form (Halliday, 2014, p. 477).

Hypotaxis and expansion: enhancement: casual-conditional: cause: purpose: finite

- (79) *Vengo **a que** me des las patatas del mercado.*
*I come **so that** you give me the potatoes from the market.*
- (80) *Nos enseñas **a fin de que** aprendamos mucha gramática.*
*You teach us **so that** we learn a lot of grammar.*
- (81) *Estudio español **con vistas a que** me comprendáis.*
*I study Spanish **so that** you can understand me.*

So that is an explicit subordinator for purpose that introduces a finite clause (Eastwood, 1994). The fine line between <<purpose>> and <<result>> must be drawn when defining *so that*. The chief semantic difference is that result clauses are factual, and the outcome is achieved. Purpose clauses are putative with the consequence yet to be achieved. Therefore, result clauses employing *so that* cannot accept a modal auxiliary, yet purpose clauses can (Quirk et al., 1985, p. 1107). One o'clock displays a variety of <<purpose>> subordinators yet <<result>> needs to be properly identified when it appears.

The authors of *Systemic Functional Grammar of Spanish: A Contrastive Study with English* (2010) explain the differences between reason and purpose. Reason presents real facts while purpose deals with wishes and intentions. Reason in Spanish is typically represented by the indicative and purpose by the subjunctive in the secondary clauses (p. 57). In line with this current study, the authors make a reference to *para que* (82) translated as *so that* when realizing purpose. The English translation of the below example as well as the examples taken from MR provide evidence that *para que*, when signifying purpose, is generally translated to *so that*. The resulting English V2 appears in a finite form with a different subject which would be the counterpart to the Spanish subjunctive in these types of sentences.

Hypotaxis and expansion: enhancement: casual-conditional: cause: purpose: secondary clause finite

- (82) *Chusa trata de llamar la atención **para que** dejen de estar pendientes de él.*
*Chusa tries to draw attention **so that** they stop watching him.*

The subordinator *para que* (83) appears in translations with the preposition *for*. The meaning induced by *for* tends to delimit purpose and is a subordinator that introduces a *-ing* form (Lindstromberg, 2010, p. 229; Eastwood, 1994, p. 253). The phrase *a cambio de que* (84) also includes *for* in its structures when translated to English as seen with *in exchange for*. Either previously mentioned English subordinator can be categorized under the category of hypotaxis and expansion through enhancement in casual-conditional cause and purpose with a non-finite preposition (Halliday, 2014, p. 477).

Hypotaxis and expansion: enhancement: casual-conditional: cause: purpose: non-finite preposition

- (83) *Yo estoy aquí **para que** tú aprendas español.*
*I am here **for** you to learn Spanish.*
- (84) *Yo trabajo **a cambio de que** mi familia viva dignamente.*
*I work **in exchange for** my family to live decently.*

The basic meaning of *para que* and *a cambio de que* falls into the same parameters as *so that* meaning a purposeful outcome. The difference from the previous examples involving a finite form is that a non-finite preposition is used. A salient observation was made with the preposition *for* with both examples in which the V2 appeared as the *to*-infinitive with a different subject in the sentence. The parallel Spanish examples showed the V2 appearing in the subjunctive form.

The basic meaning of *para* is purpose and can be translated into various subordinators that signify *so that*. (85) provides an example in Spanish of a sentence with an infinitive form. In this case, the English translation appears without a subordinator and the *to*-infinitive directly follows the verb which signifies a clause of purpose. Clauses of purpose tend to be more infinitival than finite in English (Quirk et al., 1985, p. 1107). The interpretation with the resulting translation of *para* appearing as *in order to* must also be considered when translating Spanish subordinators to English. This type of construction also appears with a *to*-

infinitive in which the non-finite preposition *to* denotes purpose. With the inclusion of *that*, the phrase *in order that* prompts a finite form (Collins Cobuild, 1990; Halliday, 2014).

(85) *Yo estudio para aprender.*

I study to learn.

I study in order to learn.

The translation of *con la intencion de* (86) resulted in being *with the intention of*. This subordinator is categorized in hypotaxis and enhancement in causal-conditional through cause and purpose with the inclusion of a non-finite preposition (Halliday, 2014, p. 477). When compared to the previous example in (85), *with the intention of* appears with a non-finite *-ing* form instead of a *to*-infinitive.

Hypotaxis and expansion: enhancement: casual-conditional: cause: purpose: non-finite preposition

(86) *Trabaja mucho con la intención de ahorrar dinero para el viaje de estudios.*

He works hard with the intention of saving money for the study trip.

There are a variety of syntactical patterns that incorporate meaning-based subordinators for casual-conditional under condition through cause and purpose with a finite form. Example (1) in Table 39 employs *so that* with the following V2 in the indicative form. The particle *for* expressing <<purpose>> appears before the V2 *-ing* form to signify <<purpose>> in (2). The translation for *para* in (3) materializes with *in order* plus a V2 in the *to*-infinitive while (4) shows *in order that* introducing a finite form. The construction *in exchange for* takes shape introducing a V2 *-ing* form (5). Lastly, *with the intention of* in (6) comes between the V1 and the V2 *-ing* form.

Example	Structure
1	V1 + <i>so that</i> + V2 finite
2	V1 + <i>for</i> + V2 -ing form
3	V1 + <i>in order</i> + V2 to-infinitive
4	V1 + <i>in order that</i> + V2 finite form
5	V1 + <i>in exchange for</i> + V2 -ing form
6	V1 + <i>with the intention of</i> + V2 -ing form

Table 39 English R2 One o'clock <<casual-conditional: condition: cause: purpose: finite>>.

The second group detailed in R2 at One o'clock involves the category <<cuando/when>>. The prototypical particles in this group include *antes de que* (87) and *sin que* (88) and are translated to *before* in English. Different than the R2 <<purpose>> category, *before* appeared in hypotaxis and expansion through enhancement in the temporal different time earlier subcategory introducing a finite verb form (Halliday, 2014, pp. 477-487). The V2 in the Spanish examples appear in the present subjunctive form while English examples appear in the present form with the V2 appearing in the indicative form.

(87) *Hago la tortilla **antes de que** tú cuentes ocho.*

*I make the omelet **before** you count to eight.*

(88) *Hago el desayuno **sin que** te sientes.*

*I make breakfast **before** you sit down.*

Antes de que is categorized in the same hypotactic network as *before*. However, an important distinction arises when *que* is removed from the structure. The sentence in (89) shows how with the inclusion of *que* prompts a present subjunctive V2 form in Spanish when there are different subjects. The manifestation in English takes place with *before* giving rise to a V2 finite form with different subjects. When *que* is dropped from the sequence, the remaining structure *antes de* (90) causes the Spanish V2 verb form to appear in infinitive with the same subject.

Hypotaxis and expansion: enhancement: temporal: divergent: earlier: secondary clause finite

(89) *Concilió un sueño breve, lleno de grumos, **antes de que** sonara el despertador.*

*He fell into a brief slumber, full of lumps, **before** the alarm clock went off.*

Hypotaxis and expansion: enhancement: temporal: divergent: earlier: secondary clause non-finite

(90) *Aprovechó para ir de compras **antes de** tomar el vuelo de vuelta a casa.*

*He took advantage of the situation to go shopping **before** taking the flight back home.*

The English word that represents both *antes de* and *antes de que* is *before*. Referring to *antes de*, the V2 appears in a non-finite form in both languages although this is manifested as the infinitive in Spanish and the *-ing* form in English (Arús et al., 2010, p. 49). *Before* as a preposition signals a time or event that marks two actions happening at distinct moments. *Before* can also introduce finite and non-finite clauses (Eastwood, 1994; Quirk et al., 1985).

English sentence patterns for R2 at One o'clock signifying temporal in an earlier different time take two different forms (Table 40). In (1) *before* appears between the V1 and the V2 with an indicative form while in (2) it appears as a *-ing* non-finite form.

Example	Structure
1	V1 + <i>before</i> + V2 finite
2	V1 + <i>before</i> + V2 <i>-ing</i> form

Table 40 English R2 One o'clock <<temporal: different time: earlier>> Structures.

3.3.2 Two o'clock

The subordinators found at R2 at Two o'clock signify <<porque/because>> and <<consecuencia/consequence>> (Pérez, 2018, pp. 126-127). Subordinators with the meaning <<porque/because>> include four forms which appear as *porque* (91), *a causa de que* (92), *gracias a que* (93) and *debido a que* (94). The primary translations of *porque* and *a causa de* coincide with *because* with a finite V2 or *because of* with a non-finite *-ing* V2. *Gracias a que* corresponds to *thanks to* and *debido a que* appears as *due to*. Both *thanks to* and *due to* appear with a non-finite *-ing* V2. The four previously mentioned subordinators share a correlation with hypotaxis and expansion under enhancement through a casual-conditional relationship that includes cause and reason (Arús et al., 2010, p. 65; Halliday, 2014, p. 477; Quirk et al., 1985, pp. 1103-1105 & 1070).

Hypotaxis and expansion: enhancement: casual-conditional: cause: reason: finite

- (91) *Estoy en Santander **porque** llueve.*
*I am in Santander **because** it is raining.*
- (92) *Me caes bien **a causa de que** eres muy simpática.*
I like you because you are very nice.

Hypotaxis and expansion: enhancement: casual-conditional: cause: reason: non-finite

- (93) *Estoy vivo **gracias a que** tú me ayudas.*
*I am alive **thanks to** you helping me.*
- (94) *Me duele la cabeza **debido a que** cambia el tiempo.*
*My head hurts **due to** the changing weather.*

The blueprint for causal-conditional under condition in cause and reason subordinators can be seen in Table 41. Example (1) takes form with a V1 plus *because* following the V2 in the finite form. The differential element in (2), (3) and (4) is the inclusion on the preposition *of* into the syntactical framework which prompts the V2 *-ing* non-finite form.

Example	Structure
1	V1 + <i>because</i> + V2 finite
2	V1+ <i>because of</i> + V2 <i>-ing</i> form
3	V1 + <i>thanks to</i> + V2 <i>-ing</i> form
4	V1 + <i>due to</i> + V2 <i>-ing</i> form

Table 41 English R2 Two o'clock <<causal-conditional: condition: cause: reason>> Structures.

The second subordinator group representing <<consecuencia/consequence>> is made up of *tan que*. There are two sequences that compromise this group which firstly include *tan* + adjective/adverb + *que* which translates into English as being *so* + adjective/adverb + *that*. This structure in English is categorized as an adverb as a clause element which is represented by the sequence *so* + adjective/adverb + *that*-clause. *So* is a linking adverb that is an amplifier in academic prose and semantically intensifies the adjective following it (95). The construction *so* + adjective + *that* also introduces patterns that combine the notion of sufficiency or excess with the notion of result (96) which is suggested to be the consequence of an action (Biber et al., 1999, pp. 550; Quirk et al., 1985, pp. 435 & 1143).

- (95) *He wants it **so urgently that** he fidgeted in his chair.*
 (96) *I'm **so happy** to hear your good news **that** I could kiss you.*

The second sequence is *tantos/a/os/as* + noun + *que* which in English is represented by two constructions. Firstly, *so many* + countable noun + *that* and secondly *so much* + uncountable noun + *that*. In this case, the linking adverbial *so* is combined with countable or uncountable nouns to form a clause of result. In this type of comparative clause *so...that* expresses result which is considered the consequence of an action (97) (Quirk et al., 1985, p. 1109).

- (97) *Her parents gave her **so many toys that** she couldn't possibly play with them all.*

The notion of <<consecuencia/consequence>> in the MR R2 with both previously mentioned structures involving *so* + adjective + *that* (98) and *so many/much* + countable/uncountable noun + *that* (99) are categorized under hypotaxis through causal-conditional result with a finite verb form (Halliday, 2014, pp. 477-487).

Hypotaxis and enhancement: casual-conditional: cause: result

- (98) *Vengo **tan** cansado que me duermo.*
*I come **so** tired **that** I fall asleep.*
 (99) *Tu hermano escribe **tantas** cartas con el boli, que le duele la mano.*
*Your brother writes **so** many letters with his pen, **that** his hand hurts.*

An important point to highlight in the above translations is a *that*-clause appears at the end of both constructions making it a form of adverbial complementation (Quirk et al., 1985, p. 1049). There are three constructions which introduce finite clauses. The first construction is *so* + adjective + *that*. The second construction is *so* + *many* + countable noun + *that*. The third and final construction is *so* + *much* + uncountable noun + *that*.

The prototypical English R2 Two o'clock patterns incorporating the adverbial intensifier *so* structures representing causal-conditional through condition in cause and result share properties with what Pérez (2018) describes as <<consecuencia/consequence>>. A commonality that appears in every English sequence is the V1 plus the V2 taking shape in the finite form. In Table 42 (1) *so* appears with an adjective followed by *that*. Sentence (2)

exemplifies *so many* which obligates the noun to be countable. On the other hand, in (3) *so much* conditions the following noun to be uncountable.

Example	Structure
1	V1 + <i>so</i> + adjective/adverb + <i>that</i> + V2 finite
2	V1 + <i>so many</i> + countable noun + <i>that</i> + V2 finite
3	V1 + <i>so much</i> + uncountable noun + <i>that</i> + V2 finite

Table 42 English R2 Two o'clock <<causal-conditional: condition: cause: result>> Structures.

3.3.3 Three o'clock

R2 at Three o'clock has an important aspect in which the V2 appears in the indicative form in 1: and 2: to present an idea that is real while subjunctive forms appear with the V2 in 3:, 4: and 5: to present an idea that is hypothetical. There are three meaning groups at this hour including <<cuando/when>>, <<así/like>> and <<donde/where>>. The first group of subordinators represents <<cuando/when>> and incorporates seven separate forms which include *cuando/when*, *siempre que*¹⁶/*whenever*, *apenas/as soon as*, *según/as soon as*, *tan pronto como/as soon as*, *una vez que/once* and *hasta que/until*. The second group signified by <<así/like>> only contains the subordinator *como*¹⁷ which is translated to *as/like*. The third group <<donde>> has one subordinator being *donde*¹⁸/*where* (Pérez, 2018, p.129).

The two subordinators in <<cuando/when>> are *cuando* translated to *when* and *siempre que* translated to *whenever*¹⁹. Both *when* and *whenever* are simple subordinators which introduce finite forms, *-ing* forms and *-ed* forms. The use of these subordinators underlines the simultaneity of the situation in the main and subordinate clause or an overlap in time between the two situations. An important aspect of *when* and *whenever* is they can signify a sequence in which a situation in the main clause occurs after the subordinate clause. Either subordinator may indicate a non-durative situation. *When* (100) corresponds with hypotaxis and expansion through enhancement with temporal same time point introducing

¹⁶ *Siempre que* appears at Three o'clock and Nine o'clock in different <<meaning>> categories.

¹⁷ *Como* appears at Three o'clock, Eight o'clock and Nine o'clock in different <<meaning>> categories.

¹⁸ *Donde* appears at Three o'clock and Four o'clock in different <<meaning>> categories.

¹⁹ It must be noted for future reference that *siempre que* was translated to the subordinator *whenever* in wordreference.com. However, the reverse translation for *whenever* resulted in *cuando sea* which happens to fall outside the definition of subordination taken by the MR at this hour as per the explanation of the author Manuel Pérez Saiz.

finite and non-finite forms. The only difference with *whenever* (101) is that it takes place over a spread in time. Either subordinator can prompt a V2 which is finite form, a *-ing* form (102) or a *-ed* form (103) (Halliday, 2014, pp. 477-487; Quirk et al., 1985, p. 1078).

Hypotaxis and expansion: enhancement: temporal: same time point: finite

(100) *Compro el pan **cuando** vuelvo del trabajo.*

*I buy bread **when** I return from work.*

Hypotaxis and expansion: enhancement: temporal: same time spread: finite

(101) *Yo veo la tele **siempre que** ellos se van.*

*I watch TV **whenever** they leave.*

Hypotaxis and expansion: enhancement: temporal: same time point: non-finite

(102) *Be careful **when** crossing the street.*

Hypotaxis and expansion: enhancement: temporal: same time [point]: non-finite

(103) *Spinach is delicious **when** eaten raw.*

The English R2 Three o'clock patterns incorporating the adverbial intensifier *when* structure meaning temporal at same time point and *whenever* meaning temporal same time spread can be seen in Table 43. There are three different patterns with (1) finite form, (2) *-ing* non-finite form and (3) *-ed* non-finite form. It should be noted that an additional category of verbless clauses could be considered. Nevertheless, it is an ellipted form of the sentence and in its entirety would be in one of the three categories.

Example	Structure
1	V1 + <i>when/whenever</i> + V2 finite form
2	V1 + <i>when/whenever</i> + V2 <i>-ing</i> form
3	V1 + <i>when/whenever</i> + V2 <i>-ed</i> form

Table 43 English R2 Three o'clock <<temporal: same time: [point] & [spread]>> Structures.

Apenas (104), *según* (105) and *tan pronto como* (106) are translated into the *as soon as* in English. *As soon as* is a simple time subordinator that can introduce a finite or a *-ed* form. Clauses with *as soon as* appear to be mainly limited to a few adjectives such as

available, feasible, necessary and possible. As soon as indicates a sequence in which the thing happening in the main clause occurs after that in the subordinate clause. Moreover, *as soon as* adds the notion of proximity or closeness between the two clauses. These units of analysis are classified in hypotaxis and expansion through enhancement in temporal same time point constructions introducing a finite verb form (Halliday, 2014, p. 477; Quirk et al., 1985, p. 1078).

Hypotaxis and expansion: enhancement: temporal: same time: point: finite

(104) *Empieza a discutir conmigo **apenas** llega a casa del trabajo.*

*He starts arguing with me **as soon as** he gets home from work.*

(105) *Se ducha con agua fría **según** se desnuda.*

*He showers with cold water **as soon as** he undresses.*

(106) *Llega a las reuniones **tan pronto como** puede.*

*He arrives at meetings **as soon as** he can.*

The English R2 Three o'clock patterns with *as soon as* structures classified under temporal at same time point are listed in Table 44. There are two different structures incorporating (1) a finite form and (2) an *-ed* non-finite form.

Example	Structure
1	V1 + <i>as soon as</i> + V2 finite form
2	V1 + <i>as soon as</i> + V2 <i>-ed</i> form

Table 44 English R2 Three o'clock <<temporal: same time [point]>> Structures.

Una vez que/once and *hasta que/until* are two temporal markers in hypotaxis that are not in temporal same time. *Once* indicates a sequence in which the situation in the main clause occurs after that in the subordinate clause. The subordinator can also add proximity in time in two different situations. *Once* is found in hypotaxis and expansion through enhancement in the subcategory temporal at a later different time which prompts a finite form (107). *Once* is a correlative subordinator for finite clauses which proposes a time after in combination with optional conjuncts *then* and *in that case*. Clauses with the subordinator *once* can additionally introduce a non-finite *-ing* form (108) or *-ed* form (109) (Halliday, 2014, pp. 477-487; Quirk et al., 1985, pp. 998, 1078 & 1084).

Hypotaxis and expansion: enhancement: temporal: different time later: finite

(107) *Una vez **que** llegamos a un acuerdo, son imposibles los problemas.*

Once we reach an agreement, problems are impossible.

Hypotaxis and expansion: enhancement: temporal: different time later: non-finite

(108) ***Once** having made a promise, you should keep it.*

(109) ***Once** seen, the painting will never be forgotten.*

Table 45 shows English R2 Three o'clock patterns using *once* structures in temporal at a different time later. There are three different structures incorporating (1) finite form, (2) *-ing* non-finite form and *-ed* non-finite form.

Example	Structure
1	V1 finite + <i>once</i> + V2 finite form
2	V1 finite + <i>once</i> + V2 <i>-ing</i> form
3	V1 finite + <i>once</i> + V2 <i>-ed</i> form

Table 45 English R2 Three o'clock <<temporal; different time later>> Structures.

Unlike the previous two temporal categories same time and different time later, *hasta que/until* is in different time earlier. *Until* specifically marks the time up to when a situation in the matrix clause applies. The duration lasting to the time is indicated by the clause and the situation is presupposed to be true. This temporal marker is classified as hypotaxis and expansion through enhancement in temporal different time earlier which can introduce finite and non-finite forms. *Until* is a simple subordinator which can introduce a V2 which is finite (110), *-ing* form (111) or an *-ed* form (112) (Collins Cobuild, 1990; Halliday, 2014, pp. 477-487; Quirk et al., 1985, pp. 998, 1078 & 1084).

Hypotaxis and expansion: enhancement: temporal: different time earlier: finite

(110) *Come patatas **hasta que** se atraganta.*

*He eats potatoes **until** he chokes.*

Hypotaxis and expansion: enhancement: temporal: different time earlier: non-finite

(111) *I'll wait here **until** hearing from you again.*

(112) *The dog stayed in **until** told to come in.*

The English R2 Three o'clock patterns applying *until* subordination structures representing temporal at a different time earlier can be seen in Table 46. There are three different structures incorporating (1) finite form, (2) *-ing* non-finite form and (3) *-ed* non-finite form.

Example	Structure
1	V1 + <i>until</i> + V2 finite form
2	V1 + <i>until</i> + V2 <i>-ing</i> form
3	V1 + <i>until</i> + V2 <i>-ed</i> form

Table 46 English R2 Three o'clock <<temporal: different time earlier>> Structures.

The second particle group in R2 at Three o'clock is <<así/like>> formed by *como* which is translated to the simple subordinators *as* and *like*. *As* and *like*, in certain cases, can be finite adverbial subordinator of time. However, at this hour in the MR they signify similarity and comparison in the sub-class of manner. When the verb is dynamic in an *as*-clause, similarity is combined with manner. Moreover, *as* and *like* introduce finite and non-finite clauses. *As* (113) and *like* are classified under hypotaxis and expansion in enhancement under manner through comparison introducing a finite form (Halliday, 2014, p. 477; Pérez, 2018, p. 129; Quirk et al., 1985, pp. 1110-1111).

Hypotaxis and expansion: enhancement: manner: comparison: finite

(113) *Yo conduzco como quiero, la policía no me da miedo.*

I drive as I want, the police don't scare me.

The English R2 Three o'clock patterns utilizing *as* and *like* subordination structures representing manner through comparison can be seen in Table 47. There are two structures incorporating (1) finite form and (2) non-finite *-ing* form.

Example	Structure
1	V1 + <i>as/like</i> + V2 finite form
2	V1 + <i>as/like</i> + V2 <i>-ing</i>

Table 47 English R2 Three o'clock <<manner: comparison>> Structures.

The third and final category at Three o'clock is represented by <<donde/where>> and takes the shape of *where* in English. *Where* is considered a simple subordinator and appears in adverbial clauses of place in the first or second position in the sentence. *Where* may also appear in clauses of contingency yet for this hour it will be considered to only indicate location. *Where* (114) is listed in hypotaxis and expansion through enhancement in spatial same time point structures introducing a finite form (Halliday, 2014, p. 477; Quirk et al., 1985, p. 1110).

Hypotaxis and expansion; enhancement; spatial [point]: finite

(114) *Voy a **donde** tú quieres.*

*I go **where** you want me to go.*

The English R2 Three o'clock pattern utilizing *where*²⁰ subordination structures representing <<spatial point>> (Table 48). There is one structure incorporating (1) a finite form.

Example	Structure
1	V1 + <i>where</i> + V2 finite form

Table 48 English R2 Three o'clock <<spatial point>> Structures.

3.3.4 Four o'clock

R2 at Four o'clock is a unique hour since the defining factor of this hour is what Pérez (2018) defines as ConNombres, or noun phrases in English. There are six different groups that make up this category each named CI-CVI. However, since this study only deals with CI, as explanation of the remaining five will be absent. C-groups are characterized by different types of relative clauses which coincide with hypotaxis and elaboration. Relative clauses in English generally function as restrictive or non-restrictive modifiers of noun phrases and are functionally parallel to attributive adjectives (Quirk et al., 1985, p. 1048). The following section details how CI is manifested in the English language.

²⁰ *Where* with the as the translation of *donde* appears in different <<meaning>> categories at Three o'clock and Four o'clock.

3.3.4.1 CI

CI is characterized by non-restrictive relative clauses which take place when the reference of the head noun is a member of a class which can be identified through the modification that has been supplied. Any modification given to the head noun is additional information. Therefore, “the girl” in example (115) is only identifiable as *Mary Smith* given that we understand that said girl is tall and wants to meet you (Collins Cobuild, 1990; Quirk et al., 1985, pp. 1238-1329).

(115) *Mary Smith, who is the tall girl, wants to meet you.*

Since non-restrictive relative clauses are not essential for identification, it allows the speaker to make different parts of a sentence into a relative clause (116) and (117) in which the highlighted information is located between commas (Quirk et al., 1985, pp. 1239-1240).

(116) *My brother, who is an engineer, lives in America.*

(117) *My brother, who lives in America, is an engineer.*

CI contains a head noun which is a unit or a group of units that can point out something unique. Therefore, it is marginalized from any other reference and in this context the two grammatical pauses, in the form of commas, are obligatory. The first comma is after the head noun and the second before V2. In this sense, it can be understood the pauses work as a parenthesis and are characterized by containing extra information about the noun (Pérez, 2018, p. 133).

Given the information and examples presented in the MR, CI constitutes what would be considered as non-restrictive relative clauses in English. In (118), *que* is translated into the relative pronoun *who*. *La niña* was independently identified thus modification of the head noun is additional information (Pérez, 2018, p.133; Quirk et al., 1985, pp. 1239-1240). Halliday (2014) defines example (118) as hypotaxis and elaboration through non-defining relative clauses which are finite and is symbolized by $\alpha = \beta$. Apart from *who*, head noun modifiers for this category include *which* and *that* (p. 598).

Hypotaxis and elaboration: non-defining relative clause: finite

(118) *La niña, **que** llora, está loca.*

*The girl, **who** cries, is crazy.*

Halliday (2014) refers to non-defining, or non-restrictive, relative clauses as a combination of elaboration with hypotaxis. It serves as a descriptive gloss to the primary clause (119). Hypotactic elaboration allows for the introduction of background information, a characterization, an interpretation of some aspect of the dominant clause or some form of evaluation. There may be an explanatory comment, as with paratactic elaboration when clarifying. Elaborating hypotactic clauses may be finite or non-finite.

(119) *Yu, who has been visiting Taiwan this week, did not elaborate.*

A hypotactic elaborating finite non-defining relative clause happens when the secondary clause includes a *wh*-form, which is embedded as the qualifier in a nominal group and has the verb form as the dependent clause. A non-defining relative clause adds further characterization of something that is already understood as fully specific. These clauses are divided into three groupings. The first grouping of finite clauses is with *which* whose domain is either the whole clause or the primary clause (120). The sequence in this example is $\alpha \wedge = \beta$ (Halliday, 2014, pp. 464-466).

(120) */// He talks down to people, // **which** automatically puts people's backs up. ///*

The second group involves clauses with *which*, *that*, *who* or *whose* and has a nominal group as its domain. When the nominal group is non-final in the primary clause, the secondary clause is often enclosed $\langle \rangle$ signifying a grammatical pause thus (121) making the structure $\alpha \langle = \beta \rangle$ (Halliday, 2014, pp. 464-466).

(121) */// Inflation, < **which** was necessary for the system, > > became also lethal. ///*

The third group encapsulates clauses with *when* or *where*²¹ which have domain over place or time (122). In this group the secondary clause [you can give it] may be enclosed (Halliday, 2004, pp. 464-466).

(122) /// *The first few days are a time for adjustment, || <<when the kitten needs all the love and attention>> [you can give it]. ///*

Table 49 shows hypotaxis and elaboration with finite non-defining relative clauses in English in CI at Four o'clock. In (1), *who, whose, which, that, where* or *when* are the only subordinators that can be an enclosed *wh*-form and set apart by commas. The V2 appears as a finite form when combined with previously mentioned *wh*-forms.

Example	Structure
1	V1 # << <i>who, whose, which, that, where</i> and <i>when</i> >># + V2 finite form

Table 49 English R2 Four o'clock Hypotaxis and Elaboration <<non-defining relative clause>>.

3.3.5 Five o'clock

R2 at Five o'clock includes the subordinator *si/If*²² and functions with normal tense *amo* form in 3: and in the subjunctive in 4:-5:. This form is a simple subordinator for finite clauses which appears in the English as type 1 conditional clauses (123) using variations of the pattern *if* + present + *will*. (124) shows a type 2 conditional which is manifested by *if* + past + *would* and interprets an irreal situation. Lastly, (125) exemplifies type 3 conditionals inferring an irreal situation through the structure by *if* + past perfect + *would have* + past participle (Eastwood, 1994; Pérez, 2018, p. 359; Quirk et al., 1985).

(123) 3: *Iré contigo al cine si llegas pronto del trabajo.*

3: *I will go with you to the movies if you come home early from work.*

(124) 4: *Iría contigo al cine si llegaras pronto del trabajo.*

4: *I would go with you to the movies if you came home early from work.*

²¹ *Where* as the English translation for *donde* appears in different <<meaning categories>> at Three o'clock and Four o'clock.

²² *Si* appears at Five o'clock, Seven o'clock and Nine o'clock in different <<meaning>> categories.

- (124) 5: *Habría ido al cine si hubieras llegado pronto del trabajo.*
 5: *I would have gone to the movies with you if you had come home early from work.*

The main use of conditional clauses is to express a direct condition. This type of clause conveys the situation in the main clause is directly contingent on that of the conditional clause. The truth of the proposition in the main clause is a consequence of the fulfillment of the condition in the conditional clause. The most common and most versatile subordinators are *if* for positive conditions and *unless* for negative conditions. Other conditional subordinators are *as long as*, *so long as*, *assuming that*, *given that*, *in case*, *in the event that*, *on the condition that*, *provided that* and *supposing that*. *If*, *unless* and all the previously mentioned subordinators are used in finite clauses. Additionally, *if* and *unless* have the capability to introduce non-finite clauses which are mainly *-ed* (125) and verbless (126) (Quirk et al., 1985, pp. 1088-1091).

- (125) *The grass will grow more quickly if watered regularly.*

- (126) *If wet, the pipe won't give you good smoke.*

Halliday (2014) classifies *if* in hypotaxis and expansion through enhancement under casual-conditional positive condition with both finite (127) and non-finite subordinators. *Unless* (128) is classified under the same criteria with the exception that it construes a negative condition in both finite and non-finite clauses. It is important to note the translation of *unless* can take the form of several subordinators including *a menos que*, *salvo que* or *a no ser que* which are categorized in the MR at Nine o'clock in 3:-5:.

Hypotaxis and expansion: enhancement: casual-conditional: condition: positive: finite

- (127) ||| *If I had a different view*, || *then perhaps I would write more novels*. |||

Hypotaxis and expansion: enhancement: casual-conditional: condition: negative: finite

- (128) ||| *You will cherish them on your bookshelves for a long time* – || *unless, of course, someone borrows them* || *and somehow 'forgets' to return them*. |||

Systemic Functional Grammar of Spanish: A Contrastive Study with English (2010) highlights a construction using the casual-conditional in positive condition with a finite verb (129) using *si* in Spanish, yet in English this doesn't appear in the subjunctive form (p. 59). In

English, this type of construction is known as the type 0 conditional in which the pattern *if* + present form + present form appears (Eastwood, 1994).

Hypotaxis and expansion: enhancement: causal-conditional: condition: positive: real

(129) *Si tú conoces un poco más a tu mamá, pregúntale cualquier cosa que tú creas que le pueda dar un motivo de conversación.*

If you know your mom a bit more, ask her anything you think might give her a reason to converse.

The English R2 Five o'clock pattern incorporating *if* representing casual-conditional under positive condition can be seen in Table 50. The plethora of conditional subordinators outlined earlier are not included under this category since their meaning denotes casual-conditional through a concessive condition. Conditional sentences in this category are type 0 and 1 in 1: real structures, while types 2/3 are 4:/5: unreal structures. All the constructions outlined below can appear with V2 finite (1), V2 *-ing* form (2) and V2 *-ed* form (3).

Example	Structure
1	V1 + <i>if</i> + V2 finite
2	V1 + <i>if</i> + V2 <i>-ing</i> form
3	V1 + <i>if</i> + V2 <i>-ed</i> form

Table 50 English R2 Five o'clock <<casual-conditional: condition: positive >> Structures.

3.3.6 Six o'clock

R2 at Six o'clock is the second hour in which quadruple choice, the decision between normal or subjunctive use, is available. This choice is largely left to the speaker and their view of how they want to construe reality. English is considered to have a much narrower definition of the subjunctive and the quadruple choice schema is not applicable here. This study will outline the subordinators at this hour in relation to English and SFG. There are four groups at this hour which include <<cuando>>, <<cuánto>>, <<obstáculo>> and <<y>>.

The first category <<cuando>> has the meaning of *después de que/after* (130). *After* in a simple subordinator that appears in adverbial finite clauses of time and serves as a preposition introducing a subordinate clause. This subordinator indicates the situation occurs

after the main clause marking the sequence of main and dependent clause (Pérez, 2018; Quirk et al., 1985, pp. 1078-1084).

- (130) *Nosotros dormimos **después de que** esa señora hace/haga la cama.*
*We sleep **after** that lady makes the bed.*

After is categorized in *Introduction to Functional Grammar* (2014) under hypotaxis and expansion through enhancement at a temporal different time later than the main clause. *After* can be a finite subordinator (131) or a non-finite preposition (132) (Halliday, 2014, p. 477).

Hypotaxis and expansion: enhancement: temporal: different time: later: finite

- (131) ***After** he survived the disgrace, he became a fine citizen.*

Hypotaxis and expansion: enhancement: temporal: different time: later: non-finite: preposition

- (132) *I took a bath **after** working in the garden all day.*

The model for the English version employing temporal different time later constructions using *after* can be seen in Table 51. The basic model involves (1) *after* with a V2 finite verb and (2) with a non-finite *-ing* form V2.

Example	Structure
1	V1 + <i>after</i> + V2 finite
2	V1 + <i>after</i> + V2 <i>-ing</i> form.

Table 51 English R2 Six o'clock <<temporal: different time: later >> Structures.

The second group involving <<cuánto>> is represented by *tan...como*, *más que* and *menos que*. All three subordinators are used in comparative clauses. At this point in the study subordinators for comparative clauses will be referred to as comp-elements. It is believed that comparative clauses answer *how*-questions. In a comparative construction the proposition expressed in the main clause is compared to a proposition expressed in the subordinate clause with respect to a standard of comparison (Pérez, 2018; Quirk et al., 1985, pp. 1027-1029).

Ellipsis of a part of the comparative clause is likely to occur when that part is a repetition of something in the main clause. As both clauses are likely to be parallel in structure and content, ellipsis tends to be a rule rather than an exception in comparative clauses. The comparative clause, if it overlaps with the content of the main clause in respects to the comp-element, can be an independent structure. The most common type of comparative clause is one that imitates the structure of the main clause and repeats the whole content except for the differing element which provides contrast (Quirk et al., 1985, pp. 1030-1031).

Tan...como, translated to *as...as* (133) in English, expresses equality in a comparative clause. An adjective is placed between the two comp-elements in this construction like with *as handsome as*. *As...as* is a comparison of equivalence that also contains the constructions *as many* incorporating a countable noun (134) and *as much* incorporating a noncountable noun (135) (Pérez, 2018; Quirk et al., 1985).

(133) *Pedro es **tan guapo** como yo (soy).*

*Peter is **as handsome** as I (am).*

(134) *Regamos **tantas flores** como ellos (riegan).*

*We water **as many flowers** as they (water).*

(135) *I agree with you as much as **I agree** with Robert.*

The English comp-element in (136) is a comparative adjective plus *than* used to compare something with a greater quality. (137) shows the second manner to compare a quality greater than that of something else by using *more* (Eastwood, 1994).

(136) *Yo camino **más rápido** de lo **que** caminas (tú).*

*I walk **faster than** you (walk).*

(137) *Most hotels are **more** comfortable **than** motels.*

Less is used as a comparison to say something lacks the same qualities as the proposition of the main clause. The countable *fewer (than)* appears in (138) and the noncountable *less (than)* appears in (139). The fact that there are two words in English with the meaning of *menos* marks a difference between languages (Eastwood, 1994; Pérez, 2018).

(138) *(Él) Tiene **menos** hijas de las **que** tienes.*

*He has **fewer** daughters **than** you (have).*

(139) *Juan es **menos** bobo de lo **que** tú (eres).*

*Juan is **less** dumb **than** you (are).*

No references to comparative clauses were found in SFG in parataxis neither hypotaxis. However, a few references indicate comparatives like *as...as* being in the realm of expansion (Halliday, 2014, p. 22). Therefore, the same comparative examples listed above are repeated to show how they are classified in SFL. Quirk et al. (1985) detail comparative clauses under the chapter describing the syntactic and semantic functions of subordinate clauses. This study sets comparative clauses aside as a subclause type of expansion with the attributes of the comp-element displaying greater, equal or less quality, and the verb appearing in the finite form. (140), (141) and (142) are examples showing comp-elements appearing with an equal quality.

Expansion and comparison: equal quality: finite

(140) *Peter is **as handsome as** I am.*

(141) *We water **as many flowers as** they water.*

(142) *I agree with you **as much as** I agree with Robert.*

In contrast to comp-elements exhibiting equal status, (143) and (144) are manners in which a greater status is represented. Lastly, the comp-element in examples (145) and (146) demonstrates a status of being less than the principle comparative subject.

Expansion and comparison: greater: finite

(143) *I walk **faster than** you do (walk).*

(144) *Most hotels are **more** comfortable **than** motels.*

Expansion and comparison: less: finite

(145) *He has **fewer** daughters **than** you (have).*

(146) *Juan is **less** dumb **than** you (are).*

Table 52 displays the comparison structures found at Six o'clock. (1) shows comparative constructions of equal status. In contrast, (2) and (3) exhibit comparative

structures of greater status. Lastly, (4) shows a comparative construction indicating less of a quality taking shape with *fewer* plus countable noun and *than* and (5) introduced by *less* plus uncountable noun plus *than*.

Example	Structure
1	V1 + <i>as</i> + adjective + <i>as</i> + V2 finite
2	V1 + adjective + <i>than</i> + V2 finite
3	V1 + <i>more</i> + adjective + <i>than</i> + V2 finite
4	V1 + <i>fewer</i> + countable noun + <i>than</i> + V2 finite
5	V1 + <i>less</i> + uncountable noun + <i>than</i> + V2 finite

Table 52 English R2 Six o'clock <<manner: comparison>> Structures.

The third classification involving <<obstáculo>> contains *aunque* (147), *a pesar de que* (148) and *pese a que* (Pérez, 2018, p.139). There are various subordinators that can be extracted from the translation of *aunque* and *a pesar de que* including *even though*, *even if*, *although* and *though*. They are informal subordinators for concessive clauses that may introduce V2 finite or V2 non-finite forms (Quirk et al., 1985, p. 1097). *Even though*, *even if*, *although* and *though* are found in SFG under hypotaxis and expansion through enhancement under causal-conditional in concessive condition introducing a finite verb form as well as a non-finite conjunction (Halliday, 2014, p. 478).

Hypotaxis and expansion: enhancement: condition: concessive: finite

(147) *Yo como la tortilla **aunque** tiene/tenga grasa.*

*I eat the tortilla **even though** it has fat in it.*

(148) *Te odio **a pesar de que** tú eres/seas buena persona.*

*I hate you **even though** you are a good person.*

A pesar de que and *a pese que* differs from *aunque* since they can additionally be translated to *despite* (149) and *in spite* which are also is categorized under hypotaxis and expansion through enhancement under concessive condition introducing non-finite prepositions.

Hypotaxis and expansion: enhancement: condition: concessive: non-finite preposition

(149) *Tú me odias **a pesar de** ser yo el más inocente.*

*You hate me **despite being** the most innocent.*

Table 53 displays the structures for causal-conditional through concessive condition. (1) shows the concessive condition for *even though*, *even if*, *although* and *though* with V2 finite form. On the other hand, (2) and (3) respectively appear in the *-ing* and *-ed* non-finite forms. (4) details the additional translation of *pese a que* and *a pesar de que* which appear as *in spite* and *despite* with a preposition introducing a non-finite form.

Example	Structure
1	V1 + <i>even though/even if/although/though</i> + V2 finite
2	V1 + <i>even though/even if/although/though</i> + V2 <i>-ed</i> form
3	V1 + <i>even though/even if/although/though</i> + V2 <i>-ing</i> form
4	V1 + <i>in spite/despite</i> + preposition + non-finite form

Table 53 English R2 Six o'clock <<causal-conditional: condition: concessive>> Structures.

The last category formed by <<y>> is comprised of *además de que* (150) which is translated into *besides* in English (Pérez, 2018, p.139). *Besides* is considered an additive conjunct (Quirk et al., 1985). Halliday (2014) classifies *besides* as a non-finite preposition under hypotaxis and expansion through extension under positive addition (p. 471). (1) in Table 54 shows the single *-ing* form in which *besides* can appear.

Hypotaxis and expansion: extension: additive: positive: non-finite preposition

(150) ***Además de que*** *no tiene/tenga inteligencia, es feo como el culo de un mono.*

Besides *having no/he has no intelligence, he's ugly as a monkey's ass.*

Example	Structure
1	V1 + <i>besides</i> + V2 <i>-ing</i> form

Table 54 English R2 Six o'clock <<addition: positive>> Structures.

3.3.7 Seven o'clock

R2 at seven o'clock has five groups encompassing <<si>>, <<quién>>, <<cómo>>, <<cuándo>>, <<dónde>> and <<por qué>>. The only possible combination for this group is with a V1 which expresses a substantive subordinate proposition in combination of a V2 that can only take shape in the normal tense or the *amar* form with a limited set of predefined verbs (Pérez, 2018, p. 143). Since the main focus of this study is centered on subordination

and coordination, the previously mentioned predefined verb group is not described. The main difference between these words, other than <<si>>, is the addition of an accent mark, which implies interrogation or exclamation and does not contain a grammatical pause.

The first group meaning <<si>>²³ (151) is translated to *whether* in English and is considered a correlative subordinator which can introduce finite (1), *to*-infinitive (2), *-ing* (3) and *-ed* (4) forms as seen in Table 55. Furthermore, *whether* is a hypotaxis through causal-conditional in alternative condition concessive subordinator which combines the conditional meaning of *if* with the disjunctive meaning of *either ... or* (Eastwood, 1994; Halliday, 2014, p. 453; Quirk et al., 1985, pp. 999 & 1100).

Hypotaxis: causal-conditional; alternative condition; concessive

(151) *Yo sé si tú matas a personas indefensas o no.*

*I know **whether** you kill defenseless people or not.*

Example	Structure
1	V1 + <i>whether</i> + V2 finite
2	V1 + <i>whether</i> + V2 <i>to</i> -infinitive form
3	V1 + <i>whether</i> + V2 <i>-ing</i> form
4	V1 + <i>whether</i> + V2 <i>-ed</i> form

Table 55 English R2 Seven o'clock <<causal-conditional: alternative condition: concessive>> Structures.

The categories <<quién>> (152), <<cómo>> (153), <<cuándo>> (154), <<dónde>> (155) and <<por qué>> (156) are different since these types of subordinators appear in nominal relative clauses which resemble *wh*-exclamatory and *wh*-interrogative clauses (Pérez, 2018; Quirk et al., 1985).

(152) *Ella adivina **quién** viene a cenar esta noche.*

*She guesses **who** is coming to dinner tonight.*

(153) *Yo sé **cómo** vienes cada día al instituto.*

*I know **how** you come to school every day.*

(154) *Tú siempre recuerdas **cuándo** es mi cumpleaños.*

*You always remember **when** my birthday is.*

²³ *Si* appears at Five o'clock, Seven o'clock and Nine o'clock in different <<meaning>> categories.

- (155) *Tú sabes **dónde** está la maleta con el dinero.*
*You know **where** the suitcase with the money is.*
- (156) *Nosotros entendemos **por qué** llegas tarde a clase todos los días.*
*We understand **why** you are late for class every day.*

Wh-interrogative clauses can function as a subject (157), direct object (158), subject complement (159), appositive (160), adjectival complement (161) and a prepositional complement (162). This variety of subordinate clauses resemble *wh*-questions semantically in that they leave a gap of unknown information represented by the *wh*-element (Quirk et al., 1985, pp. 1050-1052).

- (157) ***How the book will sell** depends on the reviewers.*
- (158) *I can't imagine **what they want with your address**.*
- (159) *The problem is **who will water my plants when I am away**.*
- (160) *Your original question, **why he did not report it to the police earlier**, has not yet been answered.*
- (161) *I'm not sure **which she prefers**.*
- (162) *They did not consult us on **whose names should be put forward**.*

Wh-exclamatory clauses tend to function as an extraposed subject (163), direct object or prepositional complement (164). Like independent exclamative clauses, the exclamative element is formed with *what* as predeterminer in a noun phrase and *how* as intensifier of an adjective, adverb, or clause. Furthermore, the exclamative element is positioned initially regardless of its normal position in a declarative clause (Quirk et al., 1985, pp. 1055-1056).

- (163) *It's incredible **how fast she can run**.*
- (164) *I remember **what a good time I had at your party**.*

Wh-interrogative clauses and *wh*-exclamatory clauses, although considered a form of subordination, appears to have a commonality with paratactic extension through clarification. The secondary clause clarifies the thesis of the primary clause by providing background information in the way of an explanatory comment or explanation (165) (Halliday, 2014).

Parataxis: extension: clarification

(165) ||| *They used to work over here*; || *that's [how they met]*.

The English equivalent to <<quién>>, <<cómo>>, <<cuándo>>, <<dónde>> and <<por qué>> are *wh*-interrogative clauses and *wh*-exclamatory clauses that happen to be categorized in parataxis in extension through clarification. Therefore, the parameters for these elements would be a V1 plus *wh*-interrogative/exclamatory clause with the V2 as seen in Table 56.

Example	Structure
1	V1 + <i>wh</i> -interrogative/exclamatory clause + V2

Table 56 English R2 Seven o'clock Parataxis <<extension: clarification>> Structures.

3.3.8 Eight o'clock

R2 at Eight o'clock includes three subordinator groups involving <<porque>>, <<así>> and <<consecuencia>> which introduce normal form V2s. The first category involving <<porque>> is composed of *puesto que* (166), *ya que* (167), *como*²⁴ (168) and *dado que* (169) (Pérez, 2018, p. 149). The apparent translation for this word grouping is *since* in English. *Since* is a simple subordinator for reason clauses that expresses the speakers inference of a connection and introduces a finite form (Quirk et al., 1985, pp. 1104).

Hypotaxis and expansion: enhancement: causal-conditional: cause: reason: finite verb

(166) *Tengo hambre **puesto que** no he comido desde ayer.*

*I am hungry **since** I have not eaten since yesterday.*

(167) *Tiene mucho dinero **ya que** trabaja durante muchísimas horas.*

*He has a lot of money **now that/considering that** he works very long hours.*

(168) ***Como** soy pequeño, voy en el asiento de atrás.*

As I am small, I ride in the back seat.

(169) *Comemos plátanos **dado que** tenemos hambre.*

*We eat bananas **given that** we are hungry.*

²⁴ *Como* appears at Three o'clock, Eight o'clock and Nine o'clock in different <<meaning>> categories.

Ya que is also translated into *now that* as well as *considering that*, *como*²⁵ into *as* and *dado que* into *given that*. Halliday (2014) defines the below examples as being classified under causal conditional in condition through cause and reason (p. 477). Table 57 exhibits the beforementioned subordinators and the structure they have in English.

Example	Structure
1	V1 + <i>since/now that/as/given that</i> + V2 finite

Table 57 English R2 Eight o'clock <<causal conditional: condition: cause: reason>> Structures.

The second category representing <<así>> includes the phrases *de forma/ manera/ modo que* (170) (Pérez, 2018, p. 149). This translates to *in a way* in English which *way* is considered a manner adjunct. It is a circumstantial element that appears as a prepositional phrase in which *in* or *with* is the head of the nominal group that expresses manner commonly illustrated by *way* (Quirk et al., 1985, p. 557).

Hypotaxis and expansion: enhancement: manner: comparison: finite

(170) *Aloísio viste de forma que parece estúpido para que yo pase vergüenza.*

Aloísio dresses in a way that looks stupid so that I will be embarrassed.

This form is classified in hypotaxis and expansion through enhancement under manner and comparison introducing a finite verb form (Halliday, 2014, pp. 319 & 477). Table 58 demonstrates the single form that appears under <<manner comparison>> with a preposition plus *the way* introducing a finite secondary verb form.

Example	Structure
1	V1 + preposition + <i>the way</i> + V2 finite

Table 58 English R2 Eight o'clock <<manner: comparison>> Structures.

The final category at this hour takes the meaning of <<consecuencia>> and is made up of *de forma/manera/modo que* (171) (Pérez, 2018, p. 149). Although certain translations of the *de forma/manera/modo que* with a subjunctive V2 can be translated as *in order to* or *so that*, at Seven o'clock using a V2 normal form the resulting translation is *therefore* which is a resultative conjunctive adjunct (Quirk et al., 1985, p. 635). Halliday (2014) notes the

²⁵ *Como* appears at Three o'clock, Eight o'clock and Nine o'clock in different <<meaning>> categories.

combination of *and... therefore* falls into paratactic enhancement. However, without the inclusion of *and*, the resultative conjunct *therefore* cannot be classified under hypotaxis nor parataxis and is outside the parameters of this study.

- (171) *Marta es egoísta, **de forma que** no le voy a regular nada en la fiesta.*
*Marta is selfish, **therefore** I'm not going to give her anything at the party.*

3.3.9 Nine o'clock

R2 at Nine o'clock is made up of the categories <<si>>, <<consecuencia>>, cuánto>> and <<para>> that can only combine with a V2 in the subjunctive form (Pérez, 2018, pp. 149-150). The first classification embodying <<si>> appears with a grammatical pause and is composed of *como/if*²⁶. It is important to note *if* appears in this category of subordinators, yet it has already been covered at Five o'clock. *Siempre y cuando/as long as* or *provided that* (172), and *siempre que*²⁷/*provided that* (173) also appear in casual-conditional under positive condition as a finite subordinator at Nine o'clock. It is important to note that there is significant overlap when these Spanish subordinators are translated to English. As with *siempre y cuando* being translated to *as long as* or *provided that*, two English subordinators can be applied for the single Spanish one.

Hypotaxis and expansion; enhancement: casual-conditional: condition: positive: finite

- (172) *El viaje sigue en pie **siempre y cuando** no nieve.*
*The trip continues **as long as/provided that** there is no snow.*
 (173) *Trabajaré los domingos **siempre que** ellos me den más dinero.*
*I will work on Sundays **as long as/provided that** they give me more money.*

The appearance of *as long as* and *provided that* add another component to the list of subordinators in the conditional category. Both subordinators combine condition with time and are considered to reflect a positive state of affairs in the same manner as *if*. Unlike *if*, *as long as* and *provided that* introduce a finite dependent clause yet cannot introduce non-finite forms as seen in Table 59 (Quirk et al., 1985). These conditional elements are found under

²⁶ *Como* appears at Three o'clock, Eight o'clock and Nine o'clock in different <<meaning>> categories.

²⁷ *Siempre que* appears at Three o'clock and Nine o'clock in different <<meaning>> categories.

hypotaxis and expansion under enhancement through casual-conditional in positive condition (Halliday, 2014).

Example	Structure
1	V1 + <i>as long as/provided that</i> + V2 finite

Table 59 English R2 Nine o'clock <<casual-conditional: condition: positive: finite>> Structures.

Within Nine o'clock there is another subordinator group represented by *salvo que/unless* (174), *a menos que/unless* (175) and *a no ser que/unless* (176). All three Spanish subordinators are translated to *unless* which signifies a negative state and is categorized as casual-conditional through negative condition (Halliday, 2014; Pérez, 2018).

Furthermore, *unless* can introduce finite and non-finite clauses which are *-ing* forms and *-ed* forms as seen in Table 60 (Quirk et al., 1985, pp. 1088-1091).

Hypotaxis and expansion: enhancement: casual-conditional: condition: negative: finite

(174) *Te compraré estas flores salvo que huelan mal.*

I will buy you these flowers unless they smell bad.

(175) *A menos que me pagues el dinero, no puedes irte.*

Unless you pay me the money, you can't leave.

(176) *Iré solo, a no ser que tú vengas conmigo.*

I'll go alone, unless you come with me.

Example	Structure
1	V1 + <i>unless</i> + V2 finite
2	V1 + <i>unless</i> + V2 <i>-ing</i> form
3	V1 + <i>unless</i> + V2 <i>-ed</i> form

Table 60 English R2 Five o'clock <<casual-conditional: condition: negative >> Structures.

The second group involving <<consecuencia>> is made up of *de ahí que* (177). The English translation is *so* and *so that* which are subordinators that introduce clauses of result. Clauses of result can overlap with purpose. Nevertheless, result clauses express an achieved action. When *that* is removed from the equation, it is difficult to distinguish the resultative *so* from the one that means asyndetic condition. Accordingly, result clauses introduced by *so* must be separated by a comma (Quirk et al., 1985, p. 1109).

Hypotaxis and expansion: enhancement: casual-conditional: cause: result: finite

(177) *Mis amigos roban en el súper, **de ahí que** yo no vaya con ellos.*

*My friends steal at the supermarket, **so** I don't go with them.*

Halliday (2014) includes *so* in hypotaxis and expansion through enhancement representing a positive casual-conditional introducing cause and result with a finite form.

Table 61 shows how *so* introduces a finite dependent clause.

Example	Structure
1	V1 + # <i>so</i> + V2 finite

Table 61 English R2 Nine o'clock <<casual-conditional: cause: result>> Structures.

The third category <<cuánto>> contains *como si* (178) with a grammatical pause (Pérez, 2018). The English translation is *as if* and it is a complex subordinator which is a clause of comparison. *As if* expresses dependency and materializes under hypotaxis and expansion through enhancement under manner by comparison (Halliday, 2014).

Hypotaxis and expansion: enhancement: manner: comparison: finite

(178) *Sagrario come patatas **como si** fuera una vaca holandesa.*

*Sagrario eats potatoes **as if** he were a Dutch cow.*

As seen in Table 62, *as if* can introduce secondary clauses that are finite (1), non-finite *-ed* (2), non-finite *-ing* (3) and non-finite *to*-infinitive (4) (Quirk et al., 1985).

Example	Structure
1	V1 + # <i>as if</i> + V2 finite
2	V1 + # <i>as if</i> + V2 <i>-ed</i> clause
3	V1 + # <i>as if</i> + V2 <i>-ing</i> clause
4	V1 + # <i>as if</i> + V2 <i>to</i> -infinitive

Table 62 English R2 Nine o'clock <<manner: comparison>> Structures.

The final grouping in R2 at Nine o'clock indicates <<para>> through *de forma/manera/modo que* (179) (Pérez, 2018). Like the forms covered at Eight o'clock, the translation of *de forma/manera/modo que* to *in a way* falls outside the criteria of this study since it is not included in either hypotaxis or parataxis.

- (179) *Raúl viste con ropa ridícula **de forma que** podamos verle a distancia.*
*Raúl dresses in ridiculous clothes **in a way** that we can see him from a distance.*

3.3.10 Ten o'clock

R2 at Ten o'clock materializes as three groups being <<cuando>><<después>>, <<porque >> and <<si>> (Pérez, 2018). The first category <<cuando>><<después>> encompasses two words being *al*²⁸ (180) and *tras* (181). The second category signifying <<porque >> has one subordinator *al*. The commonality that interpretations of *al* share is the English translation to *on/upon* (Arús et al., 2010, p. 49). The crux of this interpretation is that the action in the main and dependent clauses must happen at the same time. *On/upon* are non-finite prepositions that are in hypotaxis and expansion under temporal enhancement with both main and dependent clauses happening at the same point in time (Halliday, 2014, p. 477).

Hypotaxis and expansion: temporal: same time: non-finite preposition

- (180) ***Al** no tener suficiente dinero para el café, Pedro se puso colorado.*
***On/Upon** not having enough money for coffee, Pedro blushed.*
- (181) *Marina siempre espera a Juaco **al** terminar el trabajo.*
*Marina always waits for Juaco **on/upon** finishing work.*

Table 63 displays the two constructions found in hypotaxis and expansion under temporal enhancement with both main and dependent clauses happening at the same point in time involving *on/upon*. The difference in (1) is the prepositions introduce a *-ing* form followed by a grammatical pause then a secondary clause while in (2) there is no grammatical pause and *on/upon* prompts the *-ing* form in the secondary clause.

Example	Structure
1	<i>On/upon</i> V1 <i>-ing</i> # + V2
2	V1 + <i>on/upon</i> V2 <i>-ing</i> form

Table 63 English R2 Ten o'clock in <<temporal: same time: [point]>> Structures.

²⁸ *Al* appears at Ten o'clock in different <<meaning>> categories.

The translation of *tras* in English appears as the non-finite preposition *after* (182). To avoid repetition, this type of temporal different time later construction is detailed at Six o'clock in this chapter.

- (182) *Sale rápidamente de la comisaría **tras** encontrar una pista del asesino.*
*He quickly leaves the police station **after** finding a clue to the killer.*

The final category in R2 at Ten o'clock involves <<si>> and the single subordinator *de* materializes in (183) as *if* in English. *If* is a subordinator which signals a casual-conditional positive condition and is detailed at Five o'clock in this chapter. Therefore, a further description will not be carried out to avoid repetition.

- (183) ***De** tener problemas, te llamaré.*
***If** I have problems, I will call you.*

3.4 The Meaning-Based Framework

The concluding section of this chapter details the theoretical origin of the Meaning-based Framework (MBF) which is the culmination of the previously described methodology as well as the qualitative cross-linguistic analysis carried out between English and Spanish through the focal points of SFG, MR R2 and CEFR proficiency levels. The following presentation of measurements in the MBF are described through the concepts of extension, enhancement and elaboration.

The framework is organized to give a researcher the ability to analyze different segments of the structure. For example, one could investigate general categories like Level 1 taxis. The next step would be dissecting Level 2 which would be expansion through extension, enhancement and elaboration. Level 3 details logico-semantics relations through MBCs. Lastly, Level 4 gives an account of individual subordinators and coordinators which make up the MBU category. The idea of this multilevel framework is to give a researcher the ability to analyze from broad to specific or specific to broad within the parameters of the CEFR proficiency scale and the focal lens of the MR R2.

3.4.1 Extension

The MBF for hypotaxis and extension includes seven separate measures (Table 64). In this case, subordinators extend the meaning of another clause by adding something new to it (Halliday, 2014). As there are MBUs that are repeated in enhancement, it is crucial to remember the definition of extension and its application to subordination in its own context.

#	MBC	MBSC	MBU	MR R2 Hour
1	addition	additive: positive	<i>whereas</i> (finite) <i>mientras que</i> (finite)	Eight o'clock <<obstáculo>>
2	addition	additive: positive	<i>besides</i> (non-finite) <i>además de que</i> (finite/subjunctive)	Six o'clock <<y>> <<incluso>>
3	addition	additive: positive	<i>apart from</i> (non-finite) <i>aparte de que</i> (finite/subjunctive)	Six o'clock <<y>>
4	addition	adversative	<i>without</i> (non-finite) <i>sin que</i> (subjunctive/infinitive)	One o'clock <<así>> <<no con>>
5	variation	replative	<i>instead of</i> (non-finite) <i>en vez de</i> (subjunctive/infinitive)	One o'clock <<así>> <<sustitución>>
6	variation	subtractive	<i>except that</i> (finite) <i>salvo que</i> (subjunctive)	Nine o'clock <<si>> <<si no>>
7	alternation	-	<i>if ... not (...then)</i> (finite) <i>si ... no</i> (subjunctive)	Five o'clock <<si>>

Table 64 MBF for Hypotaxis and Extension.

In the case of parataxis and extension, the combination of 1st clause + 2nd clause leads to the concept of coordination (Halliday, 2014). There are eight different measurements in this grouping (Table 65). Since these types of structures do not cause a hypotactic interaction between a V1 and a V2, they are located outside of the parameters of R2. MBUs under extension appear in enhancement with additional elements, yet the definition between extension and enhancement clarifies how each MBU is defined when overlapping exists.

#	MBC	MBSC	MBU	MR R2 Hour
1	addition	additive: positive	(both ...) <i>and</i> (ambos ...) <i>y</i>	-
2	addition	additive: positive	<i>not only ... but also</i> <i>no solo... sino también</i>	-
3	addition	additive: negative	(neither ...) <i>nor</i> (ni...) <i>ni</i>	-
4	addition	adversative	(and) <i>yet</i> (y) <i>aun así</i>	-
5	addition	adversative	<i>but</i> <i>pero</i>	-
6	variation	subtractive	<i>except</i> <i>excepto</i>	-
7	variation	subtractive	<i>but</i> <i>pero</i>	-
8	alternation	-	<i>or</i> <i>o</i>	-

Table 65 MBF for Parataxis and Extension.

3.4.2 Enhancement

Hypotaxis and enhancement is the category in the MBF with the highest number of measures totaling 45 indices and is the largest representative sample of MR R2 (Table 66). Enhancement gives way to traditional adverbial subordination in which $\alpha \times \beta$ (Halliday, 2014). The cross-linguistic comparison of subordinators was made so that the closest meaning between English and Spanish was matched for accuracy. However, the use of any given subordinator depends on personal preference or pragmatic function at the moment of utterance.

#	MBC	MBSC	MBU	MR R2 Hour
1	temporal	same time: [extent]	<i>as</i> (finite) <i>a medida que</i> (finite/subjunctive)	Three o'clock <<cuando>> <<acción larga>>
2	temporal	same time: [extent]	<i>while</i> (finite) <i>mientras</i> (finite/subjunctive)	Three o'clock <<cuando>> <<acción simultánea>>
3	temporal	same time: [point]	<i>when</i> (finite/ non-finite conj.) <i>cuando</i> (finite/subjunctive)	Three o'clock <<cuando>>
4	temporal	same time: [point]	<i>as soon as</i> (finite) <i>tan pronto como</i> (finite/subjunctive)	Three o'clock <<cuando>> <<después>>
5	temporal	same time: [point]	<i>the moment</i> (finite) <i>en el momento en que</i> (finite/subjunctive)	Three o'clock <<cuando>> <<después>>
6	temporal	same time: [point]	<i>upon</i> (non-finite prep.) <i>al</i> (infinitive)	Ten o'clock <<cuando>> <<después>>
7	temporal	same time: [spread]	<i>every time</i> (finite) <i>cada vez que</i> (finite/subjunctive)	Three o'clock <<cuando>> <<acción larga>>
8	temporal	different time: later	<i>after</i> (finite/ non-finite prep.) <i>después de (que)</i> (finite/subjunctive/ infinitive)	Six o'clock <<cuando>> <<después>>
9	temporal	different time: later	<i>since</i> (finite/non-finite conj.) <i>desde que</i> (finite/subjunctive)	Three o'clock <<cuando>> <<acción larga>>
10	temporal	different time: later	<i>once</i> (finite/non-finite conj.) <i>una vez (que)</i> (finite/subjunctive)	Three o'clock <<cuando>> <<después>>
11	temporal	different time: earlier	<i>before</i> (finite/non-finite) <i>antes de que</i> (subjunctive/infinitive)	One o'clock <<cuando>> <<antes>>
12	temporal	different time: earlier	<i>until</i> (finite/non-finite conj.) <i>hasta que</i> (finite/subjunctive)	Three o'clock <<cuando>> <<acción larga>>
13	spatial	same place [point]	<i>where</i> (finite) <i>donde</i> (finite/subjunctive)	Three o'clock <<donde>>

14	manner	means	<i>by</i> (means of) (non-finite prep.) <i>por</i> (infinitive)	Two o'clock <<porque>>
15	manner	comparison	<i>as if</i> (finite) <i>como si</i> (subjunctive)	Nine o'clock <<cuánto>> <<como>> <<como si>>
16	manner	comparison	<i>as</i> (finite) <i>como</i> (finite/subjunctive)	Three o'clock <<así>>
17	causal- conditional:	cause: reason	<i>because</i> (finite) <i>porque</i> (finite/subjunctive/ infinitive)	Two o'clock <<porque>>
18	causal- conditional:	cause: reason	<i>as</i> (finite) <i>ya que</i> (finite)	Eight o'clock <<porque>>
19	causal- conditional:	cause: reason	<i>since</i> (finite) <i>puesto que</i> (finite)	Eight o'clock <<porque>>
20	causal- conditional:	cause: reason	<i>in case</i> (finite) <i>en caso de (que)</i> (finite/subjunctive)	One o'clock <<si>>
21	causal- conditional:	cause: reason	<i>seeing that</i> (finite) <i>visto que</i> (finite)	Eight o'clock <<porque>>
22	causal- conditional:	cause: reason	<i>given that</i> (finite) <i>dado que</i> (finite)	Eight o'clock <<porque>>
23	causal- conditional:	cause: reason	<i>considering</i> (finite) <i>considerando que</i> (finite)	Eight o'clock <<porque>>
24	causal- conditional:	cause: reason	<i>with</i> (non-finite prep.) <i>con que</i> (finite/subjunctive)	One o'clock <<si>>
25	causal- conditional:	cause: reason	<i>as a result of</i> (non-finite prep.) <i>como consecuencia de</i> (finite/subjunctive/ infinitive)	Six o'clock <<obstáculo>>
26	causal- conditional:	cause: reason	<i>because of</i> (non-finite prep.) <i>a causa de</i> (finite/subjunctive/ infinitive)	Two o'clock <<porque>>

27	causal- conditional:	cause: reason	<i>in case of</i> (non-finite prep.) <i>en caso de</i> (non-finite)	One o'clock <<si>>
28	causal- conditional:	cause: reason	<i>due to</i> (non-finite prep.) <i>debido a que</i> (finite/subjunctive/ infinitive)	Two o'clock <<porque>>
29	causal- conditional:	cause: reason	<i>thanks to</i> (non-finite prep.) <i>gracias a que</i> (finite/subjunctive/ infinitive)	Two o'clock <<porque>>
30	causal- conditional:	purpose	<i>so that</i> (finite) <i>para que</i> (finite/subjunctive)	One o'clock <<para>>
31	causal- conditional:	purpose	<i>(in order to/so as to) to</i> (non-finite prep.) <i>para</i> (infinitive)	One o'clock <<para>>
32	causal- conditional:	purpose	<i>with the aim of</i> (non-finite prep.) <i>con el objetivo de</i> (finite/subjunctive)	One o'clock <<para>>
33	causal- conditional:	purpose	<i>in exchange for</i> (non-finite prep.) <i>a cambio de que</i> (finite/subjunctive)	One o'clock <<si>> <<solo si>>
34	causal- conditional:	purpose	<i>with the intention of</i> (non- finite prep.) <i>con la intención de que</i> (finite/subjunctive)	One o'clock <<para>>
35	causal- conditional:	condition: positive	<i>if</i> (finite/non-finite conj.) <i>si</i> (subjunctive)	Five o'clock <<si>>
36	causal- conditional:	condition: positive	<i>provided that</i> (finite) <i>siempre que</i> (subjunctive)	Nine o'clock <<si>>
37	causal- conditional:	condition: positive	<i>as long as</i> (finite) <i>siempre y cuando</i> (subjunctive)	Nine o'clock <<si>>
38	causal- conditional:	condition: positive	<i>so long as</i> (finite) <i>con tal de que</i> (finite/subjunctive)	One o'clock <<si>>
39	causal- conditional:	condition: positive	<i>on the condition that</i> (finite) <i>a condición de que</i> (finite/subjunctive)	One o'clock <<si>>
40	causal- conditional:	condition: negative	<i>unless</i> (finite/non-finite conj.) <i>a menos que</i> (subjunctive)	Nine o'clock <<si>> <<si no>>

41	causal- conditional:	condition: negative	<i>without</i> (non-finite prep.) <i>sin que</i> (subjunctive)	Nine o'clock <<si>> <<si no>>
42	causal- conditional:	condition: concessive	<i>even if</i> (non-finite conj.) <i>aun si</i> (subjunctive)	Five o'clock <<si>>
43	causal- conditional:	condition: concessive	<i>even though</i> (finite) <i>a pesar de que</i> (finite/subjunctive/ infinitive)	Six o'clock <<obstáculo>>
44	causal- conditional:	condition: concessive	<i>although</i> (finite) <i>aunque</i> (finite/subjunctive/ infinitive)	Six o'clock <<obstáculo>>
45	causal- conditional:	condition: concessive	<i>despite</i> (non-finite prep.) <i>pese a que</i> (finite/subjunctive) infinitive)	Six o'clock <<obstáculo>>

Table 66 MBF for Hypotaxis and Enhancement.

The MBF for parataxis and enhancement contains fewer measures than hypotaxis with only 12 in total (Table 67). Parataxis through enhancement is the notation of 1 x2 which incorporates a circumstantial feature to the notion of coordination (Halliday, 2014).

#	MBC	MBSC	MBU	MR R2 Hour
1	temporal	same time	(and) <i>meanwhile</i> (y) <i>mientras tanto</i>	-
2	temporal	different time: later	(and) <i>then</i> <i>y luego</i>	-
3	temporal	different time: later	<i>and + afterwards</i> <i>y después</i>	-
4	temporal	different time: earlier	<i>and + before</i> <i>y + antes</i>	-
5	temporal	different time: earlier	<i>but + before</i> <i>pero + antes</i>	-
6	spatial	same place	<i>and there</i> <i>y allí</i>	-
7	manner	means	<i>and + in that way</i> <i>y de esa manera</i>	-
8	manner	comparison	<i>thus</i> <i>así</i>	-
9	causal-conditional	cause: reason	(and) <i>so</i> (y) <i>por lo tanto</i>	-
10	causal-conditional	condition: positive	<i>and + in that case</i> <i>y en este caso</i>	-
11	causal-conditional	condition: concessive	(and) <i>yet</i> <i>y aún así</i>	-
12	causal-conditional	condition: concessive	<i>pero aun así</i> <i>but + nevertheless</i>	-

Table 67 MBF for Parataxis and Enhancement.

3.4.3 Elaboration

Elaboration through hypotaxis is the notation in which a =b. This category is composed of non-defining relative clauses. An important element that must be highlighted in this category is that the non-defining relative clause must be enclosed by commas. Table 68 gives a summary of the five measure which constitute the MBF for this category.

#	MBC	MBSC	MBU	MR R2 Hour
1	non-defining relative clause	-	<i>which</i> (finite) <i>que</i> (finite)	Four o'clock CI
2	non-defining relative clause	-	<i>that</i> (finite) <i>que</i> (finite)	Four o'clock CI
3	non-defining relative clause	-	<i>who</i> (finite) <i>quien</i> (finite)	Four o'clock CI
4	non-defining relative clause	-	<i>when</i> (finite) <i>cuando</i> (finite)	Four o'clock CI
5	non-defining relative clause	-	<i>where</i> (finite) <i>donde</i> (finite)	Four o'clock CI

Table 68 MBF for Hypotaxis and Elaboration.

While elaboration through hypotaxis includes non-defining relative clauses, parataxis deals with the notation of 1 =2. This happens through exposition, exemplification and clarification. Except for R2 at Seven o'clock, the majority of measures in this category do not meet the conditions of R2. There are a total of 13 measures in this category (Table 69). As the list for measures could extend even further, the indices included in this study are the ones found in *Introduction to Functional Grammar* (2014). A matter for future research may be to focus on this category and greatly extend number of measures.

#	MBC	MBSC	MBU	MR R2 Hour
1	exposition	-	<i>in other words</i> <i>en otras palabras</i>	-
2	exposition	-	<i>that is to say</i> <i>es decir</i>	-
3	exposition	-	<i>I mean</i> <i>quiero decir</i>	-
4	exemplification	-	<i>such as</i> <i>tal como</i>	-
5	exemplification	-	<i>for example</i> <i>por ejemplo</i>	-
6	exemplification	-	<i>in particular</i> <i>en particular</i>	-

7	clarification	-	<i>how</i> <i>cómo</i>	Seven o'clock
8	clarification	-	<i>when</i> <i>cuándo</i>	Seven o'clock
9	clarification	-	<i>what</i> <i>qué</i>	Seven o'clock
10	clarification	-	<i>in fact</i> <i>de hecho</i>	-
11	clarification	-	<i>actually</i> <i>de verdad</i>	-
12	clarification	-	<i>indeed</i> <i>en efecto</i>	-
13	clarification	-	<i>at least</i> <i>por lo menos</i>	-

Table 69 MBF for Parataxis and Elaboration.

4. Corpus Study: Application of the Meaning-Based Framework

4.1 Introduction

Section 2.2 and 2.3 overviewed how L2 writing has played a significant role in EFL research and how complexity has become a salient construct to assess and investigate L2 writing performance and development (Lahuerta Martínez, 2018a). Complexity is composed of several sub-constructs and components which can be assessed on their own (Norris & Ortega, 2009). Syntactic complexity is a component of linguistic complexity and is generally defined as the range of sophistication of grammatical resources exhibited in language production.

Syntactic complexity gauges the capacity to use language in ever more mature and skillful ways that involve a full range of linguistic resources offered by a given grammar to fulfill communicative goals (Ortega, 2015, p. 82). Subordination is thought to be a valid indicator of L2 proficiency and is often used to measure syntactic complexity, describe performance and benchmark L2 development (Chen et al., 2021, p. 811).

A wide range of studies have incorporated syntactic complexity as their central focus using a broad range of indices while on an individual basis each study applied few measures to relatively small amounts of data (Bulté & Housen, 2012; Lu, 2011; Lu & Ai, 2015; Ortega, 2003; Wolfe-Quintero et al., 1998). In general, the overall scope of syntactic complexity measures was not created equal and some of them redundantly measure the same thing (Norris & Ortega, 2009). Recent studies have investigated how traditional versus fine-grained syntactic complexity indices could predict second language raters' quality ratings (Kyle & Crossley, 2018; Zhang & Lu, 2022).

There are studies looking into different facets of SFL in Chinese-English translation (Li & Yu, 2021), Spanish heritage language learners (Achugar & Colombi, 2009), development of clause complexity in children with a language impairment (Aray et al., 2023) and grammatical metaphor (Velázquez-Mendoza, 2015). Nonetheless, there appears to be a lack of cross-linguistic studies researching frequencies of parataxis and hypotaxis, expansion and logico-semantic relations incorporating L2 Spanish and English writers in conjunction

with CEFR levels A1-C2. Additionally, there are no known studies which include the fine-grained MBU indices of the proposed MBF to measure L2 acquisition and L1 to L2 transfer.

This investigation will carry out a corpus study to explore how parataxis and hypotaxis signals L2 cognitive development when analyzed through the focal lens of the MBF. Broadly defined, a corpus is a collection of naturally occurring spoken or written data in electronic form which is selected by external criteria to represent language, a language variety or domain of language use. Corpus linguistics has revolutionized language theory and description by placing special emphasis on the frequency of words and patterns, the way words combine in collocations, the connection between grammar and lexis and the ways in which situational factors act as explanatory variables (Phakiti et al., 2018).

Corpus linguistics is based on empirical and inductive forms of analysis which rely on real-world instances of language use to be able to derive rules or explore trends about the way people produce language. Theoretical justification for corpus linguistics is that humans do not make accurate judgements. They rely on cognitive or social bias in language use. Computers calculate frequency and carry out statistical tests quickly and accurately to give researchers access to linguistic patterns and trends (Litosseliti, 2010, pp. 94-95). The basic idea is that frequency of form and meaning is the most reliable predictor of what can be usefully taught at different points in the learning process (Phakiti et al., 2018, p. 360). The current study takes a corpus-based approach in which a selective corpora is used to test a specific research question.

A key concept of a corpora is that it be of a large enough representative sample to reveal something about a frequency of a linguistic phenomenon (Phakiti et al., 2018, p. 360). In order to ensure an accurate representative sample, this study employs L2 learner corpora. Learner corpora are electronic collections of natural or near natural data produced by foreign or second language learners and are designed based on a strict criterion. Learner corpora have allowed for the opportunity to access a wide range of features typical of learners at different proficiency levels (Granger, 2017). The three learner corpora used in this study and their individual attributes will be covered in the methodology section of this chapter.

The objectives of the study at hand are multi-faceted, yet the central focus is to measure the frequencies of taxis, expansion, logico-semantic relations and MBUs using the

MBF. Moreover, the MBF will be used to test the developmental sequence of syntactic interlanguage complexification hypothesis proposed by Norris and Ortega (2009). The issue at hand, is if coordination prevails at the beginning A1-A2 levels (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012) while increased subordination takes the place of coordination at the intermediate level of B1 and B2 and subsides at the advanced levels of C1 and C2 when grammatical metaphor increases (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). While this study is not measuring grammatical metaphor, the main objective is to research coordination and subordination trends in English and Spanish L2 written texts.

The objective of the MBF is not only to analyze taxis, but to look at the phenomenon through the various focal points of expansion through MBCs which represent logico-semantic markers and MBUs. The top-down method starts by investigating taxis and ending with MBUs. At every level the methodology imposes more fine-grained meaning-based measures to paint a clearer picture of L2 acquisition and L1 to L2 transfer through CEFR proficiency levels. The following chapter will include the methodology section in which the corpora and data, as well as the technical details and evaluation will be reviewed. The analysis of the results will then be reviewed followed by the discussion. The last section in this chapter will draw conclusions from the research.

4.2 Methodology

4.2.1 Corpora

This study adheres to the definition of corpus linguistics in the literal linguistic interpretation. Corpus linguistics is seen as the investigation of linguistic research questions based upon the complete and systematic analysis of the conditional distribution of linguistic phenomena in a corpus (Stefanowitsch, 2020). The corpora used for this study were chosen to incorporate large amounts of information in which learners produced authentic data generated through real life activity for the purpose of communication. The data for this study comes from the three different corpora to ensure an equal representation.

The second version of the *Education First-Cambridge Open Language Database* provided access to L2 English texts written by L1 Spanish. The *Corpus escrito del español L2* and *Corpus de aprendices de español* provided data from L2 Spanish texts written by L1 English learners. The three corpora were deemed necessary to provide a large enough representative sample for cross-linguistic analysis. By using the previously mentioned L2 learner corpora, this study aimed to include a representative sample, not only through sheer size, but also through language variety, genre, register and style to gauge the phenomena in question (Stefanowitsch, 2020). All the relevant variables were not fully controlled, since the constraints of amalgamating a word count representative of a large data set within the parameters of CEFR proficiency scales was the principal focus used to uncover frequency patterns as with previous studies involving the EFCAMDAT (Alexopoulou et al., 2015) and the EFCAMDAT2 (Chen et al., 2021). Nevertheless, comparative variability within the two Spanish corpora was not a crucial factor since the CEDEL2 was designed to complement the CAES as an existing resource for L2 Spanish acquisition (Lozano, 2022). The following section details the characteristics of each of the three corpora used in this study.

4.2.1.1 EFCAMDAT2

The second release of the *Education First-Cambridge Open Language Database* (EFCAMDAT2)²⁹ resulted in an 83-million-word database consisting of over one million assignments written by 174,743 learners from all over the world who were enrolled between 2011 and 2015 in the online school Englishtown (now English Live) run by *Education First*. The second version is considerably bigger than the first release which contained only 551,036 assignments with a total of about 33 million words from 84,864 learners (Tan & Römer, 2022).

The global reach of *English First* has given researchers a resource with significant diversity including 128 activities in which learners produce scripts across sixteen teaching levels aligned with CEFR proficiency levels (Table 70). A script is defined as a writing piece a learner submits as an answer to a writing task. There are two issues which raise

²⁹ The EFCAMDAT2 website appeared to go offline in July 2023 and appears to not be publicly available at the time this section was revised in May 2024. That data for the following study had already been downloaded and saved. However, any reference or access to the EFCAMDAT2 website appears to not be an option, but information from published sources about said website are.

methodological concerns. The first challenge is the size of the database. Learner corpora is annotated manually for linguistic information thus requiring the need for automated data processing thus calling for automated natural language processing (Alexopoulou et al., 2015).

EF Teaching Levels	CEFR Levels
1-3	A1
4-6	A2
7-9	B1
10-12	B2
13-15	C1
16	C2

Table 70 Correspondence between EF Teaching Levels and CEFR (Alexopoulou et al., 2015).

The second challenge relates to the nature of EFCAMDAT2 data regarding the number of variables. A certain number of variables which are standardly controlled for by research design are not easily recoverable in learner corpus compiled in *Englishtown* activities. Firstly, obtaining scripts from students provides a set of circumstances in which the progress rates of each student vary in the year-long course. Second, the tasks, topics and prompts of the different topics (Table 71) across proficiency levels are developed for teaching, not creating a representative set of writing tasks for research purposes. Lastly, written texts are produced through computer assisted learning which puts a limit on words used in texts thus leading to shorter texts (Alexopoulou et al., 2015).

Level	Writing Topic	Level	Writing Topic
1:1	Introducing yourself by email	7:1	Giving instructions to play a game
1:3	Writing an online profile	8:2	Reviewing a song for a website
2:1	Describing your favorite day	9:7	Writing an apology email
2:6	Telling someone what you are doing	11:1	Writing a movie review
2:8	Describing your family's eating habits	12:1	Turning down an invitation
3:1	Replying to a new pen pal	13:4	Giving advice about budgeting
4:1	Writing about what you do	15:1	Covering a news story
6:4	Writing a resume	16:8	Researching a legendary creature

Table 71 Examples of Writing Topics/Unit Number Across Teaching Levels (Alexopoulou et al., 2015).

An important factor regarding EFCAMDAT2 is that it lacks detailed information on learners' L1. Alexopoulou et al. (2015) used background L1 information on Brazilians,

Chinese, Russian, Mexican and Germans learners by matching nationality and country of residence. It is believed this combination provides a reasonable match when the L1 is the dominant or official language of a country.

Bearing this in mind, EFCAMDAT2 cannot fully ascertain the variety in L1 backgrounds or multilingualism. There remains uncertainty on whether some of the sampled learners are indeed speakers of the national language even though they are residents or nationals of that country. Despite these uncontrolled variables, the nature of big data is to contain large overall numbers from a specific country so that the representative sample is sufficient to allow strong national language affects to emerge in data (Alexopoulou et al., 2015). Moreover, the Chen et al. (2021) study using the EFCAMDAT2 did not mention any of the previously mentioned variables when they reported their findings.

4.2.1.2 CEDEL2

The *Corpus escrito del español L2* (CEDEL2) is a systemic collection of authentic and contextualized written and spoken language with 744,950 words produced by L2 learners which are subdivided into 11 learner corpora depending on the L1. CEDEL2 incorporates these aspects in data collected between 2006 and 2016 through its corpus design via what is termed as metadata and adheres to specific design principles focusing on L2s other than English. Design principles take into account learners from all proficiency levels, L1 ages and backgrounds. The goal of the CEDEL2 is to offer a freely available resource to the research community that promotes cross-linguistic comparison, incorporates more learner and task variables, includes varied hypothesis-testing research and considers different perspectives regarding corpora control (Lozano, 2022, pp. 965-966).

The main objective behind the CEDEL2 is that it was created according to corpus-design principles that were adapted for SLA purposes (Sinclair, 2005). The CEDEL2 features nine design features making it a suitable and valuable tool for SLA research. According to Lozano (2022), they include:

1. Same design across sub corpora to ensure maximum comparability
2. SLA motivated variable (learner profile)
3. SLA-relevant variables (task profile)
4. Multiple L1 backgrounds
5. Cross-sectional, developmental corpus
6. Bidirectionality
7. Bimodal contrasts
8. Dual native control subcorpora
9. Heterogeneous sample

Online written data collection for the CEDEL2 spanned over 10 years collecting 4,399 written and spoken files from 4,399 speakers with 1,931 files from L1 English learners producing material in L2 Spanish. With SLA research as a priority, the CEDEL2 has twenty learner variables factoring in linguistic background and five task variables as shown in Figure 2 (CEDEL2, 2023; Lonzano, 2022).

Linguistic background variables	Task variables
<ol style="list-style-type: none"> 1. L1 of the learner 2. L1 of the learner's father 3. L1 of the learner's mother 4. Language(s) spoken at home 5. Placement test score (1-43 points) 6. Proficiency level (lower beginner up to upper advanced) 7. Proficiency level self-evaluation on each skill in Spanish (speaking, listening, writing, reading). 8. Proficiency level self-evaluation on each skill in additional foreign language (speaking, listening, writing, reading). 9. Spanish language certificates held, if any 10. Sex 11. Age 12. Age of exposure to L2 Spanish (AoE) 13. Years studying Spanish (Length of Instruction, LoI) 14. Stays in Spanish-speaking countries? (yes/no): 15. Stay(s): Where? 16. Stay(s): When? (period(s) of residence) 17. Stay(s): How long? (length of residence) 18. School/University/Educational institution (if any) 19. Major degree (if any) 20. Year at university/school (if any) 	<ol style="list-style-type: none"> 1. Task title 2. Task text (written text/spoken text transcription/audio file) 3. Approximate time to produce the task (in minutes). 4. Where was the task done? (in class/outside class/both) 5. Resources used to produce the task (help from Spanish native/bilingual dictionary/monolingual dictionary/spellchecker/grammar book/background readings/none)

Figure 2 Linguistic background and task variables (CEDEL2, 2023).

The CEDEL interface offers multiple and sophisticated search and download options which allow results which can be refined through 12 filters (Figure 3). The features with the

highest importance for this study include concordance lines which show the searched for element being displayed in the center of accompanied text. The searched for element is also known as key word in context (KWIC). Features also include a filter for L1 medium, task medium and proficiency level (Lonzano, 2022).

The screenshot displays the CEDEL2: Corpus Escrito del Español L2 (version 2) web interface. At the top, a red navigation bar contains links for CEDEL2, Search / Download, User guide, Statistics, About, and Contact. The main content area is divided into several sections:

- Corpus to search:** A dropdown menu for 'Subcorpus' is set to 'Learners of L2 Spanish'.
- Sensitivity:** Two checkboxes are present: 'Case' and 'Orthographic accents', both of which are unchecked.
- Result (Output):** This section includes:
 - Result type:** A dropdown menu set to 'Concordances (KWIC)'.
 - Result subtype:** A dropdown menu set to 'Words'.
 - Sorting:** Two buttons, 'Concordance match' and 'Previous element', with a dropdown menu.
 - Page size:** A dropdown menu set to '50'.
- Filters:** A large section with multiple filters:
 - L1:** A dropdown menu set to 'L1 any - L2 Spanish'.
 - Medium:** A dropdown menu set to 'Any'.
 - Sex:** A dropdown menu set to 'Any'.
 - Proficiency level:** A dropdown menu set to 'Any'.
 - Placement test score (%):** A horizontal slider.
 - Proficiency (self-assessment):** A horizontal slider.
 - Task title:** A dropdown menu set to 'Any'.
 - Filename:** A dropdown menu set to 'Any'.
 - Age:** A horizontal slider.
 - Age of exposure to Spanish:** A horizontal slider.
 - Years studying Spanish:** A horizontal slider.
 - Stay abroad (months):** A horizontal slider.
- Words:** A text input field with the placeholder 'Five words maximum'.

At the bottom, there is a red button with a left arrow and a red button labeled 'Search Clean'.

Figure 3 Example of CEDEL2 web-based interface (Lonzano, 2022).

The data encompassed in the CEDEL2 is from learners of Spanish at six proficiency levels ranging from low beginner to higher intermediate. Firstly, for the objective measurement, learners are given a 43-point standardized placement test which gauges their levels according to six levels (Figure 4) (CEDEL2, 2023).

Proficiency level	Placement test score	Corresponding % score
Lower beginner	0-12	0%-28%
Upper beginner	13-20	30%-47%
Lower intermediate	21-28	49%-65%
Upper intermediate	29-35	67%-81%
Lower advanced	36-40	84%-93%
Upper advanced	41-43	95%-100%

Figure 4 Objective Measurements Proficiency Levels (CEDEL2, 2023).

A distinction from other learner corpora is that the CEDEL2 does not contain standardized measures for learners' proficiency, rather only two measures being objective and subjective. While objective measurements are proficiency levels, the CEDEL2 (2023) describes subjective measures as:

Subjective measurement: Learners self-rate their proficiency in Spanish for each of the four skills (speaking, listening, reading, writing) according to a six-point ordinal scale. The subjective measurement for each skill is then transformed into a 1-6 numeric scale and a new variable is created called 'Proficiency self-assessment', which is an average of the four observations. For example, suppose a learner self-rates their Spanish as follows: speaking A1, listening B1, reading A2, writing A1. These ordinal values are transformed into their corresponding numeric values: 1, 3, 2, 1. The final average for the variable 'proficiency self-assessment' is 1.75 (out of a maximum of 6).

The tasks outlined in the CEDEL2 include 14 tasks (Table 72) which are not assigned to any proficiency level. Learners have the option to choose any task independently of the level they have (CEDEL2, 2023).

Task #	Task Title	Task #	Task Title
1	Region where you live	8	Terrorism
2	Famous person	9	Anti-smoking law
3	Film	10	Gay couples
4	Last year holidays	11	Marijuana legalization
5	Future plans	12	Immigration
5	Recent trip	13	Frog
7	Experience	14	Chaplin

Table 72 Example of CEDEL2 Writing Tasks (CEDEL2, 2023).

4.2.1.3 CAES

The project involving the *Corpus de aprendices de español* (CAES) started in October 2011 and concluded in December 2020 with the newest version 2.1 being released to the public in March 2022. The CAES combines texts written by L2 Spanish learners according to five CEFR levels from A1-C1 which are applied to Spanish through *Plan curricular del Instituto del Cervantes*. The newest version of the CAES contains samples from 2,544 students from eleven different languages, each writing two or three texts according to their level producing 6,561 tasks integrated in 2,544 tests (CAES, 2023).

Samples were collected according to uniform criteria and a common protocol using a computer application designed for this purpose. The texts produced by the students received automatic morphosyntactic annotation and were organized according to a system of categories. The L2 Spanish written text data was dumped into a user-friendly computer application that allowed a large number of simple and combined queries with linguistic, personal and social variables (CAES, 2023).

The objective for the CAES is to carry out applied research based on solid and objective data so that it can provide information on learning difficulties, most common errors, vocabulary, etc. The CAES was promoted and financed by the Instituto Cervantes whose goal was to fill the gap in the specific area of language learner corpora for researchers looking to investigate L2 Spanish writing (CAES, 2023).

The search criteria of the CAES web application can be configured to consider all the parameters used in the corpus design thus allowing a researcher to select data according to

five indices including acquired level of knowledge of Spanish (from A1 to C1), L1 (initial or family language), country of residence, age and gender (CAES, 2023).

Since the samples were labeled and sorted manually, each of the graphic forms in the texts have a label associated with it that indicates each of the grammatical elements that it is composed of, the word class that corresponds to it, the values of the categories that apply to it and the lemma to which it belongs. Considering the problems that exist in texts of this nature, the application allows the results to be sensitive to the differences between upper and lower case and to the presence or absence of graphic accentuation. It is also possible to search for elements, lemmas or subcategories that occupy contiguous positions in the texts and even elements, lemmas or subcategories that appear in a nearby context, even if they do not occupy immediate positions (CAES, 2023).

Results show a simple statistic, a complete statistic or show the sequences containing the fragment. An additional display will show the relevant features of the person who wrote the text, the sequence of graphic forms used, as well as the corresponding labels and lemmas. Important for the current study, results show an utterance in which the grammatical element appears in concordance lines (CAES, 2023). *Llegaré* was sample grammatical element used in the KWIC as seen in Figure 5.

1	A1/Inglés	Lo siento , pero	llegare tarde y , por_favor , no me esperais .
2	B1/Árabe	Recuerde mi tu email y tu numero de telefono porque voy a llamar te cuando	llegare a el aeropuerto .
3	B2/Chino mandarín	- Si	llegare a universidad para aprender , podria solicitud dormitorio universitario ?
4	A1/Portugués	Lo siento , pero no	llegare en casa para cenar con tú .
5	A1/Portugués	Lo siento pero hoy	llegare mas tarde .
6	A1/Árabe	hoy creo que	llegare mas tarde por_motivo_de probleme de transporte .
7	A1/Portugués	Hoy	llegare mas tarde porque tengo que hacer trabajos extras en la editora .
8	A1/Portugués	Te escribo porque	llegare más tarde hoy , no mi espere para el jantar , sólo retorno a nuestra casa después de las 🇪🇸
9	A1/Árabe	pues es un parte de mi , vais a mi conocer mas quando	llegare proximamente .
10	A1/Portugués	Te escribo para decir te que no me	llegare tan temprano quanto esperava , por que voy salir de tapas con mis compañeros de trabajo .
11	A1/Portugués	Estoy le escribendo para decir que	llegare tarde , pues que yo tengo que hacer mi estudios en la Universidad .
12	A1/Portugués	Me voy a el cine con una chica muy guapa y	llegare tarde en casa .
13	A1/Árabe	hola amigo soy redouan puedo decir te que+llegare	llegare tarde esta noche por eso no debes esperar me , por que tengo mas horas de el estudio en cervantes 🇪🇸
14	A1/Portugués	Mama ,	llegare tarde hoy , no deben esperar me .
15	A1/Portugués	Yo	llegare tarde hoy .
16	A1/Chino mandarín	Hoy	llegare tarde porque el embotellamientom .
17	A1/Portugués	Tengo examen de español mañana todavia	llegare tarde porque voy estudiar en la biblioteca de UnB con mis compañeras de clase .
18	A1/Chino mandarín		Llegare a casa antes_de las ocho .
19	A2/Inglés		Llegare en Madrid de 1/12/11 a 5/12/11 a las diez_de_la_mañana .
20	A1/Chino mandarín		Llegare tarde , porque mi jefe me da mas trabajo que hacer .
21	A1/Portugués	No me esperen , pues yo no tengo idea de la hora qué	llegaré .
22	A2/Chino mandarín	Cuando	llegaré a Amsterdam , la primera cosa que haré es visitar La_Casa_de_Anne .

Figure 5 Example CAES concordance lines with KWIC (CAES, 2023).

The variety of tasks included in the CAES are made up of 12 different written text categories (Table 73) (CAES, 2023).

Task #	Task Title	Task #	Task Title
1	Job change	8	Vacation post card
2	Letter to a friend	9	Complaint to airline
3	Family	10	Movie review
4	Smoking in public places	11	Room reservation
5	Funny story	12	Application for admission
5	Late arrival note		
7	Person you admire		

Table 73 Example of CAES Writing Tasks (CAES, 2023).

4.2.2 Data Summary

The data set used for this study involved the EFCAMDAT2 corpus for Spanish learners of L2 English. As the EFCAMDAT2 is a large corpus, it was necessary to utilize and merge data from the CEDEL2 and the CAES to have a large enough representative sample from English learners of Spanish L2. Table 74 shows the corpus summary. The total combined word count for the three corpora is 1,478,492 words from 3,645 learners who carried out a total of 10,871 tasks or writing scripts. Following the global corpus summary, each corpus will be detailed individually.

Corpus	Words	Learners	Scripts
EFCAMDAT2	771,162	1,437	8,187
CEDEL2 & CAES	708,330	2,208	2,684
Total	1,478,492	3,645	10,871

Table 74 Corpus Data Summary.

A breakdown by CEFR proficiency level considering corpora, words, scripts and learners is provided in Table 75. At the A1 level the EFCAMDAT2 has a much higher word count, 125,656 versus 23,020, than both the CEDEL2 and the CAES. At the A2 level, although still high, the margin decreases between the English and Spanish corpora. EFCAMDAT2 contains 163,098 words while the CEDL2/CAES contain 113,938.

Level	Words		Scripts		Learners	
	EFCAMDAT2	CEDEL2/ CAES	EFCAMDAT2	CEDEL2/ CAES	EFCAMDAT2	CEDEL2/ CAES
A1	125,656	23,020	2,571	200	459	119
A2	163,098	113,938	2,065	794	339	565
B1	227,225	110,041	2,004	509	337	418
B2	184,831	166,613	1,175	494	222	440
C1	64,358	185,434	340	455	69	434
C2	5,994	109,284	32	232	11	232

Table 75 Data Summary by CEFR Proficiency Level.

The EFCAMDAT2 contained more words at the B1 level with a comparison of 227,225 versus the CEDEL2/CAES with 110,041. The same trend continued at B2 with the EFCAMDAT2 having a relatively higher word count at 184,831 versus 166,613 in the CEDEL2/CAES. The tendency for a higher word count in the EFCAMDAT2 changes at the C1 and C2 levels. The CEDEL2/CAES has more than twice the number of words at C1 with 185,434 versus 64,358 in the EFCAMDAT2. At the C2 level there is quite a big difference in word count between the English and Spanish corpora. The CEDEL2/CAES contains 109,284 words while the EFCAMDAT2 contains 5,994.

The corpus data taken from the EFCAMDAT2 corpus was the biggest sample in the entire study in comparison to the CEDEL2 and the CAES. Table 76 shows proficiency levels by words, learners and scripts. The two levels with the highest word count are B1 with 227,225 and B2 with 184,831 words. However, both B1 and B2 levels have fewer learners and fewer scripts than A1-A2. The C1 and C2 levels have relatively fewer words, learners and scripts than both A1-A2 beginner levels and B1-B2 intermediate levels. The beginner levels have the highest amounts of learners and scripts, yet they have fewer words than B1-B2 and more words than C1-C2. The EFCAMDAT2 does not contain either gender or age variables.

Level	Words	Learners	Scripts	Gender	Age	Dates
A1	125,656	459	2,571	N/A	N/A	2011-2014
A2	163,098	339	2,065			
B1	227,225	337	2,004			
B2	184,831	222	1,175			
C1	64,358	69	340			
C2	5,994	11	32			
Total	8,187	771,162	1,437			

Table 76 EFCAMDAT2 Data Summary.

The CEDEL2 is the largest of the two Spanish corpora with a total word count of 563,643 with 1,907 scripts and learners (Table 77). The total word count increases from the proficiency level A1 with 9,917 to C1 with 166,882 and decreases from C1 to C2 with 109,284. The number of learners and scripts are the same throughout each proficiency level. The lowest number of scripts and learners appears at the A1 level with 78. However, the highest number of learners and scripts appears at the A2 level with 438 while the second highest is the C1 level with 408. Only a small variation occurs at the B1 to B2 level with the difference only being 19 learners and scripts. Although the C2 level has the second lowest number of scripts and learners with 232, it is much higher than the A1 level. In this corpus both female and male variables were included in each proficiency level with a total age range between 13 and 88.

Level	Words	Learners	Scripts	Gender	Age	Dates
A1	9,817	78	78	28 Male/50 Female	15-52	2006-2016
A2	60,296	438	438	163 Male/275 Female	13-66	
B1	84,258	366	366	119/247 Female	14-75	
B2	133,106	385	385	99 Male/286 Female	15-75	
C1	166,882	408	408	93 Men/315 Women	15-81	
C2	109,284	232	232	74/158 Women	17-88	
Total	563,643	1,907	1,907			

Table 77 CEDEL2 Data Summary.

The CAES is the smaller of the two Spanish corpora with 144,687 words from 301 learners producing 777 scripts (Table 78). The CAES only has five proficiency levels which range from A1-C1. The A2 level contains the highest word count with 53,642 with the B2 coming in second with a large difference of 33,507 followed by the B1 with 25,783, the C1 18,552 and A1 with 13,203. In the same order as the word count, the A2 has the most learners

with 127 participants producing 356 scripts. The lowest amount of data was found in C1 with 26 learners and 47 scripts followed by A1 with only 41 learners and 122 scripts. The intermediate B2, behind A2, had the second highest amount of data with 55 learners producing 109 scripts. The B1 proficiency level had only three fewer learners than the B2 with 52 while it contained more scripts with 143. The CAES includes gender data as well as age data with a range between 17 and 72 years. Data for this corpus was collected from 2011-2020.

Level	Words	Learners	Scripts	Gender	Age	Dates
A1	13,203	41	122	20 Male/21Female	15-72	2011-2020
A2	53,642	127	356	46 Male/81 Female	18-67	
B1	25,783	52	143	17 Male/35 Female	15-66	
B2	33,507	55	109	22 Male/33 Female	15-66	
C1	18,552	26	47	15 Male/32 Female	15-71	
C2	0	0	0	N/A	N/A	
Total	144,687	301	777			

Table 78 CAES Data Summary.

4.2.3 Technical Details and Evaluation

The main goal of this study revolves around extracting information from corpus databases to measure the frequency of any given measure in the MBF. Therefore, the overall procedure started with making sure the variables were in the correct order. The subcorpora was set to either L2 learners of English or Spanish with CEFR language level set appropriately, the L1 was indicated as either Spanish or English depending on the corpora. When applicable, the medium was set to written text. The result type was set to concordances with KWIC.

The next step involved inputting an MBU into the KWIC in the CEDEL2 and the CAES to view the total number of concordance lines. The KWIC in the CEDEL2 is labeled *Words*, while in the CAES it is *Elem. Gramaticales* or grammatical elements. In the case of the EFCAMDAT2, the search function did not give the desired results so the entire data set for each proficiency level was downloaded as a CVS file and then uploaded to Sketch

Engine³⁰. The same procedure as previously described was then carried out according to proficiency levels, selecting a written text and inputting a MBU in the search function to retrieve the KWIC in the concordance lines. Figure 6 gives the entire scope for the type of search, type of results, order, CEFR level, L1 and KWIC for *porque* in the CAES.

The screenshot shows the CAES search interface. At the top, there is a header with the CAES logo and the text 'Corpus de aprendices de español (CAES)'. Below the header, there is a navigation bar with links to 'CAES', 'Datos', 'Búsqueda', 'Listados', 'Ayuda', and 'Contacto'. The main search area is divided into several sections: 'Búsqueda' (Search) with a 'Tipo' (Type) dropdown set to 'Elem. gramaticales'; 'Resultado' (Result) with a 'Tipo de resultado' (Result type) dropdown set to 'Concordancias', an 'Ordenación' (Ordering) dropdown set to 'Coincidencia', and a 'Tamaño de página' (Page size) dropdown set to '50'; 'Sensibilidad' (Sensitivity) with checkboxes for 'Acentos' and 'Mayúsculas'; and 'Filtros' (Filters) with various dropdowns for 'Nivel de español' (B2), 'L1' (Inglés), 'País' (Cualquiera), 'Sexo' (Cualquiera), 'Edad desde' (Cualquiera), 'Edad hasta' (Cualquiera), 'Tipología textual' (Cualquiera), 'Tema' (Cualquiera), and 'Estudiante'. At the bottom, there is a search bar with the text 'porque' and a 'Buscar' button. There are also buttons for 'Descargar', 'Volver', 'Limpiar', and 'Buscar'.

Figure 6 Example search criteria for *porque* in CAES search function.

Once the concordance lines including the KWIC were made available, they were saved in a PDF file under the correct criteria in the MBF. Given the relatively large amount of data as well as the large number of fine-grained measures, precision was of high importance in determining the accuracy and the categorization of the MBU, not to mention its subordinate or coordinate validity. Similar to Lahuerta Martínez (2018a), this study analyzed KWIC manually for the calculation of quantitative measures. The annotation process took eight months in total. Manual analysis was deemed necessary in absence of any feasible NLP tools with the depth to incorporate the parameters of the established MBF. Perhaps future research will create an NLP in which the MBF can be included. Figure 7 shows how results for concordance lines and KWIC appear for the search for *porque* in the CAES.

³⁰ Sketch Engine (<https://www.sketchengine.eu/>) is a website which allows you to build, develop and manage corpora by uploading data sets or using previously existing and available data sets. The search functions allow a myriad of possibilities when using corpus databases for research purposes.

Resultados 1 a 50 de 104

← 1 2 3 → Ir a la página

1	B2/Inglés	oy muy trabajadora y tengo muchas preguntas sobre cómo funciona el mundo y porque .
2	B2/Inglés	Creo que hay muchas razones muy buenas para no fumar , y no lo digo solo porque es mi opinion , y no lo digo solo porque nadie en mi familia fuma .
3	B2/Inglés	El acto de fumar en lugares públicos es malo porque el humo de este puede dañar a los personas que tiene contacto con el humo .
4	B2/Inglés	No quiero controlar la gente , esto es porque quiero implementar sitios para fumar , pero a el mismo tiempo creo que es nece
5	B2/Inglés	y vería que alguien estuviera fumando cerca de mis hijos me pondría molesta porque mis hijos estuviesen oliendo ese olor toxico .
6	B2/Inglés	Sus saludes no son dentrimenales porque sus dietas son más buenas que los estados_unidos y por eso sus saludes no sc
7	B2/Inglés	al en las décadas pasadas , fumar en lugares públicos no debe ser aceptable porque no toma en consideración las preferencias y los estados de salud de las person
8	B2/Inglés	Me gustaría estudiar en su programa de postgrado porque quiero aprender más de la cultura de el pais .
9	B2/Inglés	ndo vuelves de una noche en un bar o un disco y tienes que cambiar las ropas porque están sucio por el fumo ?
10	B2/Inglés	el conocimiento de ellos están muy importantes y más gente debe estudiar los porque luego las personas de el mundo saberían más sobre sus vecinos .
11	B2/Inglés	Ahí , las reglas son más diferentes porque sus saludes no están más efectados que en los Estados_Unidos .
12	B2/Inglés	Quiero aplicar a este programa de postgrado de lengua y literatura española porque me encanta las novelas y como la lengua cambia cuando estás escribiendo un t
13	B2/Inglés	ellos peden hacer cualquiera con su salud , yo no tengo una problem con esto porque los fumadores tienen la informacion que nesicita si no quiere fuma nada más .
14	B2/Inglés	Otros pueden decir que los restaurantes y bares cerraran porque la genta fumadora ya no iría a esos lugares ... pero no creo que este argumento
15	B2/Inglés	Estoy buscando un programa en latinoamerica porque tengo mucho curiosidad de este región .
16	B2/Inglés	an no solo las personas que padecen estas enfermedades sino todos nosotros porque , como ciudadanos , los fondos recaudados por el gobierno mediante los impue
17	B2/Inglés	Por eso , me gusta la historia de un país porque siempre hay conexiones entre el pasado y el futuro .
18	B2/Inglés	Mucha gente todavía fuma aunque saben bien el riesgo , porque todo el mundo sabe ya que fumar es algo que causa bastante daño a el cuerpo
19	B2/Inglés	Hay muchas razones porque pienso así y porque estoy en favor .
20	B2/Inglés	eligroso para los no fumares que los fumares , inhalar el humo de un cigarrillo , porque el humo lleva muchas toxinas venenosas .

Figure 7 Example concordance lines and KWIC for *porque* in the CAES.

The question of accuracy in the manual analysis is of valid concern for replication of this study. Moreover, it must be considered the MBU which are duplicated in different categories of hypotaxis and parataxis. Within this conundrum lies the importance of the MBF and its ability to separate the differences between fine-grained measures. Therefore, the analysis for each MBU as a KWIC in any given concordance line adhered to the theoretical meaning found in parataxis and hypotaxis as well as it was cross-referenced with grammar resources such as *An Introduction to Functional Grammar* (2014), *A Comprehensive Grammar of the English Language* (1985), *Nueva gramática de la lengua español* (2010), *Systemic Functional Grammar of Spanish: A Contrastive Study with English* (2010), etc.

While Section 3.2 first details the methodology for operationalizing the MBF and Section 3.4 presents the MBF as an analytical tool used in the present study, Figure 8 is included give a concrete example of how the subordinator *because* is categorized in SFG under hypotaxis through enhancement as α x β in terms of binding elements of unequal status. Under enhancement and expansion, this clause type falls under causal-conditional through cause and reason. The basic definition is “because α so the result is x β ” (Halliday, 2014, pp. 452-481). In the below example, clause 1 which equals α is bound through the adverbial

subordinator *because* to enhance the unequal clause 2 which equals $x\beta$. Looking at the KWIC allows the researcher to look at the clause in its entirety, thus seeing where the subordinator appeared in the sentence and if there were any embedded features in a clause.

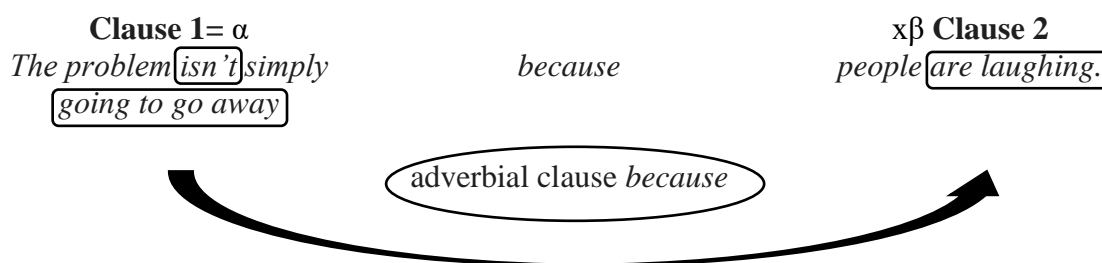


Figure 8 Example Analysis of *because* KWIC and concordance line (Halliday, 2014).

When a subordinator could be classified under more than one hypotactic category, the basic definition was considered. Case and point are with *since* which can be interpreted under hypotaxis and enhancement under temporal different time later and causal-conditional in cause and reason. Example (184) shows how the meaning of *since* is equal to α then subsequently $x\beta$ in its temporal form. When doubt appeared as to which group an MBU might be classified under, the rule of thumb was to find a synonym for the subordinator in question and analyze if it kept the same meaning. Putting this into practice, in its temporal form *since* can be substituted with *after* without losing its inherent meaning. (185) demonstrates a comparative form with *as* in which α is like $x\beta$ (Halliday, 2014). *As* can also be used as a temporal indicator, however (185) shows the manner through comparison form which cannot be substituted by the temporal *while*.

(184) ***Ever since*** 'Wildlife in America' appeared in 1959, || and ***especially since*** 'The Snow Leopard' won a 1978 National Book Award, || ...

(185) ||| *Gradually, they outgrow their baby shoes* || – *if the expression is pardoned, || as Snufkin is in fact the only one of them [who uses footwear at all].* |||

The next step is to explain how the analysis for parataxis was carried out. Parataxis is defined as the linking of elements of equal status. In the sense they could stand as a functioning whole. Looking at extension in parataxis, the overall scheme will appear as sentence 1 + sentence 2. When we look at addition through a positive additive, we can see the basic meaning is 1+2 (Halliday, 2014, pp. 450). Figure 9 shows how sentence 1 and sentence

2 are entire units and can stand freely on their own. In the analysis of KWIC for parataxis, it is important to note that co-referential subjects were excluded from the parameters of the study. An example of this would be *I ate ice cream and cake*. In the previous sentence, the co-referential subject *cake* cannot be independent of *I eat ice cream and*. Furthermore, cases of ellipsis were also excluded from this study as the writers' intentions are unknown.

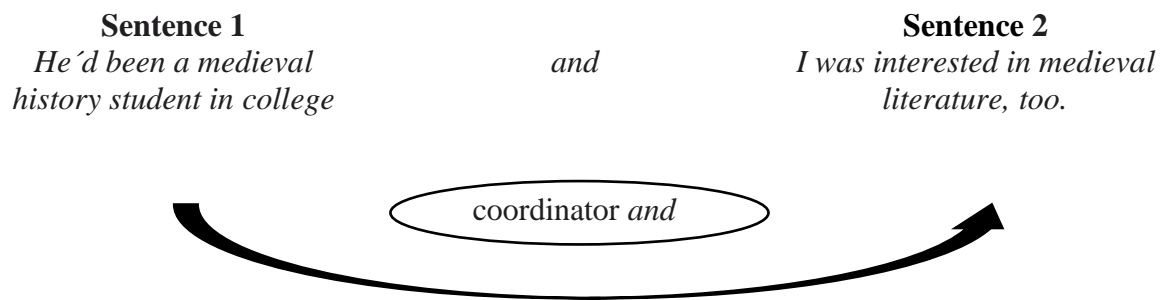


Figure 9 Example analysis of parataxis and KWIC and concordance line (Halliday, 2014).

An important element in the analysis of KWIC in concordance lines is hypotactic elaboration which takes the form of non-defining relative clauses. A hypotactic elaborating finite non-defining relative clause happens when the secondary clause includes a *wh*-form which adds additional information. The *wh*-element is embedded as the qualifier in a finite nominal group adding information to a head noun. The key reference for a non-defining relative clause is the head noun which the *wh*-element modifies and takes place between commas (Halliday, 2014). A non-defining relative clause is set apart by the designation of a proper name and describes the head noun in an essential way (Quirk et al., 1985).

Figure 10 provides an example of how non-defining relative clauses were identified in the KWIC. The non-defining relative clause begins with the *wh*-element *who* and continues as a dependent clause with a finite verb. In this case, *who* adds information to the head noun *Mary Smith*. The head noun is the subject of the sentence and continues with *wants to meet you*. There are several cases, such as *when*, in which MBUs for hypotaxis and enhancement are duplicated in elaboration. When this occurred, the main factor in determining the difference between enhancement and elaboration centered on the presence of a head noun which determined the category of hypotaxis.

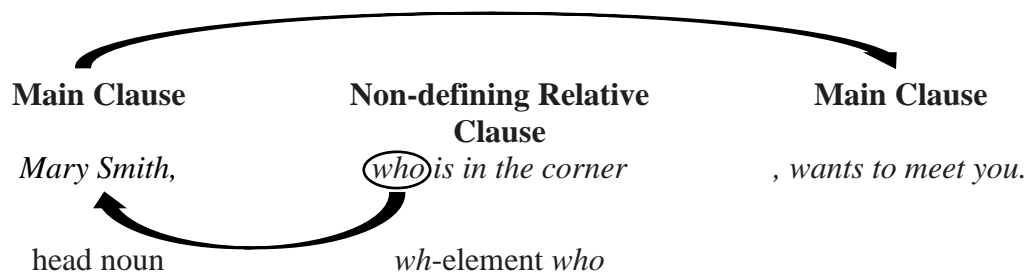


Figure 10 Example analysis of hypotactic elaboration and KWIC and concordance line (Halliday, 2014).

An important distinction needs to be highlighted with Spanish *wh*-element *que*. The ubiquitous yet ambiguous translation of *que* into English was seen to take different form in non-defining relative clauses. Therefore, the head noun was consulted to identify the correct translation of *que*. The ambiguous translations of *que* involving *which/that* lead to further analysis of translations in which the logical MBU was deduced from the context. According to Halliday (2014), *which* and sometimes *that* can appear in non-defining relative clauses (p. 466). It must be acknowledged that the speaker and their semantic intentions could have an influence on which form is used in L2 Spanish.

The next step towards achieving quantifiable results is to normalize the word count. Normalizing transforms an MBUs overall count to its normed frequency. Since text lengths vary, longer texts inflate the importance of a certain items. To offset text-length bias, one scales the frequency of a phenomenon per text, such as per 1,000 words. Second, normalizing eliminates the feature-concentration bias: some phenomena are naturally scarce in a document while others are naturally common. Normalizing converts a normed frequency to a value vis-à-vis normed frequency in each document. Consequently, one can measure the relative presence of two or more linguistic features within any given text (Asención-Delaney & Collentine, 2011, p. 305).

Once the KWIC was noted with the correct MBU, the total numbers were added and imputed into the correct proficiency level being investigated in the MBF (Figure 11). This study normalized the frequency of MBUs per 5,000 words. The logical premise behind this decision was to adequately measure the proficiency level with the lowest word count. In this case it was the C2 level in the EFCAMDAT2 which only contained 5,994 words. Therefore, it was decided that normalizing per 5,000 words would be a better measure than 1,000 because of the large size of the data set. Furthermore, normalizing per 10,000 words was

ruled out since the higher nominalization rate would diminish the scale of results for proficiency levels with lower word counts.

	Corpus	MBC	MBSC	MBU	MR R2	A1	A2	B1	B2	C1	C2
1	EFCAMDAT2	temporal	same time	<i>while</i>	Three o'clock	-	-	-	-	-	-
	CEDEL2 & CAES			<i>mientras</i>		-	-	-	-	-	-

Figure 11 Example MBF with CEFR proficiency levels.

The scope of this study and the total amount of fine-grained measures lead to the creation of four different levels of analysis starting from a broad perspective of the overall phenomenon of taxis and ending with individual MBUs. The rationale behind this was to be able to look at each fine-grained measure from a variety of angles and perspectives to see how overall categories and subcategories interacted at different proficiency levels.

Figure 12 demonstrates how the results, discussion and conclusions will be dissected and analyzed over four different levels. Level 1 includes the general categories of taxis. Level 2 adds expansion through extension, enhancement and elaboration. Level 3 includes MBCs such as indicators like temporal, spatial, manner and causal-conditional as an addition to the previous two categories. Level 4 looks at the use of each meaning-based unit, such as *because/porque*, within the framework of the previous three levels to give a fine-grained analysis of frequency of use.

- Level 1: Taxis
- Level 2: Expansion
- Level 3: Meaning-based Categories
- Level 4: Meaning-based Units

Figure 12 Levels of analysis in the MBF.

4.3 Experiment 1: Level 1 Taxis

Level 1 taxis is the top-tier level of analysis covering the two broadest categories involving hypotaxis and parataxis which encompasses expansion, logico-semantic relations and MBUs. The goal of this level is to give a global view of trends and anomalies occurring in a large data set.

4.3.1 Results

Table 79 summarizes the distribution of normalized MBUs by 5,000 words per 5k across hypotaxis and parataxis with the inclusion of CEFR proficiency levels A1 to C2 along with the total amount of measures for each Level 1 taxis.

Corpus	Level 1	Measures	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis	57	21.2	62.0	76.1	77.9	113.0	54.2
CEDEL2/CAES			37.6	63.9	90.9	86.2	103.1	100.5
EFCAMDAT2	Parataxis	33	82.4	109.8	138.5	132.3	124.5	126.8
CEDEL2/CAES			147.3	133.1	155.6	154.0	157.9	159.1

Table 79 Level 1 Results for Hypotaxis and Parataxis.

4.3.1.1 Hypotaxis

Figure 13 shows the normalized MBUs for hypotaxis comparing data using 57 fine-grained meaning-based measures from the EFCAMDAT2 and CEDEL2/CAES. At the A1 proficiency level, the CEDEL2/CAES (37.6) had a frequency of 16.2 more words than EFCAMDAT2 (21.2). At the A2 level the EFCAMDAT2 (62.0) had a 1.9-word lower use frequency than the CEDEL2/CAES (63.9).

While the mean word frequency at the A2 level was relatively close, the B1 and B2 levels exhibited the opposite characteristic with learners in the CEDEL2/CAES using hypotaxis at a higher frequency. A divergence of 14.8 words per 5k at B1 was seen with the EFCAMDAT2 having 76.1 and the CEDEL2/CAES 90.9. The B2 level displayed the same imbalanced trend with 77.9 words occurring in the EFCAMDAT2 and 86.2 in the CEDEL2/CAES.

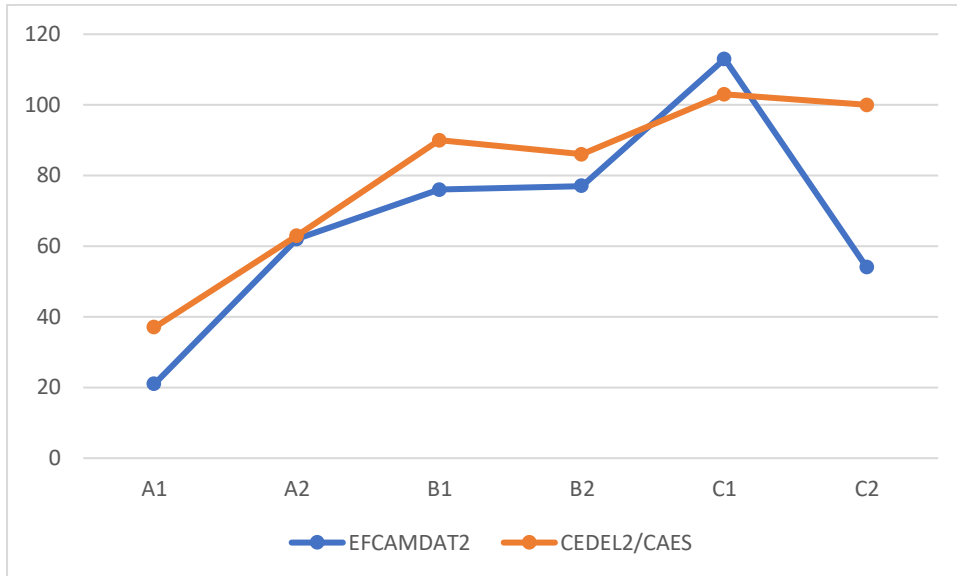


Figure 13 Level 1 Hypotaxis Result Summary.

The trend showing MBU use in the CEDEL2/CAES changed at the C1 level. There was a 9.9 difference with the EFCAMDAT2 (113.0) having the highest frequency versus the CEDEL2/CAES (103.1). At the C2 level the overall trend drastically switched with a difference of 46.3 with the EFCAMDAT2 (54.2) exhibiting a lower frequency than the CEDEL2/CAES (100.5).

4.3.1.2 Parataxis

Parataxis appeared at a much higher frequency than hypotaxis at all levels with 33 measures with the CEDEL2/CAES manifesting a higher MBU frequency than the EFCAMDAT2. Figure 14 shows a rather large dissimilarity occurred at the A1 level with a higher use frequency in the CEDEL2/CAES (147.3) versus the EFCAMDAT2 (82.4). Frequency decreased at A2 in the CEDEL2/CAES (133.1) while parataxis increased in the EFCAMDAT2 (109.8).

There was an increase from A2 to B1 in both the EFCAMDAT2 (138.5) and CEDEL2/CAES (155.6). A different trend appeared at the B2 level with frequency

decreasing in both the EFCAMDAT2 (132.3) and the CEDEL2/CAES (154.0). Frequency decreased at C1 level in data from the EFCAMDAT2 (124.5) while occurrence of parataxis slightly increased in the CEDEL2/CAES (157.9). At the C2 level the EFCAMDAT2 (126.8) displayed a slight increase with the same trend happening in the CEDL2/CAES (159.1).

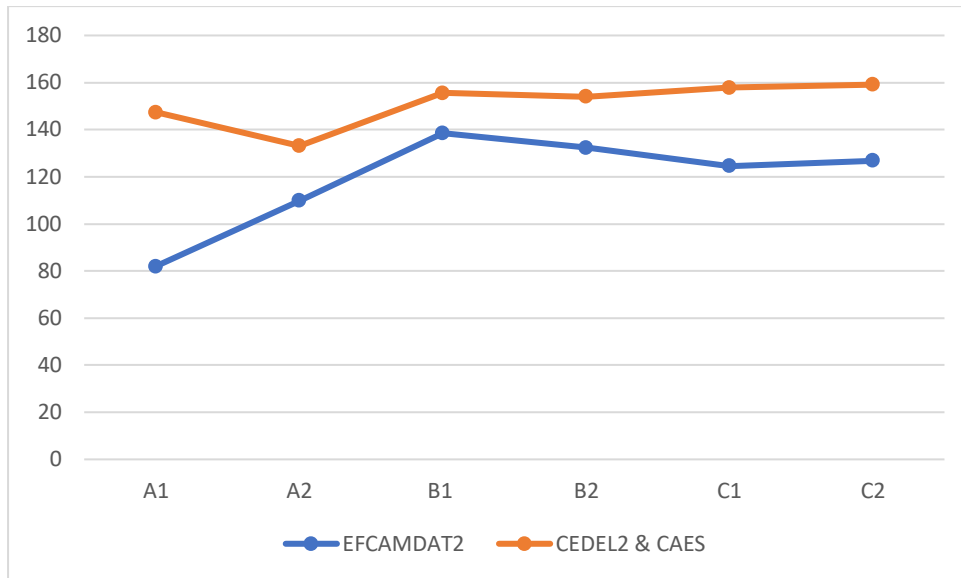


Figure 14 Level 1 Parataxis Result Summary.

4.3.2 Discussion

4.3.2.1 Hypotaxis

Results showed an overall increase of use in hypotaxis MBUs from A1 to the B1 level in both data sets. There was a plateau in frequency from B1 to B2 in the EFCAMDAT2 and the CEDEL2/CAES. Both data sets saw an increase at C1 which did not indicate a regression of hypotaxis. As there was peak frequency in hypotaxis at the C1 level, Level 1 findings are contrary to the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in which subordination indicates intermediate levels of L2 acquisition and subsides at advanced levels.

The general trend in the EFCAMDAT2 and the CEDEL2/CAES appeared to follow the same subordination pattern with a sharp increase from A1 to B1 which aligns with previous findings (Lahuerta Martínez, 2018a; Lorenzo & Rodríguez, 2014). This trend directly contradicts the Gyllstad et al. (2014) study in which subordination appeared homogeneous between A1-A2. The present study found the increased use of subordination

delineates lower levels from advanced levels (Kaushik & Huhta, 2020; Wolfe-Quintero et al., 1998). EFCAMDAT2 results showed hypotaxis leveling off at the B2 level while CEDEL2/CAES showed a minor decrease.

Results confirm the Kim (2004) study in which there is a clear difference in subordination use between A2 and B2. One discrepancy between studies is that Kim found minor differences between A2 and B1 while the present study found significant frequency increases between said levels. Moreover, Kim found a sharp increase from B1 to B2 which is inconsistent with the minor increase in the EFCAMDAT2 and the minor frequency decrease in the CEDEL2/CAES.

Findings from both data sets observed a relative frequency plateau between B1 and B2. The EFCAMDAT2 observed an increase of 1.8 words between B1 and B2 while the CEDEL2/CAES decreased by 4.7 words. Results concur with previous findings which saw the same leveling out of subordination use at intermediate levels (Alexopoulou et al., 2015; Bardovi-Harlig & Bofman, 1989; Chen et al., 2021).

Hypotaxis in both EFCAMDAT2 and CEDEL2/CAES peaked at the lower advanced level thus implying subordination increases through the C1 level as English and Spanish L2 learners use more subordination as their language proficiency increases. Data from the CEDEL2/CAES indicated that hypotaxis may begin to stabilize at the C2 level as there is only a difference of 2.6 words between C1 (103.1) and C2 (100.5).

The Chen et al. (2021) study using the EFCAMDAT2 (Table 12) reported the same phenomena as the current study in which an increase appeared in every subordination indicator from A1 to C1 with a significant frequency drop reported at C2. One possible explanation for this is the word count. The EFCAMDAT2 has 5,994 words to draw from at the C2 level while the CEDEL2/CAES has 109,284 words. This makes for a rather large margin in representative samples thus possibly accounting for the decrease in frequency in the EFCAMDAT2 and the relatively stable frequency in the CEDEL2/CAES at the C2 level.

The second possible explanation is that grammatical metaphor does precisely come into effect at C2 in the EFCAMDAT2. In comparison, the slight decrease in the CEDEL2/CAES doesn't effectively imply that grammatical metaphor or phrasal-level

sophistication replaced subordination. Nonetheless, without measuring grammatical metaphor, insight from this study regarding the matter is hypothetical.

Further research would benefit focusing on advanced levels to further investigate L1 to L2 transfer and subordination acquisition. Results show that L1 Spanish have an overall lower use of hypotaxis at the A1 through B2 than L1 English. While English appears to have a higher subordination density in the first four CEFR proficiency levels, at the C1 the equation changes. L1 Spanish writers in L2 English have a higher frequency of hypotaxis use than L1 English writers.

The current study contradicts Neff et al. (2004) findings in which formal written styles in first-language Spanish favor a much greater use of subordination than English. Evidence from this study's CEDEL2/CAES results suggest there is a higher degree of subordination density in levels A1 through B2. Level 1 findings in comparison to Neff et al. (2004) must be put into perspective since the current study incorporated a much larger representative sample with more students, words and a larger array of communicative tasks. In contrast, Neff et al. (2004) incorporated a single argumentative essay with a total of 60 students.

4.3.2.2 Parataxis

Results for parataxis indicated a different trend than hypotaxis in the EFCAMDAT2 and the CEDEL2/CAES with overall frequency increasing from A1 to B1. Data from the current study is inconsistent with DSSICH citing parataxis takes place at beginning levels and subsides at intermediate levels (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Findings from the current study do not confirm that the frequency of coordination fails to distinguish every CEFR proficiency level (Ai & Lu, 2013; Khushik & Huhta, 2020; Lu, 2011). EFCAMDAT2 data showed parataxis to be an indicator at different proficiency levels between A1 to B1 which is consistent with the Zarco-Tejada et al. (2016) study. The frequency of parataxis in the EFCAMDAT2 from B2 (132.3) to C2 (126.8) is only that of 5.5 words per 5k which might suggest parataxis stabilizes since it fails to clearly differentiate

upper intermediate and advanced proficiency levels. In view of the general lack of attention focused on meaning-based coordination in syntactic complexity, this study would recommend that future research focus on parataxis to understand the role it plays in L2 acquisition as well as L1 transfer to L2.

Data from the CEDEL2/CAES showed a decrease in parataxis MBUs from A1 to A2 which directly contrasts the notion that coordination reaches a peak at beginner levels. From A2 to B1 there is an increase that slightly peaks over the A1 level. There was a rising trend in parataxis MBUs from B1 with the categorical peak at C2. Findings from the CEDEL2/CAES appeared to be consistent with those from Zarco-Tejada et al. (2016) in which there was a frequency increase from A2 to B1. CEDEL2/CAES findings suggest A2-B1 are two levels that distinguish learner proficiency through coordination.

The trend in this discussion might be seen to contrast the idea that parataxis subsides in the intermediate to advanced levels. Regardless of the slight increase from A2 to B1 in the CEDEL2/CAES, parataxis appears to plateau from B1 (155.6) to C2 (159.1) and gives no indication of subsiding. With a plateaued frequency between upper intermediate and upper advanced levels, parataxis fails to distinguish the later three CEFR proficiency levels (Ai & Lu, 2013; Khushik & Huhta, 2020; Lu, 2011).

There is a 90,918 word difference between A1 (23,020) and A2 (113,938) in the CEDEL2/CAES giving a much higher representative sample at the upper beginner level. This might be thought to materialize in an increase in coordination. However, a frequency decrease appears from A1 (147.3) to A2 (133.1). A possible explanation for this phenomenon could be a sharp increase in hypotaxis from A1 (37.6) to A2 (63.9) thus accounting for the decrease in coordination.

In general, the data extracted from the CEDEL2/CAES showed a much higher frequency in parataxis thus suggesting coordination was used at a higher frequency by L1 English learners writing texts in L2 Spanish. The EFCAMDAT2, even though it has a higher word count, includes a word limit on communicative tasks. This design parameter in the corpus database may have influenced results thus causing MBUs extracted from the corpora to occur at a lower overall frequency. Research which resulted in the current Level 1 findings considered a wide range of studies and the general conclusion is hypotaxis received the most

attention. Based on this observation, coordination is an area of research which would benefit from more attention from researchers investigating syntactic and meaning-based complexity.

4.3.3 Conclusions

This chapter seeks to answer the second research question which concerns the DSSICH which proposes that parataxis increases through beginning CEFR levels, subordination indicates intermediate levels of L2 acquisition and subsides at advanced levels as grammatical metaphor increases (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). EFCAMDAT2 and CEDEL2/CAES results for parataxis and hypotaxis showed a significant overall rise in frequency from A1 to B1 which was consistent with previous research (Lahuerta Martínez, 2018a; Lorenzo & Rodríguez, 2014). The body of evidence showed that hypotaxis delineates lower levels from advanced levels (Kaushik & Huhta, 2020; Wolfe-Quintero et al., 1998). It appeared that hypotaxis saw a plateau from the B1 to B2 which concurs with results from previous studies (Alexopoulou et al., 2015; Bardovi-Harlig & Bofman, 1989; Chen et al., 2021).

In terms of research question two, findings cannot confirm the hypothesis that hypotaxis subsides at an advanced level in both data sets. EFCAMDAT2 results showed that subordination increased from A1 to C1 with a significant decrease at C2 which is thought the consequence of having a small representative sample at that level. The sharp decrease in subordination from C1 to C2 was consistent with results from the Chen et al. (2021) study using the EFCAMDAT2 (Table 12). Hypotaxis was seen to increase from A1 to C1 in data from the CEDEL2/CAES with a slight decrease in frequency of 2.6 words at C2 which may indicate a stabilization in use of subordination. Data from the current study indicates that hypotaxis in Spanish and English L2 written texts is used as a linguistic resource that steadily increases from A1 and peaks at C1 which is inconsistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Parataxis in data extracted from the EFCAMDAT2 and CEDEL2/CAES saw an increase in frequency from A1 to B1. After the peak at B1, there was a decrease in parataxis that lasted through C2 in the EFCAMDAT2. However, the frequency of coordination had

little variation between B1 and C2. Findings cannot confirm research question two that coordination decreases after A2 as proposed by the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Research question three explores how the MBF distinguishes syntactic complexity across CEFR proficiency levels. Findings show an increase in coordination use from A2 to B1 which are consistent with the Zarco-Tejada et al. (2016) study. The B1 level appeared to distinguish proficiency levels in data from the EFCAMDAT2 and CEDEL2/CAES. Parataxis failed to distinguish proficiency from B2 to C2 in the EFCAMDAT2. Findings in CEDEL2/CAES data demonstrated that parataxis failed to distinguish proficiency from B1 to C2.

Research question four investigates L1 to L2 transfer. Data extracted from the CEDEL2/CAES broadly indicates writers in L2 English used more hypotaxis in every proficiency level except C1 thus suggesting a higher rate of subordination is transferred from the L1 to L2. Data from the CEDEL2/CAES illustrated parataxis was used at a higher frequency than in the EFCAMDAT2 indicating English L1 transfers more coordination on to L2 Spanish writing. The caveat to these findings must consider the word limit on written tasks incorporated in the EFCAMDAT2 which may have influenced results with both subordination and coordination. Future research into both parataxis, hypotaxis and grammatical metaphor at advanced levels would help identify factors behind trends seen in this study.

4.4 Experiment 2: Level 2 Expansion

Level 2 breaks down expansion through taxis into three categories which include extension, enhancement and elaboration. The goal of Level 2 is to investigate how expansion is manifested in L2 acquisition and L1 to L2 transfer across proficiency levels. Results will be reviewed followed by the discussion and the conclusion section will detail the findings.

4.4.1 Results

The results for Level 2 are summarized in Table 80. As this study moves to the Level 2, results for extension, enhancement and elaboration show the distribution of normalized MBUs by 5,000 words across CEFR proficiency levels A1 to C2. The total quantity of measures for each Level 2 category is also included to provide a wider spectrum of the phenomena at hand.

Corpus	Level 2	Measures	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis	Extension	0.3	0.4	0.3	1.3	2.0	1.7
CEDEL2/CAES			0.2	0.7	1.4	1.8	2.0	2.7
EFCAMDAT2		Enhancement	20.5	60.4	73.6	73.1	106.5	49.2
CEDEL2/CAES			37.1	62.7	85.0	82.3	97.6	93.9
EFCAMDAT2		Elaboration	0.5	1.2	2.2	3.5	4.5	3.3
CEDEL2/CAES			0.2	0.6	4.5	2.1	3.5	3.9
EFCAMDAT2	Parataxis	Extension	75.0	101.4	125.8	114.3	109.5	102.6
CEDEL2/CAES			145.5	129.5	148.5	146.8	148.7	146.7
EFCAMDAT2		Enhancement	5.7	2.3	2.2	1.6	1.2	0.0
CEDEL2/CAES			0.9	1.2	2.1	2.0	2.6	3.6
EFCAMDAT2		Elaboration	1.7	6.1	10.5	16.4	13.8	24.2
CEDEL2/CAES			0.9	2.4	5.0	5.2	6.7	8.8

Table 80 Level 2 Normalized Expansion through Extension, Enhancement & Elaboration.

4.4.1.1 Hypotaxis and Extension

The first category covered in the Level 2 description of results is hypotaxis through extension (Figure 15). MBUs extracted from seven measures in the EFCAMDAT2 showed insignificant increase or decreases in A1 (0.3), A2 (0.4) and B1 (0.3). There was an increase from B1 (0.3) to B2 (1.3) and another increase at C1 (2.0) A slight decrease occurred at C2 (1.7). The data from the CEDEL2/CAES saw low frequency at A1 (0.2) but increased to A2

(0.7), B1 (1.4), B2 (1.8), C1 (2.0) and C2 (2.7). In general, the quantity of MBUs in hypotactic extension occurred at a much lower category than enhancement and lower than elaboration. However, CEDEL2/CAES data showed learners using extension at an overall higher frequency than seen in the EFCAMDAT2 data set.

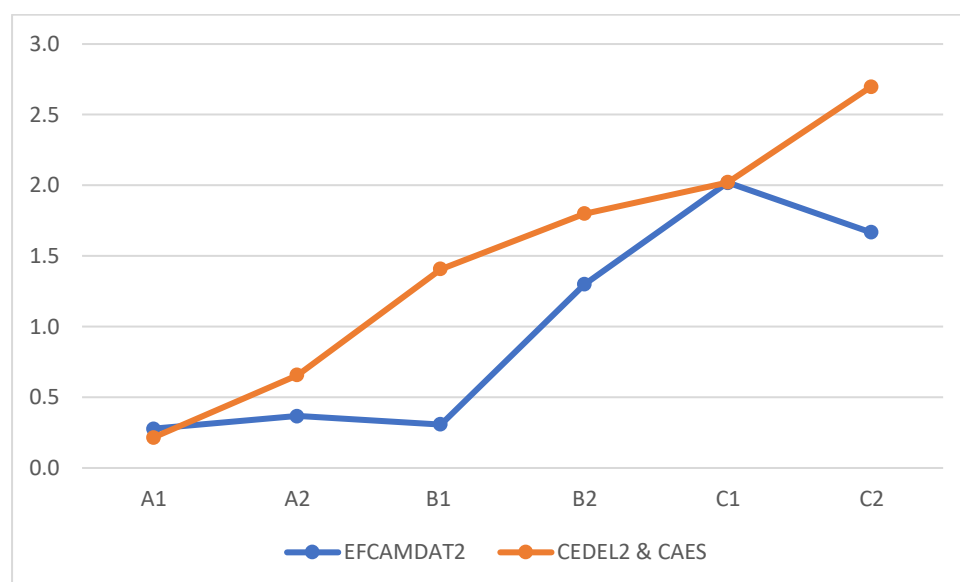


Figure 15 Level 2 Hypotaxis and Extension Results Summary.

4.4.1.2 Hypotaxis and Enhancement

Results for hypotaxis and enhancement (Figure 16) incorporated 45 measures which resulted in the largest representative sample of MBUs in the current study. MBUs extracted from the EFCAMDAT2 show a substantial increase from A1 (20.5) to A2 (60.4). Frequency in the EFCAMDAT2 continued to increase from A2 (60.4) to B1 (73.6) with a marginal decrease at B2 (73.1). There was a remarkable increase at C1 (106.5) followed by a significant decrease at C2 (49.2).

Frequency increased significantly in CEDEL2/CAES data from A1 (37.1) to A2 (62.7) with a constant upward curve continuing through B1 (85.0). From B1 (85.0) to B2 (82.3) there was a marginal decrease with a large increase at C1 (97.6). A slight decrease in frequency use was seen from C1 (97.6) to C2 (93.9). An interesting point is L2 English learners at the C1 level produced more MBUs in EFCAMDAT2 texts (106.5) versus the CEDEL2/CAES (97.6). C1 was the only proficiency level in which the MBU frequency was higher in the EFCAMDAT2.

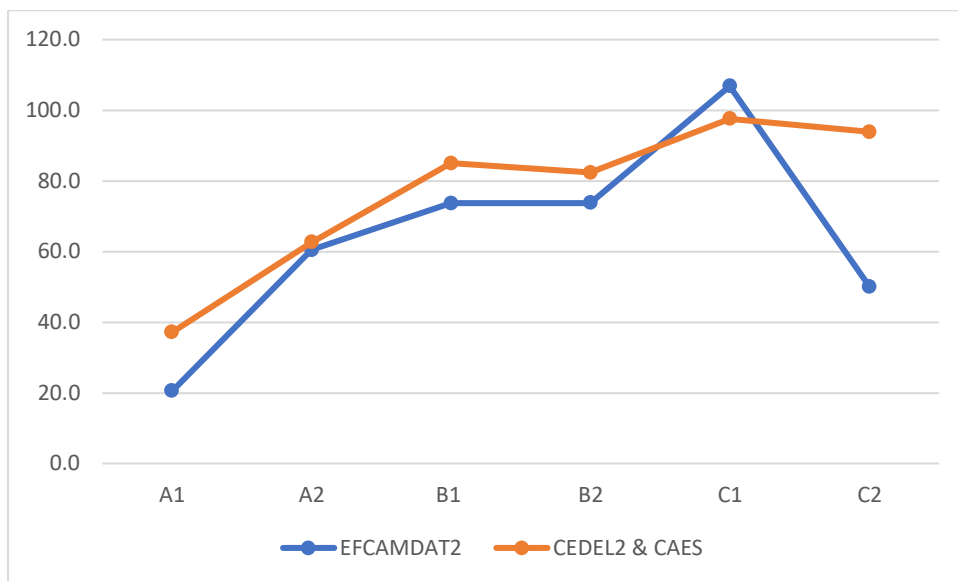


Figure 16 Level 2 Hypotaxis and Enhancement Results Summary.

4.4.1.3 Hypotaxis and Elaboration

Results for hypotaxis and elaboration incorporated a total of five measurements. The results (Figure 17) show the EFCAMDAT2 demonstrating increases from A1 (0.5) through to A2 (1.2) and B1 (2.2). Frequency increases continued in the EFCAMDAT2 from B1 (2.2) to B2 (3.5) continuing through C1 (4.5). A final decrease was seen at C2 (3.3). Data from the CEDEL2/CAES showed an increase from A1 (0.2) to A2 (0.6) and a sharp increase at B1 (4.5). An ensuing decrease appeared at B2 (2.1) with a secondary rise in frequency C1 (3.5) and C2 (3.9).

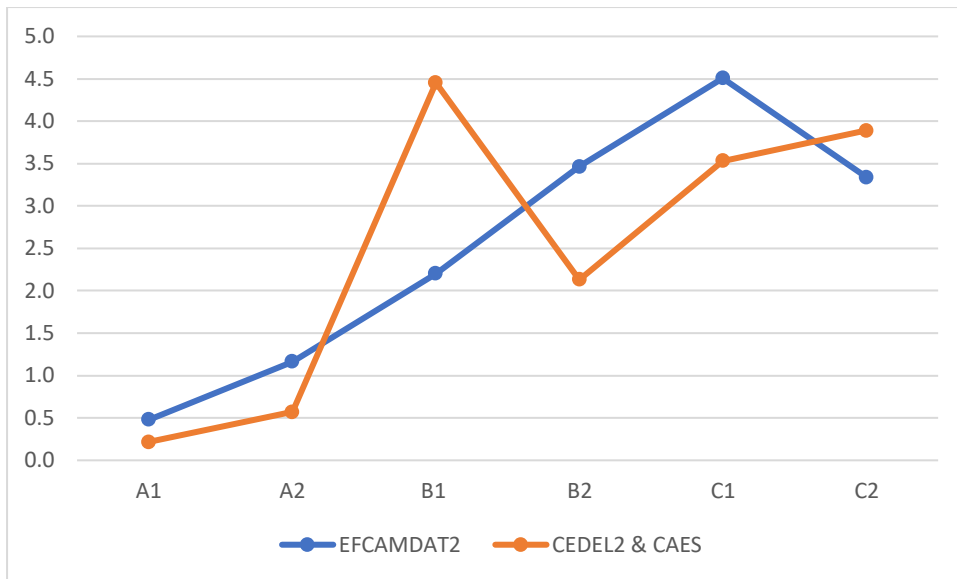


Figure 17 Level 2 Hypotaxis and Elaboration Results Summary.

4.4.1.4 Parataxis and Extension

The results for extension, the first category in parataxis, can be seen in Figure 18. Extension is by far the highest use frequency of MBUs in parataxis and in the entire study with a total of eight measures. Frequency in the EFCAMDAT2 increased from A1 (75.0) to A2 (101.4) and peaked at B1 (125.8). MBUs continually decreased from B1 (125.8) through the remaining levels of B2 (114.3), C1 (109.5) and C2 (102.6). The CEDEL2/CAES started at a higher frequency at A1 (145.5) then a decrease occurred to A2 (129.5) followed by an increase from A2 (129.5) to B1 (148.5). The resulting trend in the CEDEL2/CAES was a decrease from B1 (148.5) to B2 (146.8) followed by a slight increase from B2 (146.8) to C1 (148.7). A final decrease occurred from C1 (148.7) to C2 (146.7).

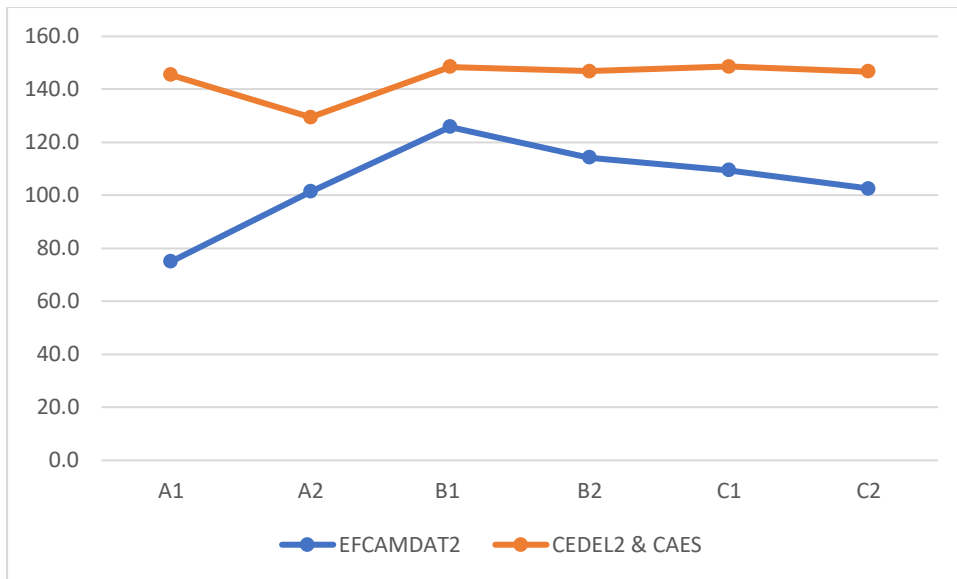


Figure 18 Level 2 Parataxis and Extension Results Summary.

4.4.1.5 Parataxis and Enhancement

Parataxis and enhancement (Figure 19) with 12 measures saw EFCAMDAT2 starting with peak frequency at A1 (5.7) and the continuous decreases at A2 (2.3), B1 (2.2), B2 (1.6), C1 (1.2) and zero frequency at C2. In contrast, the CEDEL2/CAES saw a rise in frequency from A1 (0.9) to A2 (1.2) and B1 (2.1) with a minor decrease at B2 (2.0). Increases were seen from C1 (2.6) to a peak at C2 (3.6).

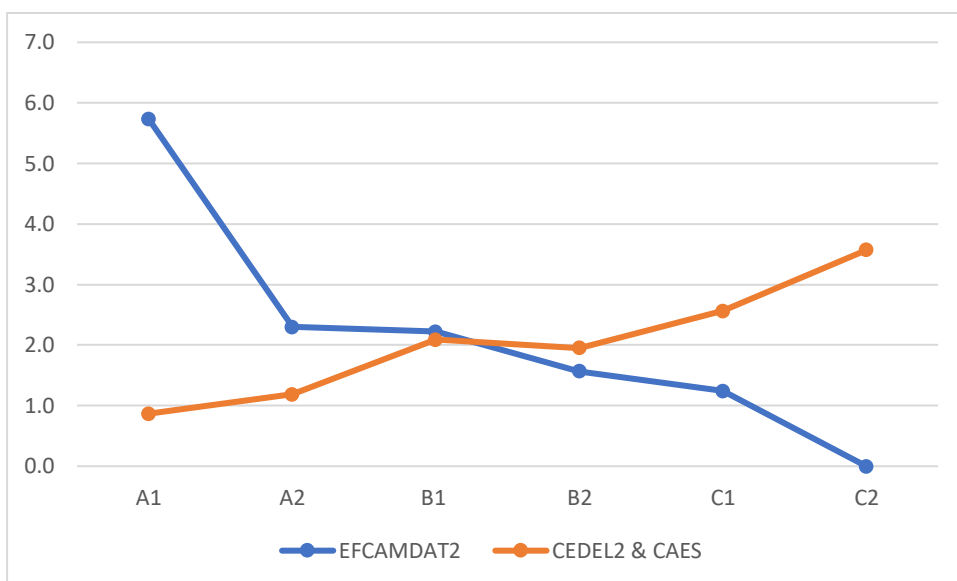


Figure 19 Level 2 Parataxis and Enhancement Results Summary.

4.4.1.6 Parataxis and Elaboration

The results for parataxis and elaboration are shown in Figure 20 with a total of 13 measures for both data sets. The EFCAMDAT2 showed an increase from A1 (1.7) through to A2 (6.1), B1 (10.5) until B2 (16.4). From B2 (16.4) to C1 (13.8) level there was a decrease in frequency followed by a sharp increase at C2 (24.2).

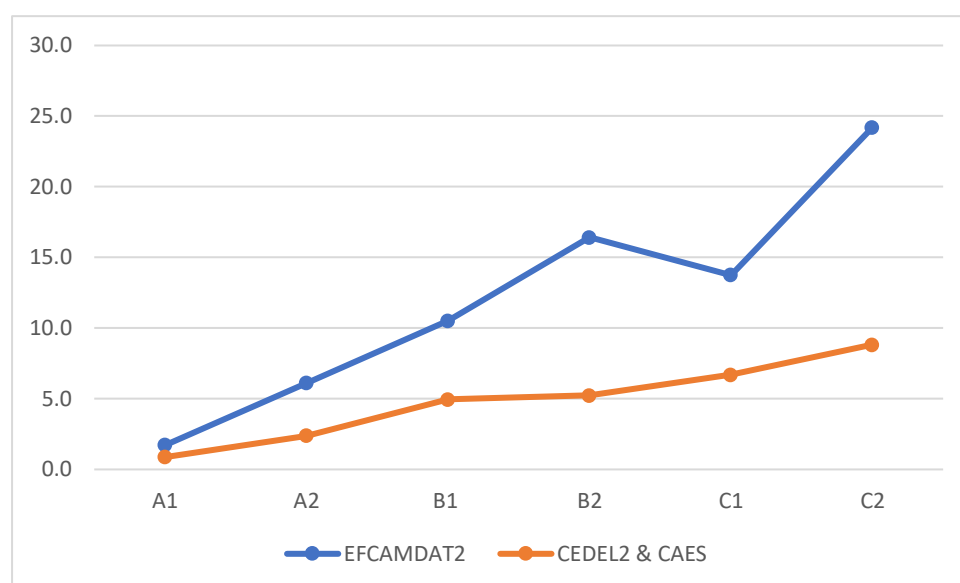


Figure 20 Level 2 Parataxis and Elaboration Results Summary.

The trajectory for the CEDEL2/CAES demonstrated a progressive increase through all CEFR proficiency levels. At the A1 (0.9) level the relative MBU frequency was rather low. Frequency increases started at A2 (2.4) and carried through to B1 (5.0), B2 (5.2) C1 (6.7) and culminated at the C2 (8.8).

4.4.2 Discussion

4.4.2.1 Hypotaxis and Extension

Results from the EFCAMDAT2 in the hypotactic expansion through extension show that Spanish L1 writing in English L2 use this resource at a very low frequency in A1 (0.3), A2 (0.4) and B1 (0.3). From B1 (0.3) to C1 (2.0) there is an increase followed by a decrease at C2 (1.7). Contrary to the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday &

Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998), hypotaxis through extension in the EFCAMDAT2 increased through the C1 level thus indicating a higher use of subordination through the lower advanced level rather than a decline.

Results involving hypotactic expansion through extension from the CEDEL2/CAES showed a continuous increase from A1 (0.2) to C2 (2.7). This upward trajectory across all six levels indicates subordination does not peak at intermediate levels and decreases with advanced proficiency thus not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in terms of hypotactic extension. Results from the EFCAMDAT2 and CEDEL2/CAES are consistent with previous studies (Sulistyaningrum & Rasyid, 2015; Wenhui Xuan, 2019) since hypotaxis and extension was used at a lower frequency than enhancement and elaboration. A line of research that is open for more investigation would involve more studies on how L1 English and Spanish use hypotaxis and extension in L2 texts.

The Sulistyaningrum and Rasyid (2015) study incorporated hypotaxis and parataxis with the categories of extension, enhancement and elaboration in recorded oral presentations from L1 Indonesian learners. The study focused on speaking and not writing with a small representative sample consisting of two groups. Since there are few SFL studies dealing with this granular line of research, results were deemed relevant. The first group composed of four female students failed to use any clauses related to hypotaxis and extension. In the second group involving four males, only 4% of the clauses were made up of hypotaxis and extension. The finding from the Sulistyaningrum and Rasyid (2015) study and the present study do correspond in the low use of hypotaxis and extension compared with a higher use frequency in hypotaxis through elaboration and enhancement.

Despite the Wenhui Xuan (2019) study having a lower representative sample and a different L1, results corresponded with the findings from the current study with hypotaxis and extension being used at a much lower percentage in written texts than either enhancement or elaboration. Like the current study, Wenhui Xuan showed elaboration was used at a slightly higher percent than extension, but enhancement was used with a much higher frequency than the other two categories.

The implications from the current study, as well as the Sulistyaningrum and Rasyid (2015) and Wenhui Xuan (2019) studies would introduce the idea that hypotaxis and extension is a category which EFL teachers may overlook or is not included in EFL curriculums. This is deduced by the low use of hypotactic extension MBUs in this category with Chinese, Indonesian, English and Spanish L1 backgrounds. The reasons for this could reside in the extremely high frequency use of hypotaxis and enhancement which is evident in the results of the current as well as the two previously mentioned studies.

4.4.2.2 Hypotaxis and Enhancement

Hypotaxis and enhancement was the largest representative sample in the EFCAMDAT2 as well as the CEDEL2/CAES. The general trend for the EFCAMDAT2 was an increase from A1 (20.5) to B1 (73.6). A leveling out from B1 (73.6) to B2 (73.1) was observed which took the form of a slight decrease in subordination. There was a sharp increase from B2 (73.6) to C1 (106.5). This finding suggests that MBUs for hypotaxis and enhancement reaches a peak at C1 which is contrary to the idea of subordination subsiding after the B2 level as per the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

A sharp decrease in hypotactic enhancement in the EFCAMDAT2 was seen from C1 (106.5) to C2 (49.2). Findings from the Chen et al. (2021) study observed the same phenomenon using the EFCAMDAT2 using a balanced sample from 31,040 texts involving 3.8 million words. While the current study and the Chen et al. (2021) study share the same corpus construct, an interesting observation is that a decrease in adverbial clauses from C1 (5,557) to C2 (1,458) appeared in L1 Brazilian, Chinese, Russian and Japanese texts. Grammatical metaphor might be a factor in the sharp decrease in adverbial subordination. Nevertheless, this study suggests a bias in the data sample with C1 having 64,358 words versus C2 with 5,994 words had a direct impact on findings in the EFCAMDAT2 data resulting in decreased subordination from the lower advanced to the upper advanced level.

The CEDEL2/CAES demonstrated a similar tendency until the C2 proficiency level. Results showed a frequency increase from A1 (37.1) until B1 (85.0) with a slight decrease at

B2 (82.3). Frequency for hypotaxis and enhancement increased from the B2 (82.3) to C1 (97.6) and showed a slight decrease at C2 (93.9). Findings for English L1 writing in Spanish L2 do not support the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

The CEDEL2/CAES decreased slightly from C1 (97.6) to C2 (93.9) possibly indicating another linguistic resource was being used or subordination was stabilizing. Since the representative sample for the CEDEL2/CAES is much higher at the upper advanced level, 109,284 words versus 5,994 words, results further indicate the EFCAMDAT2 corpus construction incorporates a global data bias which is evident at the C2 level in L2 English with L1 nationalities including Spanish, Brazilian, Chinese, Russian and Japanese.

The Sulistyaningrum and Rasyid (2015) study observed that out of a total of 63 occurrences of hypotaxis, enhancement was used a total of 43 times with the highest overall frequency. Understanding the constructs of the current study are different than EFL learners in Jakarta, there is a correlation in higher frequency of hypotaxis and enhancement use in both EFCAMDAT2 and CEDEL2/CAES when compared to elaboration and extension. This finding may suggest a cross-cultural linguistic phenomenon of higher hypotaxis through enhancement use despite the difference in the L1s. The other possible explanation may be EFL curriculums internationally have a strategic focus on adverbial subordination.

The Wenhui Xuan (2019) study also found that enhancement showed the highest frequency of use when compared to extension and elaboration. Although the study didn't incorporate a large data set, it did an analysis on 500 texts with various communicative tasks thus providing a fairly accurate representative sample for this grouping.

The Rasool and Mahmood (2023) study drew data using the International Corpus of Learner English to see what types of taxis were used. Results showed hypotaxis and enhancement was used at a much higher percentage in high and low-graded argumentative essays. Out of a total of 793 occurrences of hypotaxis in argumentative texts, 723 were categorized under enhancement. These findings from L1 Pakistani learners of L2 English correspond with the findings in which there was increased hypotactic enhancement in texts from the EFCAMDAT2 and CEDEL2/CAES.

4.4.2.3 Hypotaxis and Elaboration

Hypotaxis and elaboration, non-defining relative clauses, showed what has become a pattern with hypotaxis in the EFCAMDAT2. The overall trend increased from A1 (0.5) to C1 (4.5). This indicates non-defining relative clause usage peaked at the lower advanced level and only decreased at the higher advanced C2 level. Bearing this in mind, DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in hypotaxis and elaboration cannot be confirmed. Following the trend set in enhancement and extension, there was a decrease in elaboration from C1 (4.5) to C2 (3.3) which was also observed in the Chen et al. (2021) and thought to be because of the small representative sample.

The CEDEL2/CAES showed a different pattern emerging with an extremely low frequency from A1 (0.2) to A2 (0.6), then a sudden increase in hypotaxis and elaboration at the B1 (4.5) level followed by a substantial decrease at the B2 (2.1) level. From the B2 (2.1) to C2 (3.9) there appeared to be an increase, yet not to the height of frequency as seen at the peak B1 (4.5) level. A possible explanation for this happening is that adverbial subordination through hypotaxis and enhancement stabilizes between the B1 and B2 level thus resulting in a decrease in non-defining clause use.

The above-mentioned singular phenomenon doesn't suggest a correlation that subordination peaks at the intermediate levels since while peaking at B1, the trend did not continue through B2 and saw a secondary spike at C2, therefore not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). A phenomenon that further contraindicates the DSSICH is that elaboration increased from B2 (2.1) to C2 (3.9). This study recommends further research be carried out into English L1 use of Spanish L2 relative clauses in written tasks. An area of future research would be to expand the list of measures to see which non-defining relative clauses are used at a given proficiency level.

An explanation for the frequency spike in hypotactic elaboration at the B1 level might be explained by task-based formulaic sequences used to meet communicative needs. A formulaic sequence is a continuous or discontinuous sequence of words which appears

prefabricated. The prefabricated element is stored and retrieved from memory at time of use instead of being subject to spontaneous generation or analysis (Alexopoulou et al., 2015; Alhassan & Wood, 2015; Chen et al. 2021; Lewis, 1997; Wray, 2002).

The Alexopoulou et al. (2015) study resorted to this explanation after seeing a similar spike in the percentage of learners producing at least one relative clause from A2 to B1 in the first version of the EFCAMDAT in a study involving Brazilian, Chinese, German, Mexican and Russian L1s in L2 English written texts. They deduced that relative clause occurrence cannot be interpreted as direct evidence of grammatical knowledge. Therefore, productive use of relative clauses must be distinguished from formulaic sequencing. In an EFL context, task effects and formulaic sequences interact since the formulaic sequence can be part of a rehearsed lesson that learners are then asked to reproduce in writing.

Two vastly distinct studies (Sulistyaningrum & Rasyid, 2015; Wenhui Xuan, 2019) using different measures, L1s and methodologies found learners used slightly higher frequency of elaboration than extension which is concurrent with the results of the current study. In contrast, Rasool and Mahmood (2023) study found hypotaxis and elaboration to occur at a lower rate than extension. This might be due to their definition of hypotaxis and elaboration in (186). It should be noted, and in line with the definition taken from Halliday (2014), hypotaxis and elaboration is set apart by commas to add additional information to the clause. The below phrase doesn't contain commas, yet this type of clause would be more in tune with defining relative clauses which occur without commas. Therefore, the results of Rasool and Mahmood (2023) study for hypotaxis and elaboration are called into question by this study.

(186) They manipulate many people **which is subject of much research and discussion.**

4.4.2.4 Parataxis and Extension

Findings from EFCAMDAT2 and the CEDEL2/CAES for paratactic extension are consistent with previous studies (Rasool & Mahmood, 2023; Wenhui Xuan, 2019; Zarco-Tejada et al., 2016) involving SFL in which extension was observed to occur at a much higher frequency than either enhancement or elaboration. Results from the EFCAMDAT2

showed an increase in use of parataxis and extension from the A1 (75.0) to the B1 (125.8) level with a gradual decrease ensuing from B1 (125.8) to C2 (102.6). The trend in the EFCAMDAT2 does not support the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in which coordination in L2 written texts decreases at intermediate levels as subordination increases (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012).

In contrast, the results from the CEDEL2/CAES showed a decrease in parataxis from the A1 (145.5) to A2 (129.9) followed by an increase from A2 (129.9) to B1 (148.5). The overall tendency from B1 (148.5) to C2 (146.7) was a flattening out of the curve thus suggesting coordination does not increase or decreases as English L1 learners of Spanish L2 move through the intermediate and advanced levels. Regarding the proposal that coordination subsides at intermediate and upper levels, the data from the current study cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). Looking at the data which showed increases in MBU use in hypotactic enhancement through the C2 level, evidence would suggest there is a mutual relationship between the two largest and most indicative categories of paratactic extension and hypotactic enhancement.

This study suggests that parataxis and extension, the largest group MBUs for coordination, binds complex sentences involving adverbial subordination as English L1 learners of Spanish L2 progress through the intermediate and advanced CEFR proficiency levels. An explanation for this could be the representative sample and task requirements for the CEDEL2/CAES at the B2 through C2 levels. In the CEDEL2/CAES, there are more words produced by learners at these two levels and the average amount of words per script is greater (Table 81). While this study does not focus on sentence length, a general deduction when looking at the data would imply a text with more words would give way to longer sentences with a higher frequency of coordination as well as a higher frequency of adverbial subordination.

CEFR	EFCAMDAT2	CEDEL2/CAES
B2	157	337
C1	189	545
C2	187	471

Table 81 Average Words Per Script EFCAMDAT2 & CEDEL2/CAES CEFR Levels B2-C2.

4.4.2.5 Parataxis and Enhancement

The results for parataxis and enhancement were different from any other trend so far presented in this study in terms of L1 Spanish use of L2 English coordination MBUs in the EFCAMDAT2. From the A1 (5.7) to the A2 (2.3) there was a sharp drop in use of parataxis and extension which was due to the use of two MBUs. At the A1 level *and there* and *and then* were used at a higher frequency compared to the other ten measures. This could suggest a task-based bias in which formulaic sequences were retrieved to fulfill communicative needs (Alexopoulou et al., 2015; Chen et al., 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002).

The trend continually decreased from the peak at A1 (5.7) to C2 (0.0) as students' proficiency increased. This suggests that other linguistic resources were employed as there were notable increases in hypotaxis and enhancement and elaboration as well as parataxis and extension from the A1 to B2 level in the EFCAMDAT2. Results confirm that in the category of parataxis and enhancement coordination subsides from beginner to advanced levels (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012). Hence, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) can be confirmed for this category of coordinators.

Findings from the CEDEL2/CAES demonstrated a different trend as there was a general increase from A1 (0.9) to C2 (3.6) thus suggesting that L1 English writing L2 Spanish texts used this resource more often as acquisition increased. A possible explanation, as in parataxis and extension, is that coordination worked alongside with hypotaxis. Since there were generally longer texts in the CEDEL2/CAES and a mix of higher frequencies of parataxis and hypotaxis, the outcome may be that coordinating MBUs are linking more complex chains of subordinate clauses. The DSSICH (Byrnes et al., 2010; Colombi, 2002;

Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed since there were increases in parataxis and enhancement in L1 English writing in L2 Spanish as learners moved from A1 to C2.

When comparing the current study to the few studies found in this line of research which compared different L1s, paratactic extension and elaboration were used more often than enhancement in the Sulistyaningrum and Rasyid (2015) study. Results from the Rasool and Mahmood (2023) study showed parataxis and enhancement being used much less than extension but slightly more than elaboration. The Wenhui Xuan (2019) study found parataxis and enhancement to be used slightly less than elaboration, 1.49% versus 1.99%, while extension was by far the most often used category of coordination. What can be summarized when doing a cross-linguistic comparison across the previously mentioned studies is that enhancement and elaboration are used at a far lower frequency than extension thus suggesting students opt for simpler forms of coordination instead of more complex alternatives.

4.4.2.6 Parataxis and Elaboration

The final category in Level 2 is parataxis and elaboration. The general trend for the EFCAMDAT2 was a gradual increase from A1 (1.7) to C2 (24.2). One interesting pattern to underline is that unlike other categories in hypotaxis and parataxis, Spanish L1 writers of English L2 texts used parataxis and elaboration consistently from A1 to C2. Bearing this in mind, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed in paratactic elaboration.

One last noteworthy matter to highlight with parataxis and elaboration is that from the C1 (13.8) to the C2 (24.2) level this category shows a substantial increase in EFCAMDAT2. This is significant since in parataxis through extension and enhancement the opposite trend appeared suggesting instead a frequency decline. This particular finding is even more significant since a decline was observed in every category in hypotaxis from C1 to C2 which was generally surmised to be the result of the corpus construct having fewer words at the upper advanced level. The reason for this is two-fold. Firstly, L1 Spanish tended to use a high rate of discourse markers, words like *actually* and *for example*, when writing in L2 English.

Secondly, *wh*-exclamatory and interrogative markers were also used at higher frequencies in the EFCAMDAT2 than the CEDEL2/CAES.

The Bax et al. (2019) study researched L2 writers' use of metadiscourse markers at intermediate and advanced levels in Cambridge exams and found writers at C1 and C2 levels used significantly fewer discourse markers than the B2 level. This shows a sharp contrast to the current study in the sense that paratactic elaboration increased at the advanced levels. A caveat about the Bax et al. (2019) study is that although the authors mention that the L1 can play a part in the choice of discourse markers, in their study involving 281 metadiscourse markers in 13 categories involving 900 scripts, they never acknowledged which L1s were being researched.

Data extracted from the CEDEL2/CAES showed a gradual increase from A1 (0.9) to C2 (8.8). However, frequency of MBU use was much lower than the EFCAMDAT2 at every proficiency level. Taking this into account, the CEDEL2/CAES trend for parataxis and elaboration is not consistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Studies looking into parataxis and elaboration have shown mixed findings. Sulistyaningrum and Rasyid (2015) found parataxis and elaboration was used at a higher frequency than both extension and enhancement. The study was conducted on K-12 learners which didn't analyze results by proficiency levels. The younger age of K-12 students, which would possibly indicate a lower proficiency level, makes these results more interesting since this study found a higher use of parataxis and elaboration. This finding would contrast the idea of increased use of elaboration at higher levels thus assuming advanced students would be older. A possible explanation for this could range from L1 differences to the small representative sample versus the big data set used in the current study.

4.4.3 Conclusions

The first objective Level 2 takes on is to analyze extension, enhancement and elaboration to see how expansion is manifested in L2 acquisition and L1 to L2 transfer across proficiency levels. In terms of research question two, the DSSICH (Byrnes et al., 2010;

Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) was proven inconsistent in the EFCAMDAT2 data set in expansion through hypotactic extension, enhancement and elaboration. The MBF highlighted that subordination reached peak frequency in all three categories at the C1 proficiency level in EFCAMDAT2 data which was thought to add to the same phenomena occurring at Level 1. Results for research question three suggest the DSSICH should be offset by one level with C1 being the pivotal level as to when subordination subsides.

The current study observed a second pattern in which there was a significant decrease in frequency in the three categories at the C2 level in EFCAMDAT2 data. Said pattern also occurred in the Chen et al. (2021) study that saw a frequency decrease in subordination in four different L1s at the C2 level. This study proposes that the EFCAMDAT2 corpus construct influenced findings in both studies due to a significantly lower word count at C2.

CEDEL2/CAES data could not confirm research question two for the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in hypotactic extension, enhancement or elaboration. In terms of research question three, the MBF revealed hypotactic extension and elaboration to have much lower frequencies than enhancement. Furthermore, peak frequency for extension and elaboration was at the C2 proficiency level in the CEDEL2/CAES. On the other hand, hypotactic enhancement was the category with significantly higher frequencies over all CEFR levels, yet it reached a peak frequency at the C1. There was a small decrease from C1 to C2 in enhancement which may suggest a stabilization of subordination at the advanced level and is indicative of results seen in Level 1. Findings from the CEDEL2/CAES suggest that increased subordination is a sign of higher proficiency and that other linguistic resources, such as coordination and grammatical metaphor, might be used at the same time in conjunction with hypotaxis at advanced levels.

Research question four investigated how the MBF differentiated L1 to L2 transfer in Spanish and English written texts. L2 English writers used less subordination than L2 Spanish writers in the expansion categories of extension and enhancement. Contrary to Neff et al. (2004), the higher frequency of subordination used by L1 English writers might indicate more subordination is transferred to L2 writing. One possible factor that might influence data in the CEDEL2/CAES is that texts appeared to be longer since there was a word limit on

written tasks from the EFCAMDAT2 thus limiting length. Nevertheless, L2 English texts showed a slightly higher frequency of hypotaxis through elaboration than L2 Spanish texts thus suggesting L1 to L2 transfer might come into play with this type of linguistic resource.

Results from the EFCAMDAT2 and the CEDEL2/CAES confirm findings from previous studies using SFL meaning-based measures to gauge syntactic complexity in that hypotaxis through enhancement was used the most by L2 learners followed by elaboration and extension (Sulistyaningrum & Rasyid, 2015; Wenhui Xuan, 2019). A possible explanation comes from Halliday (2014) as he describes hypotaxis through extension as being rare and sometimes hard to detect (p. 474). Rasool and Mahmood (2023) found hypotactic extension to be used at a higher frequency than elaboration, yet enhancement had the overall highest frequency of the three classifications. Table 82 shows the frequency by ranking regarding how learners from different L1s used hypotactic extension, enhancement and elaboration. The number 1 means it was used with the highest frequency while 2 and 3 follow a descending order.

Study	L1	L2	Extension	Enhancement	Elaboration
EFCAMDAT2	Spanish	English	1	3	2
CEDEL2/CAES	English	Spanish	1	3	2
Sulistyaningrum & Rasyid (2015)	Indonesian	English	2	3	1
Zarco-Tejada et al. (2016)	Spanish	English	1	2	3
Wenhui Xuan (2019)	Chinese	English	1	3	2
Rasool & Mahmood (2023)	Urdu	English	1	2	3

Table 82 Frequency Ranking Paratactic Expansion Relations.

Data from the EFCAMDAT2 cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) for parataxis through extension or elaboration. However, data indicated parataxis and enhancement peaked at A1 and subsided through C2 thus confirming said hypothesis. In terms of the CEDEL2/CAES, the overall pattern in extension, enhancement and extension demonstrated that parataxis does not decrease after the A1 and A2 proficiency levels.

The MBF was used to distinguish syntactic complexity through CEFR proficiency levels to answer research question three. For parataxis and extension in the EFCAMDAT2, it was seen that B1 is the differentiating level as to when coordination peaks with an ensuing decrease in frequency. As for paratactic extension in the CEDEL2/CAES, the B1 level was also significant as it indicated when stabilization of coordination occurred with slight variations in frequency from B1 to C2.

Paratactic enhancement saw two opposite trends with frequency peaking in EFCAMDAT2 data at A1 and steadily decreasing to C2. The peak at A1 is thought to be due to task-based formulaic sequencing (Alexopoulou et al., 2015; Alhassan & Wood, 2015; Chen et al., 2021; Lewis, 1997; Wray, 2002) driven by increased use of *and there* and *and then*. The pattern for parataxis and enhancement revealed an increase in the CEDEL2/CAES from A1 to the highest frequency at C2. Lastly, both data sets observed an increase from A1 to C2 in paratactic elaboration thus showing that as learners acquire this type of coordination, their proficiency level increases.

Acknowledging research question four, parataxis and extension occurred at a much higher frequency in the CEDEL2/CAES than in the EFCAMDAT2, thus suggesting L1 coordination might be transferred to the L2. In parataxis and enhancement, learner data from the EFCAMDAT2 showed only slightly higher use frequency than the CEDEL2/CAES. L1 Spanish writing in L2 English used parataxis and elaboration in the EFCAMDAT2 at a much higher frequency than learners from the CEDEL2/CAES suggesting that L1 transfer to L2 might be taking place with these coordinators.

The frequency ranking by expansion categories showed several compelling trends in various studies. Findings involving paratactic extension from the EFCAMDAT2 and the CEDEL2 were consistent with three out of the four known studies (Rasool & Mahmood, 2023; Wenhui Xuan, 2019; Zarco-Tejada et al., 2016) in this line of research with only one study showing paratactic elaboration to be used at the highest frequency (Sulistyaningrum & Rasyid, 2015). This suggests that parataxis and extension is the most frequently used strategy by L2 learners from various L1s.

Results from both data sets in the current study along with studies by Sulistyaningrum and Rasyid (2015) and Wenhui Xuan (2019) showed that parataxis and enhancement was the least used category thus indicating that learners do not often rely on this type of expansion. Elaboration was the category with the most varying trends. With results from the EFCAMDAT2 and CEDEL2/CAES included, half of the studies analyzed reported elaboration as their second most frequently used type of parataxis (Table 82). Further research in parataxis, especially elaboration, would be beneficial in this line of investigation to add to the existing body of knowledge.

4.5 Experiment 3: Level 3 Meaning-Based Categories

Level 3 further dissects Level 1 taxis and Level 2 expansion by incorporating a more complex fine-grained measurement system defined by this study as meaning-based categories. The functional definition of meaning-based categories considers a more detailed level of analysis including markers such as addition, variation and alternation in extension. The rationale behind Level 3 is to see which trends are happening in a certain category and how they affect previous research questions. Level 3 sheds light on which type of SFL MBCs are used by L1s and at which CEFR proficiency level.

As a final note, hypotaxis and elaboration is not included in this section and will reappear at Level 4 in the results and discussion for meaning-based units. The reason for this is that it cannot be subdivided into MBCs due to the fact that SFL doesn't incorporate the same logico-semantic markers as for other categories such as enhancement and extension. The following section will start with results then move on to the discussion and end with conclusions.

4.5.1 Results

The results for Level 3 are summarized in Table 83. As this study moves forward, results are broken down into hypotaxis and parataxis through extension, enhancement and elaboration. The previously mentioned categories of expansion are then divided into addition, variation and alternation for extension. An example of this process would be hypotactic enhancement being separated into temporal, special, manner and causal-condition MBCs. As hypotaxis and elaboration is excluded at this level, paratactic elaboration includes exposition,

exemplification and clarification. As with Level 1 and 2, all MBCs are normalized by 5,000 words across CEFR proficiency levels.

Corpus	Levels 1 & 2	Level 3 MBC	M*	A1	A2	B1	B2	C1	C2
EFCAMDAT2 CEDEL2/ CAES	Hypotaxis Extension	Addition	4	0.2	0.3	0.2	0.9	1.5	0.8
EFCAMDAT2 CEDEL2/ CAES		Variation	2	0.2	0.2	0.4	0.6	0.7	0.9
EFCAMDAT2 CEDEL2/ CAES				0.0	0.0	0.1	0.3	0.4	0.8
EFCAMDAT2 CEDEL2/ CAES		Alternation	1	0.0	0.1	0.1	0.2	0.2	0.0
EFCAMDAT2 CEDEL2/ CAES				0.0	0.4	1.0	1.2	1.0	1.3
EFCAMDAT2 CEDEL2/ CAES		Hypotaxis Enhancement	Temporal	12	6.0	29.7	22.9	24.0	16.1
EFCAMDAT2 CEDEL2/ CAES	9.3				17.8	24.8	24.0	24.9	19.1
EFCAMDAT2 CEDEL2/ CAES	Spatial		1	1.5	2.2	3.0	3.3	4.2	0.8
EFCAMDAT2 CEDEL2/ CAES				2.0	4.2	5.3	6.1	3.8	1.9
EFCAMDAT2 CEDEL2/ CAES	Manner		3	0.2	1.0	2.2	4.0	7.1	5.0
EFCAMDAT2 CEDEL2/ CAES				0.4	0.8	0.8	2.4	3.0	4.3
EFCAMDAT2 CEDEL2/ CAES	Causal- condition	29	12.8	27.5	45.5	41.8	79.2	17.5	
EFCAMDAT2 CEDEL2/ CAES			25.4	39.8	54.2	49.8	65.9	68.7	
EFCAMDAT2 CEDEL2/ CAES	Parataxis Extension	Addition	5	64.9	91.3	117.7	106.8	99.0	85.1
EFCAMDAT2 CEDEL2/ CAES				138.1	126.6	143.0	140.4	143.0	139.1
EFCAMDAT2 CEDEL2/ CAES		Variation	2	1.5	2.3	1.4	3.3	3.3	5.8
EFCAMDAT2 CEDEL2/ CAES				0.0	0.5	1.2	0.8	0.6	1.0
EFCAMDAT2 CEDEL2/ CAES		Alternation	1	8.6	7.9	6.7	4.2	7.1	11.7
EFCAMDAT2 CEDEL2/ CAES				7.4	2.5	4.3	5.7	5.0	6.6
EFCAMDAT2 CEDEL2/ CAES	Parataxis Enhancement	Temporal	5	1.4	1.5	1.2	0.8	0.5	0.0
EFCAMDAT2 CEDEL2/ CAES				0.4	0.8	0.9	1.3	1.8	1.8
EFCAMDAT2 CEDEL2/ CAES		Spatial	1	4.3	0.6	0.8	0.5	0.4	0.0
EFCAMDAT2 CEDEL2/ CAES				0.2	0.0	0.1	0.0	0.0	0.2
EFCAMDAT2 CEDEL2/ CAES		Manner	2	0.1	0.0	0.0	0.1	0.0	0.0
EFCAMDAT2 CEDEL2/ CAES				0.2	0.4	1.0	0.5	0.6	1.5
EFCAMDAT2 CEDEL2/ CAES	Causal- condition	4	0.0	0.2	0.2	0.2	0.3	0.0	
EFCAMDAT2 CEDEL2/ CAES			0.0	0.0	0.0	0.2	0.1	0.1	
EFCAMDAT2 CEDEL2/ CAES	Parataxis Elaboration	Exposition	3	0.0	0.1	0.5	0.8	0.4	0.8
EFCAMDAT2 CEDEL2/ CAES				0.4	0.3	1.8	0.5	0.6	1.1
EFCAMDAT2 CEDEL2/ CAES		Exemplifi- cation	3	0.6	1.3	3.2	2.4	3.7	6.7
EFCAMDAT2 CEDEL2/ CAES				0.2	1.0	1.5	2.3	2.4	2.4
EFCAMDAT2 CEDEL2/ CAES		Clarification	7	1.1	4.8	6.8	13.2	9.6	16.7
EFCAMDAT2 CEDEL2/ CAES				0.2	1.1	1.7	2.4	3.6	5.4

* M equals measurements.

Table 83 Level 3 Results Summary.

4.5.1.1 Hypotaxis and Extension Through Addition

Results for hypotaxis and extension through addition (Figure 21) incorporating four measures in the EFCAMDAT2 showed a small increase from A1 (0.2) to A2 (0.3) then a slight decrease from A2 (0.3) to B1 (0.2). An increase in MBU use began at B1 (0.9) and continued through C1 (1.5) with a decrease from C1 (1.5) to C2 (0.8). There was an initial plateau in MBU frequency in the CEDEL2/CAES from A1 (0.2) to A2 (0.2). A gradual increase started at B1 (0.4) and went through to C2 (0.9).

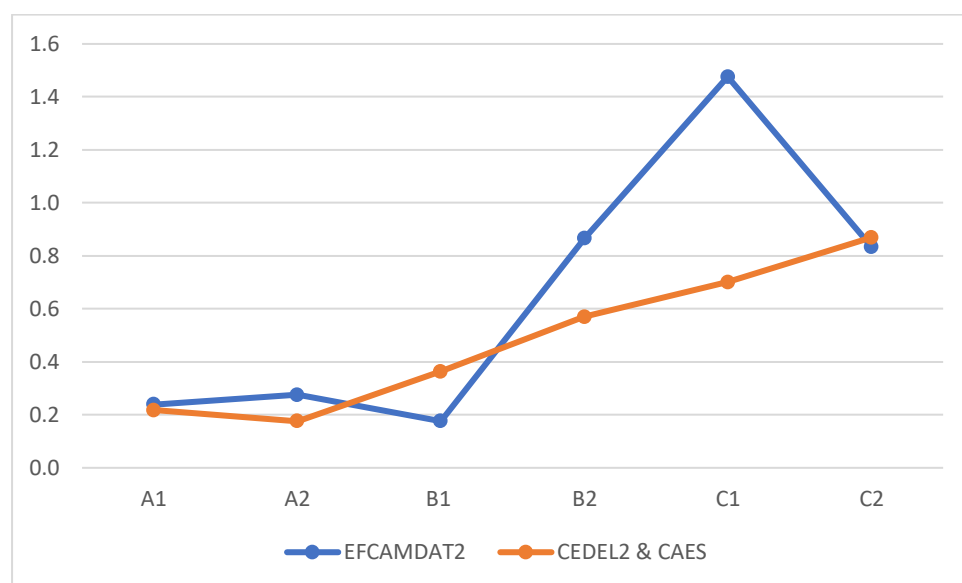


Figure 21 Level 3 Summary Hypotaxis and Extension through Addition.

4.5.1.2 Hypotaxis and Extension Through Variation

Figure 22 shows the results for hypotaxis and extension through variation with two measures. The EFCAMDAT2 and the CEDEL2/CAES showed similar tendencies as both had slight increases and decreases from A1 to B1. Data from the EFCAMDAT2 showed there was zero use of variation in A1 (0.0) and A2 (0.0) and slightly increased at B1 (0.1). An increase in the frequency of variation measures was seen in the EFCAMDAT2 from B1 (0.1) through B2 (0.3), C1 (0.4) and C2 (0.8). CEDEL2/CAES showed a slightly different trajectory, although with a lower overall use frequency in variation. There was no variation of use from A1 (0.0) to B1 (0.0). Increases were seen from B2 (0.1) to C1 (0.3) and C2 (0.5).

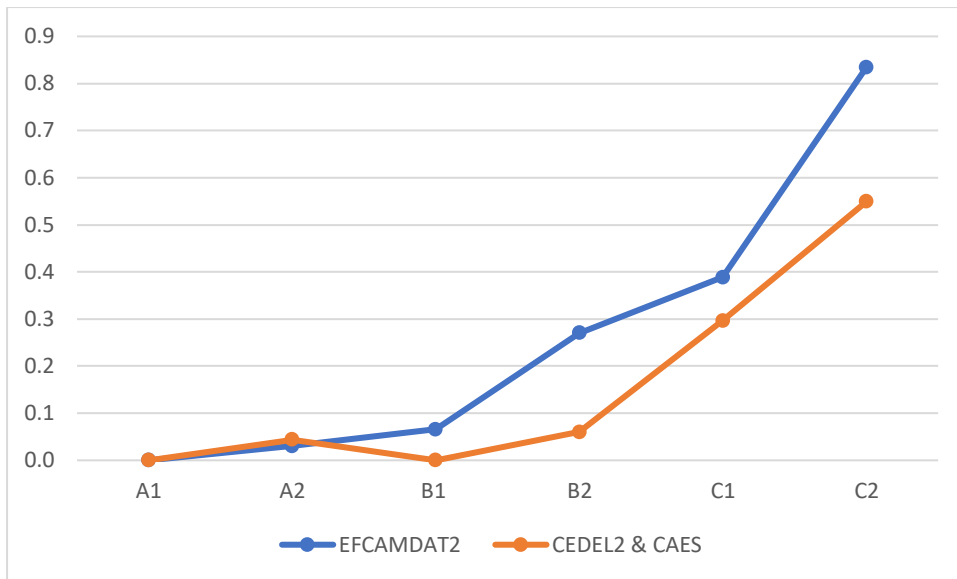


Figure 22 Level 3 Summary Hypotaxis Extension through Variation.

4.5.1.3 Hypotaxis and Extension Through Alternation

Results for hypotaxis and extension through alternation (Figure 23) show different trends for either data set with a representative sample of one measure. MBU frequency in the EFCAMDAT2 was almost nonexistent with an overall frequency starting at A1 (0.0) then slightly increasing at A2 (0.1), B1 (0.1), B2 (0.2) and C1 (0.2) followed by a final decrease at C2 (0.0). The CEDEL2/CAES displayed a different tendency since a gradual increase was seen from A1 (0.0), A2 (0.4), B1 (1.0) and ended at the B2 (1.2) level. From B2 (1.2) to C1 (1.0) there was a slight decrease with a following increase from C1 (1.0) to C2 (1.3).

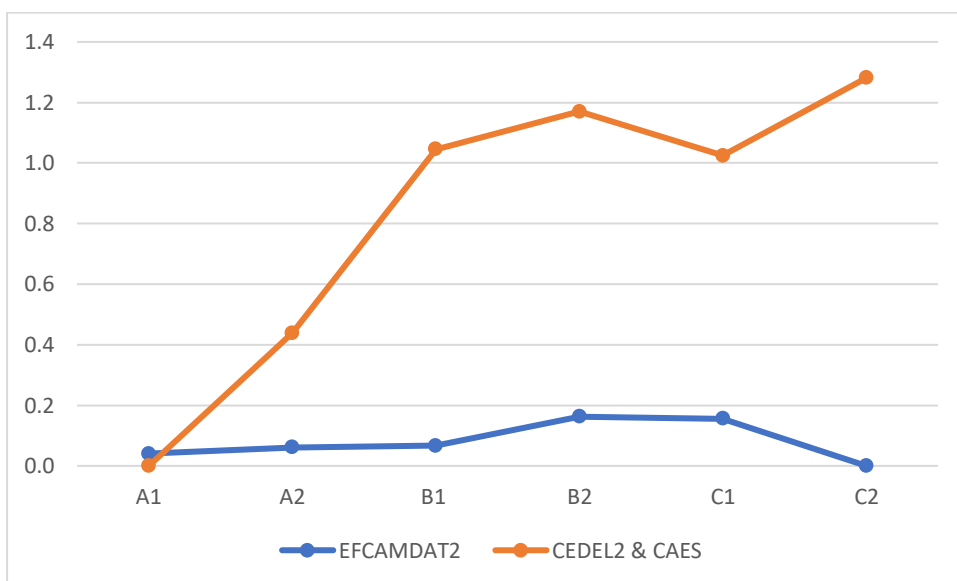


Figure 23 Level 3 Hypotaxis Extension through Alternation.

4.5.1.4 Hypotaxis and Enhancement Through Temporal Markers

MBCs found in hypotaxis and enhancement incorporated the largest number of measures in the entire study. The category representing temporal markers of enhancement (Figure 24) contained 12 measures with vastly different results. MBUs in the EFCAMDAT2 showed a large frequency increase from A1 (6.0) to A2 (29.7) with a decrease at the B1 (22.9) level. From the B1 (22.9) level to B2 (24.0) there was an incremental increase followed by a decrease from B2 (24.0) to C1 (16.1). From C1 (16.1) there was an increase to C2 (25.9), yet frequency never reached the previous peak seen at A2 (29.7). The data extracted from the CEDEL2/CAES showed an increase starting at A1 (9.3) to A2 (17.8) and B1 (24.8) with a decrease at B2 (24.0). A final increase occurred at C1 (24.9) followed by a decrease at C2 (19.1).

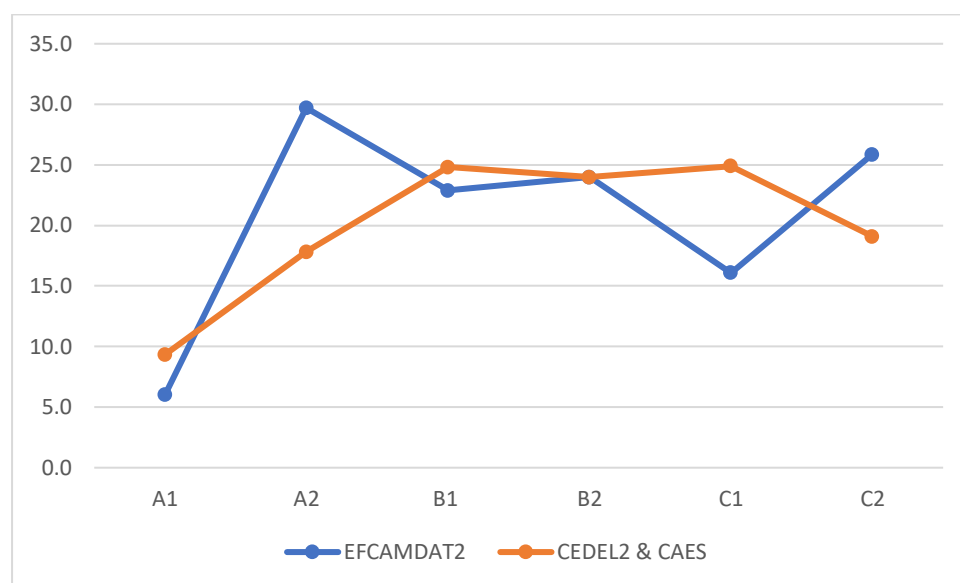


Figure 24 Level 3 Hypotaxis and Temporal Enhancement summary.

4.5.1.5 Hypotaxis and Enhancement Through Spatial Markers

Results for hypotaxis and enhancement (Figure 25) through spatial indicators with one MBU showed frequency in the EFCAMDAT2 increasing from A1 (1.5) through A2 (2.2), B1 (3.0), B2 (3.3) and C1 (4.2). There appeared a decline in frequency from C1 (4.2) to C2 (0.8). MBUs taken from the CEDEL2/CAES saw an initial increase from A1 (2.0) to A2

(4.2). A sharp increase was seen from A2 (4.2) to B1 (5.3) and then to B2 (6.1). The upward trajectory fell sharply from B2 (6.1) to C1 (3.8) and experienced another decrease at C2 (1.9).

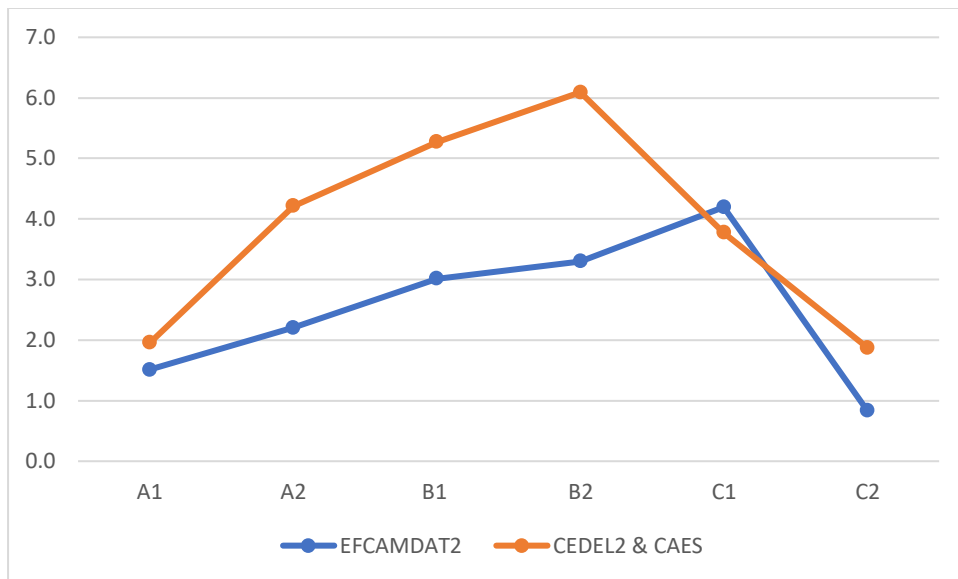


Figure 25 Level 3 Hypotaxis Spatial Enhancement Summary.

4.5.1.6 Hypotaxis and Enhancement Through Manner Markers

Findings for manner through hypotaxis and enhancement (Figure 26) showed a low frequency at the beginning levels in both data sets with a total of three measures. The EFCAMDAT2 saw increases at A1 (0.2), A2 (1.0), B1 (2.2), B2 (4.0) and C1 (7.1). From C1 (7.1) to C2 (5.0) there was a decrease. The CEDEL2/CAES had a slight increase in frequency from A1 (0.4) to A2 (0.8) and leveled off at B1 (0.8). A second larger series of increases appeared at B2 (2.4), C1 (3.0) and finally to C2 (4.3).

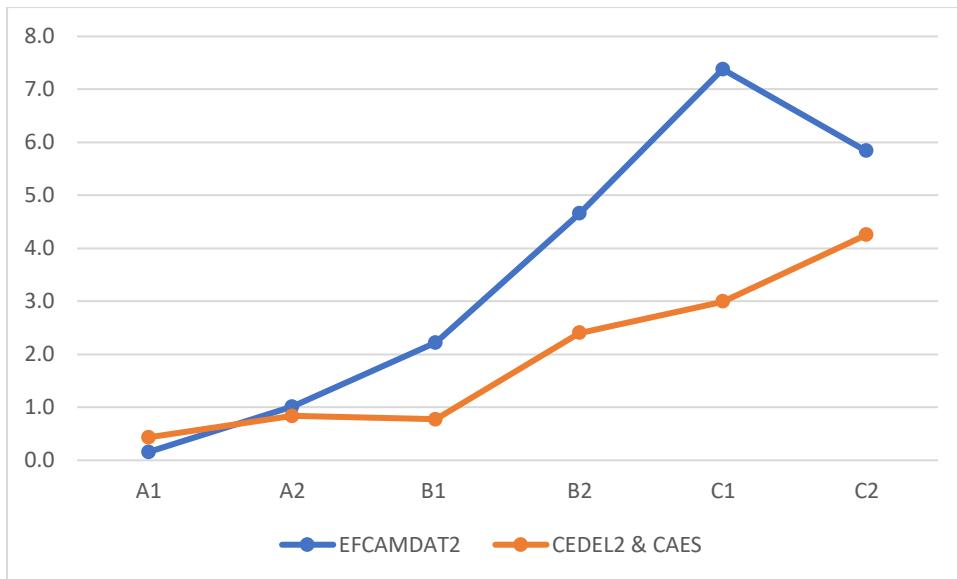


Figure 26 Level 3 Hypotaxis and Manner Enhancement Summary.

4.5.1.7 Hypotaxis and Enhancement Through Causal-Conditional Markers

Causal-conditional is the final category in hypotaxis (Figure 27). Results incorporated the overall largest number of measures with a total of 29. The EFCAMDAT2 and CEDEL2/CAES appeared to have followed a similar tendency until the B2 level with the CEDEL/CAES having overall higher MBC frequencies. EFCAMDAT2 data showed an increase from A1 (12.8) to A2 (27.5) then to B1 (45.5). There appeared a slight decrease in frequency at B2 (41.8) followed by a sharp increase to the C1 (79.2). A significant decrease appeared from C1 (79.2) to C2 (17.5). The CEDEL2/CAES showed an increase from A1 (25.4) to A2 (39.8) with the pattern ending at B1 (54.2). From B1 (54.2) to B2 (49.8) the frequency decreased. However, from B2 (49.8) there was an increase to C1 (65.9) and a second increase at C2 (68.7).

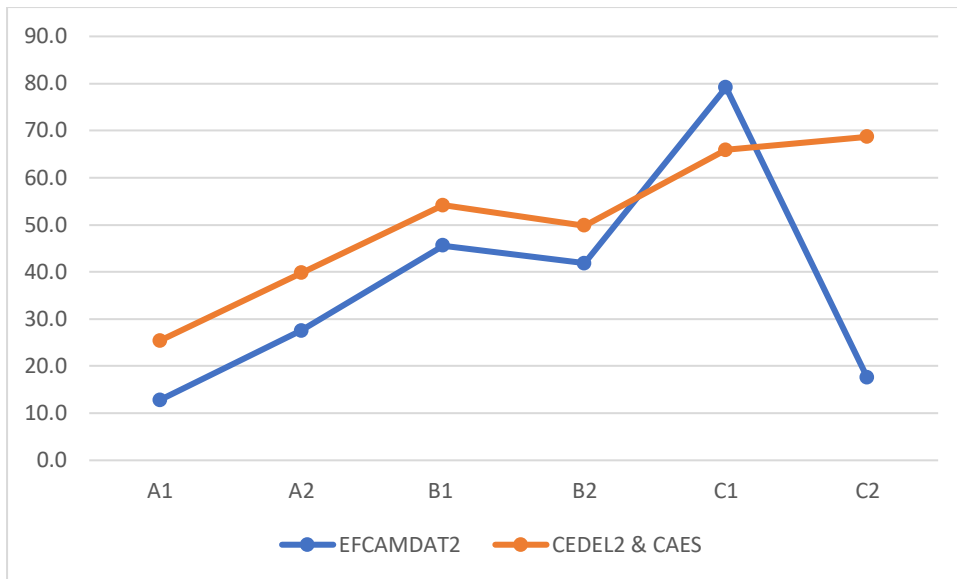


Figure 27 Level 3 Hypotaxis and Causal-conditional Enhancement Summary.

4.5.1.8 Parataxis and Extension Through Addition

Addition through parataxis and extension (Figure 28) incorporated five measurements yet had by far the highest overall MBU frequency in both data sets. The EFCAMDAT2 started with an elevated frequency from A1 (64.9) and continued through A2 (91.3) with the peak arriving at B2 (117.7). A general descending trend started at B2 (106.8) and continued through C1 (99.0) to C2 (85.1). The CEDEL2/CAES showed a decrease from A1 (138.1) to A2 (126.6) and then an increase to B1 (143.0). A relative plateau in the curve emerged starting at B1 (143.0) and continued through B2 (140.4), C1 (143.0) and finally ending without much variation at C2 (139.1).

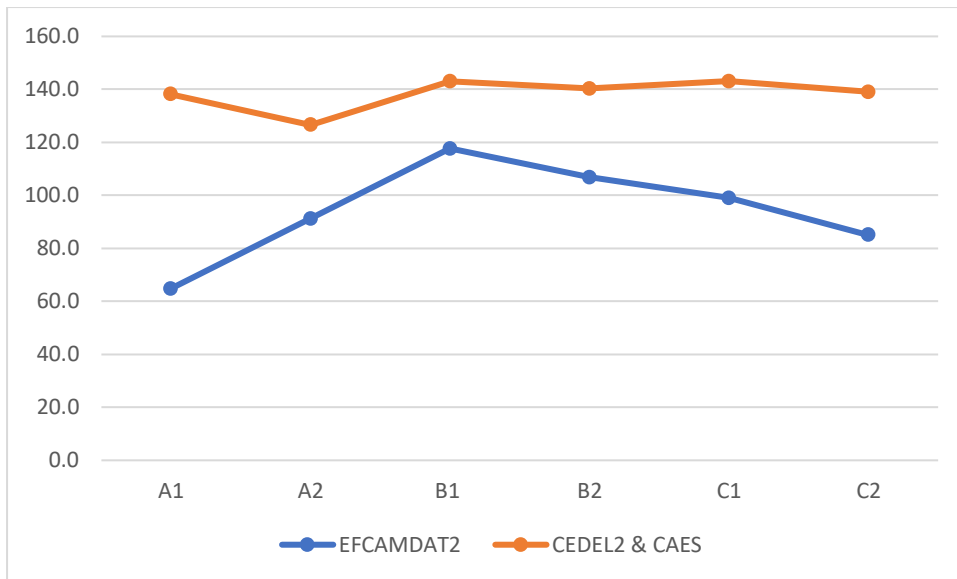


Figure 28 Level 3 Addition through Parataxis and Extension Summary.

4.5.1.9 Parataxis and Extension Through Variation

Variation in parataxis and extension (Figure 29) is the second grouping of MBUs including two measures. The EFCAMDAT2 started with an increase from A1 (1.5) to A2 (2.3) followed by a decrease at B1 (1.4). The pattern changed with an increase from B1 (1.4) to B2 (3.3). The trend continued with a leveling out of the curve from B2 (3.3) to C1 (3.3) and a final increase at C2 (5.8). The frequency trajectory in the CEDEL2/CAES manifested itself differently with zero MBU use at A1. There was an increase at A2 (0.5) which continued through B1 (1.2). From B1 (1.2) to B2 (0.8) there was a decrease which continued through C1 (0.6). From C1 (0.6) to C2 (1.0) there was a small increase.

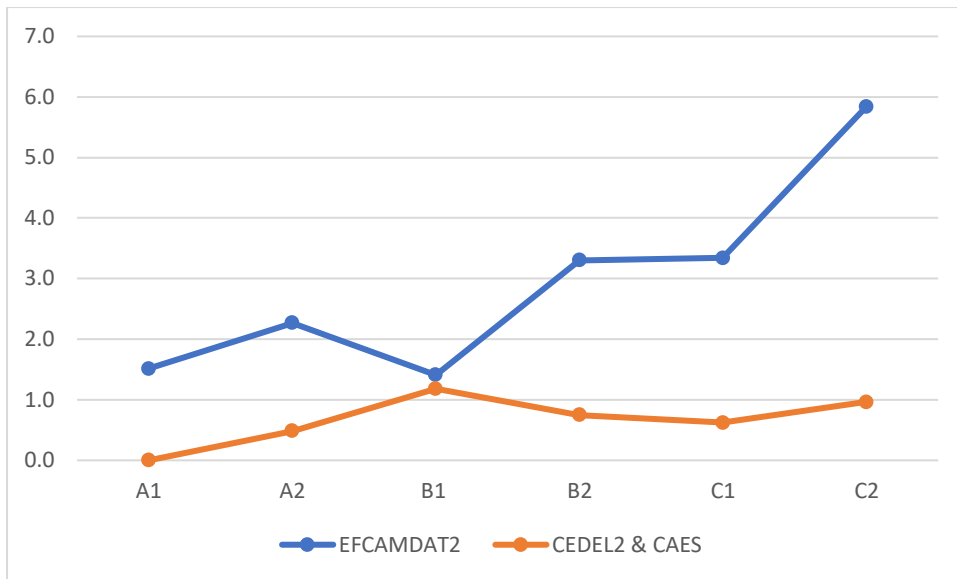


Figure 29 Variation through Parataxis and Extension Summary.

4.5.1.10 Parataxis and Extension Through Alternation

Data extracted from either the EFCAMDAT2 and the CEDEL2/CAES for alternation through parataxis and extension (Figure 30) showed a variety of patterns with only one measurement. The EFCAMDAT2 had an overall higher MBU density starting at A1 (8.6) with diminished frequency patterns seen at A2 (7.9), B1 (6.7) and B2 (4.2). An increase in frequency appeared from B2 (4.2) to C1 (7.1) and continued through C2 (11.7). Findings for CEDEL2/CAES displayed a variety of high and low frequencies. A1 (7.4) had the highest frequency with a drop in the curve to A2 (2.5). From A2 (2.5) to B1 (4.3) there was an apparent increase followed by a second increase at B2 (5.7). A slight decrease appeared at C1 (5.0) with another increase taking place at C2 (6.6).

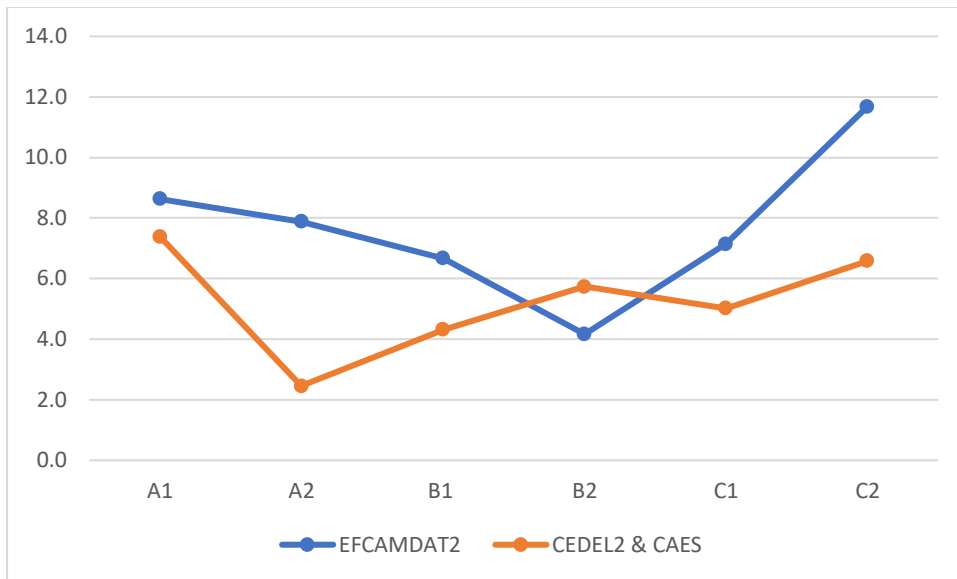


Figure 30 Level 3 Alternation through Parataxis and Extension Summary.

4.5.1.11 Parataxis and Enhancement Through Temporal Markers

The second category in parataxis is enhancement through temporal indicators (Figure 31). Results for either data set showed different patterns emerging in five measurements. The EFCAMDAT2 showed a slight increase from A1 (1.4) to A2 (1.5). Proficiency levels started decreasing at B1 (1.2), B2 (0.8) and C1 (0.5) with zero use at C2. The CEDEL2/CAES showed the opposite pattern with a low frequency at A1 (0.4) followed by a continuous augmentation in frequency at A2 (0.8), B1 (0.9), B2 (1.3) and C1 (1.8). C2 (1.8) saw a stabilization of frequency.

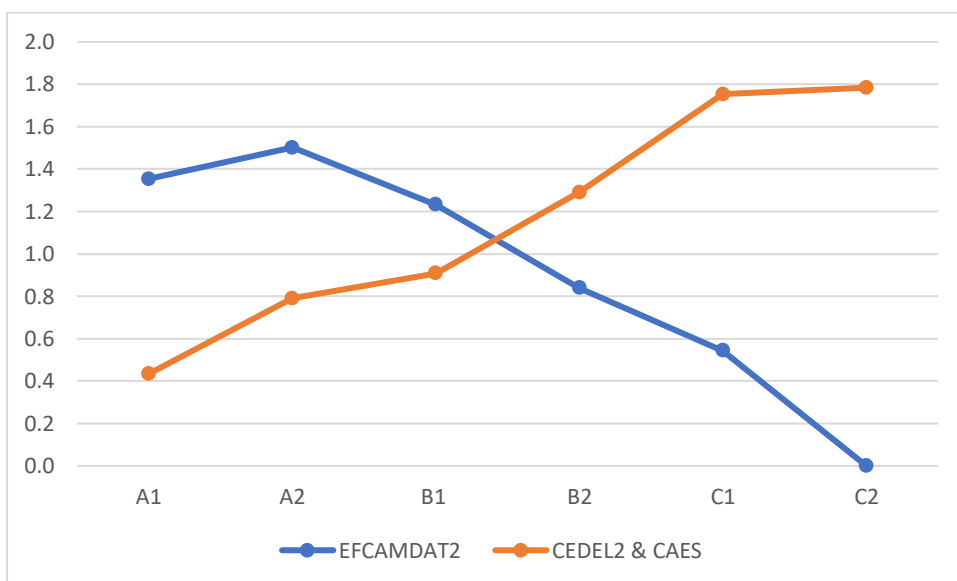


Figure 31 Level 3 Temporal Enhancement in Parataxis Summary.

4.5.1.12 Parataxis and Enhancement Through Spatial Markers

Findings in spatial enhancement in parataxis (Figure 32) with a single measurement showed an interesting pattern. The EFCAMDAT2 displayed a relatively high frequency at A1 (4.3) while decreasing at A2 (0.6). There was an increase from A2 (0.6) through B1 (0.8) then continuous decreases at B2 (0.5), C1 (0.4) and C2 (0.0). The CEDEL2/CAES exhibited a relatively low frequency with occurrence only happening at A1 (0.2), B2 (0.1) and C2 (0.2).

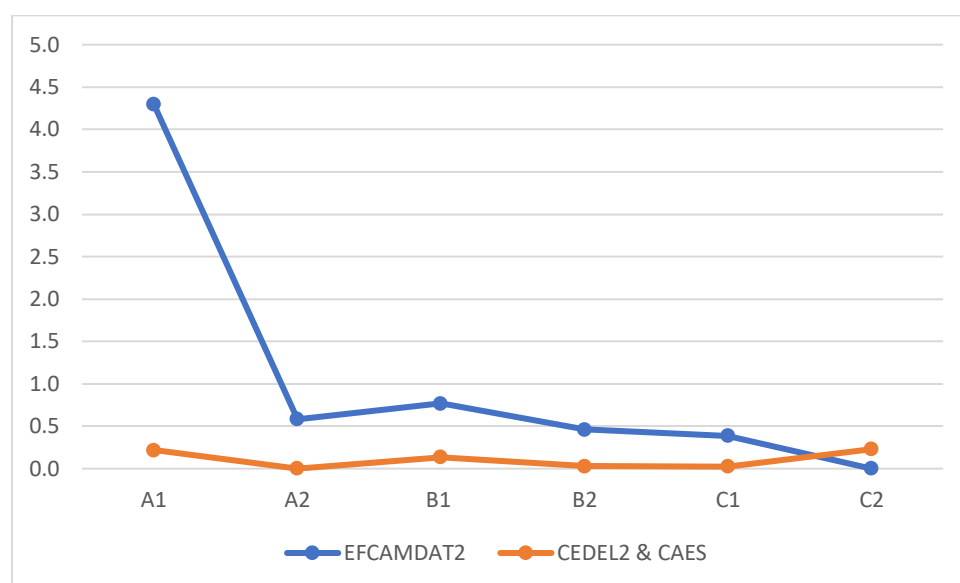


Figure 32 Level 3 Spatial Enhancement in Parataxis Summary.

4.5.1.13 Parataxis and Enhancement Through Manner Markers

The data sets for manner enhancement through parataxis (Figure 33) showed different patterns in two measurements with a much higher MBU frequency in the CEDEL2/CAES. The EFCAMDAT2 started with low frequency at A1 (0.1) followed by zero use at A2, B1, C1 and C2. B2 (0.1) was the only other level where frequency of manner MBUs appeared. The CEDEL2/CAES saw increases in frequency from A1 (0.2) through A2 (0.4) and B1 (1.0). There was a decrease at B2 (0.5) and an uptick at C1 (0.6) with another sharper increase at C2 (1.5).

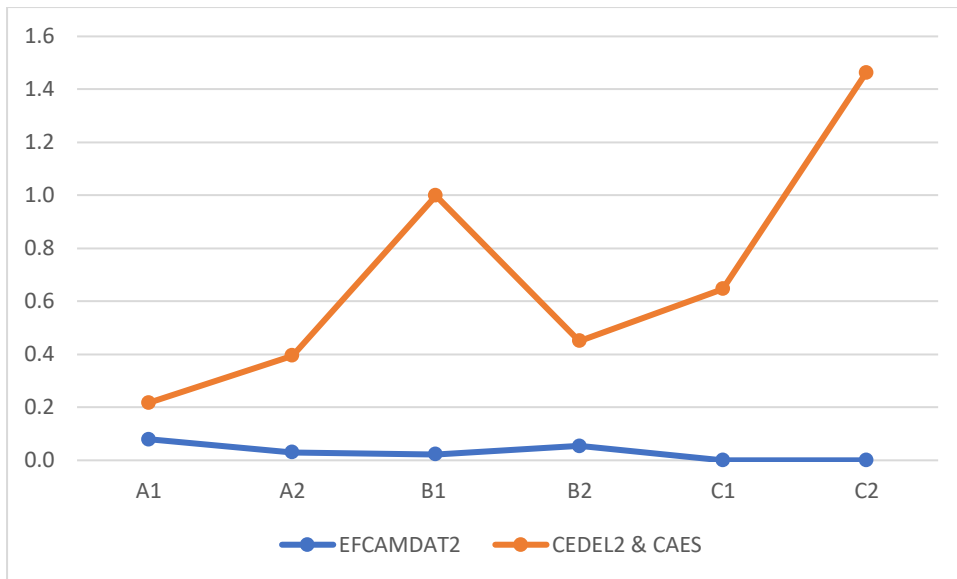


Figure 33 Level 3 Manner Enhancement through Parataxis.

4.5.1.14 Parataxis and Enhancement Through Causal-Conditional Markers

Findings for causal-conditional enhancement (Figure 34) through parataxis incorporated four measures. Frequency in the EFCAMDAT2 increased from A1 (0.0) to A2 (0.2). The pattern at A2 (0.2), B1 (0.2), and B2 (0.2) remained unchanged. There was a small increase at C1 (0.3) followed by a decrease at C2 (0.0). Data extracted from the CEDEL2/CAES showed a low frequency of MBU use with zero use at A1 through B1. There was an increase at B2 (0.2). A slight decrease occurred from B2 (0.2) to C1 (0.1) and leveled off at C2 (0.1).

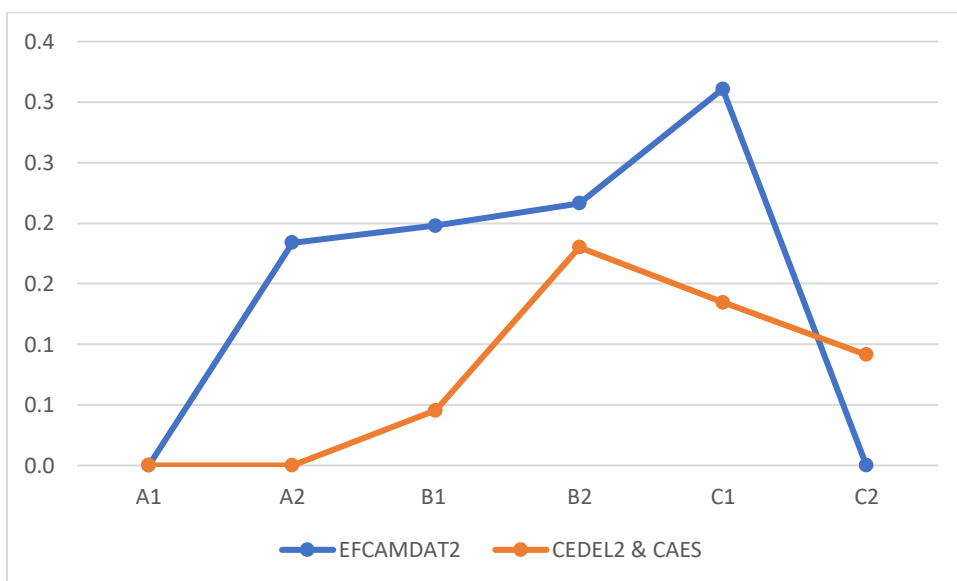


Figure 34 Level 3 Causal-conditional Enhancement through Parataxis Summary.

4.5.1.15 Parataxis and Elaboration Through Exposition

Exposition through elaboration in parataxis (Figure 35) included three measures. Findings from the EDCAMDAT2 showed an increase from A1 (0.0) to A2 (0.1), B1 (0.5) and B2 (0.8). From the B2 (0.8) to C1 (0.4) level there was a slight decrease followed by an increase at C2 (0.8). The CEDEL2/CAES exhibited a decrease from A1 (0.4) to A2 (0.3) then a peak at B1 (1.8) with a decrease at B2 (0.5). There was a small increase at C1 (0.6) and a larger one at C2 (1.1).

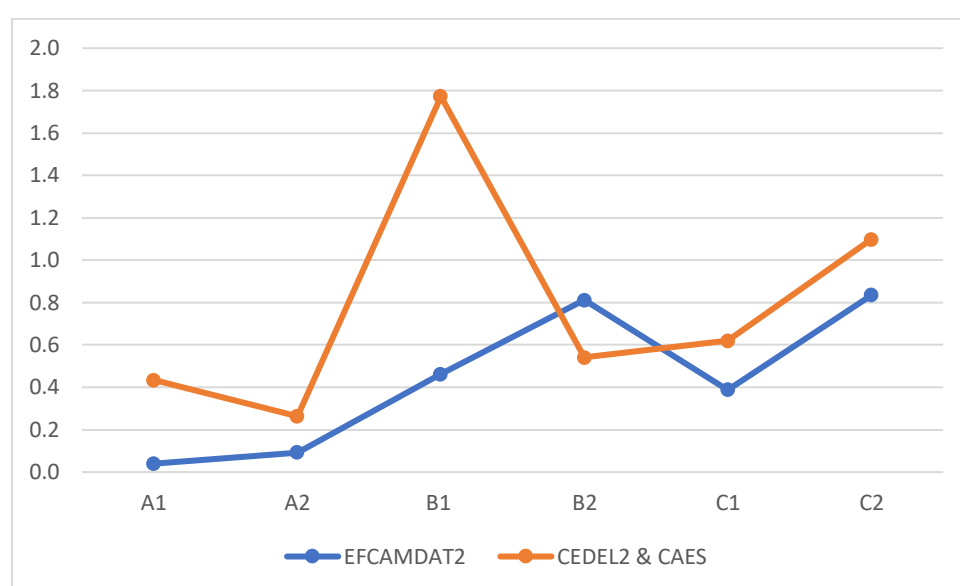


Figure 35 Level 3 Exposition through Paratactic Elaboration Results Summary.

4.5.1.16 Parataxis and Elaboration Through Exemplification

Data extracted for exemplification through elaboration in parataxis (Figure 36) entailed three measurements showing two different emerging patterns. The EFCAMDAT2 increased in frequency from A1 (0.6) to A2 (1.3) and B1 (3.2). There was a decrease at B2 (2.4) with a second upswing in frequency at C1 (3.7) and C2 (6.7). The CEDEL2/CAES had continuous increases in every level starting at A1 (0.2) and moving through A2 (1.0), B1 (1.5), B2 (2.3) and C1 (2.4). Frequency plateaued from C1 (2.4) to C2 (2.4).

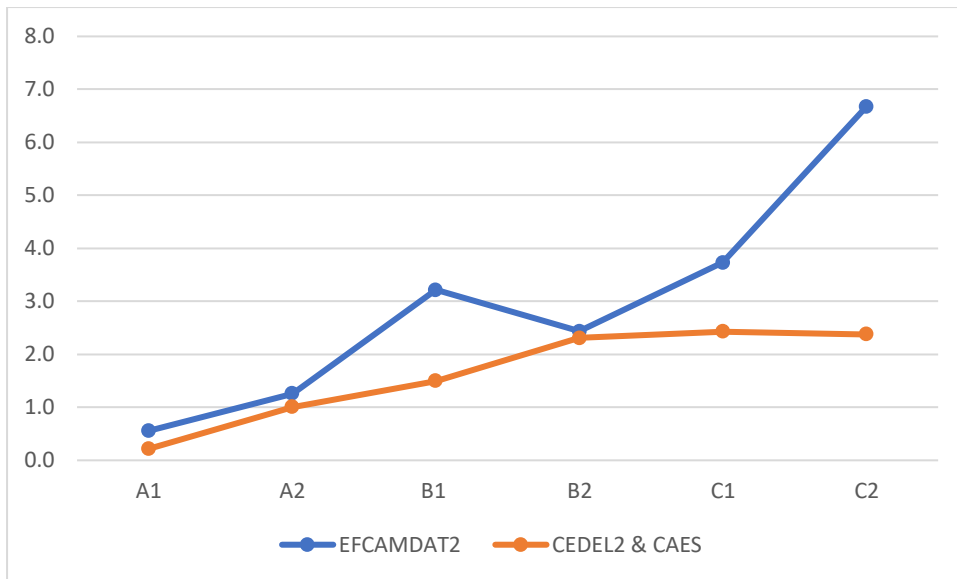


Figure 36 Level 3 Exemplification through Elaboration in Parataxis Results Summary.

4.5.1.17 Parataxis and Elaboration Through Clarification

Results for clarification through elaboration in parataxis (Figure 37) included seven different measurements with different patterns appearing in either data set. The EFCAMDAT2 had a much higher frequency of MBU use in this classification starting at A1 (1.1) and increasing through A2 (4.8), B1 (6.8) and B2 (13.2). From B2 (13.2) there was a decrease at C1 (9.6) with following increase at C1 (16.7). The CEDEL2/CAES showed a gradual increase throughout proficiency levels starting at A1 (0.2) and continuing through A2 (1.1), B1 (1.7), B2 (2.4), C1 (3.6) and peaking at C2 (5.4).

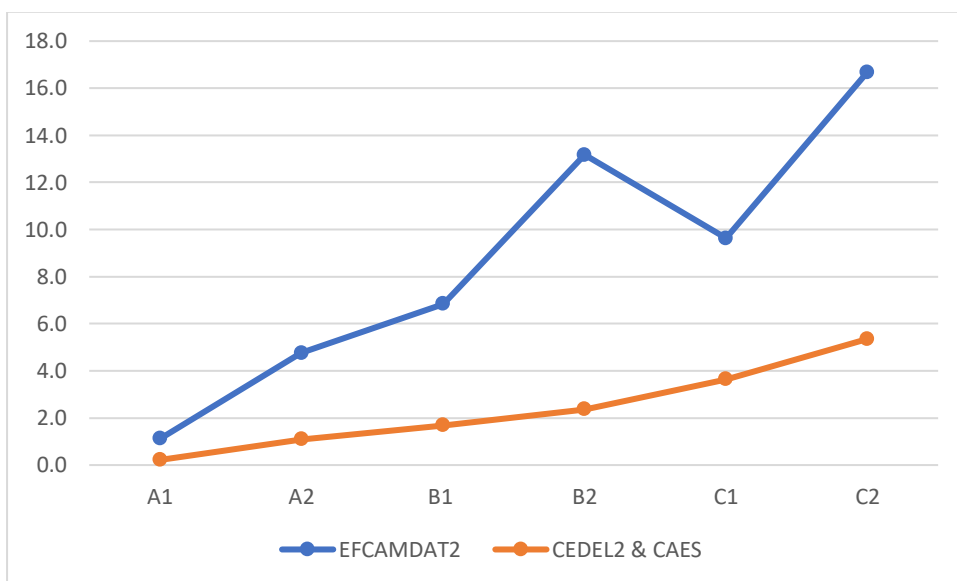


Figure 37 Level 3 Clarification and Elaboration through Parataxis Result Summary.

4.5.2 Discussion

4.5.2.1 Hypotaxis and Extension Through Addition

Findings from the EFCAMDAT2 suggest hypotaxis and extension through addition continues to increase until the C1 level thus providing evidence that subordination may be an often-used linguistic resource through the lower advanced level. Data from the CEDEL2/CAES indicates that use of extension increases through C2. In the case of both data sets, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quintero et al., 1998) cannot be confirmed for addition through extension. EFCAMDAT2 texts showed an overall higher MBU frequency than those from the CEDEL2/CAES.

4.5.2.2 Hypotaxis and Extension Through Variation

Variation through hypotaxis and extension for the EFCAMDAT2 saw a low frequency in MBU use with two measures used in the study. The overall pattern implies this type of subordination is used more in the intermediate and advanced levels. The overall high numbers in paratactic extension at beginner levels may imply that learners use hypotactic extension less. The pattern in the CEDEL2/CAES showed a steady increase, while at a low frequency, from B2 (0.1) to C2 (0.5). Findings from the EFCAMDAT2 and the CEDEL2/CAES indicate that variation, while used at a low frequency, only increases as proficiency increases thus not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quintero et al., 1998). L2 English writers used variation at a higher frequency than L2 Spanish ones.

4.5.2.3 Hypotaxis and Extension Through Alternation

Alternation is the final category in hypotaxis and extension which incorporated one measure making for a small representative sample in both data sets. Frequency for the single MBU in the EFCAMDAT2 displayed a low frequency at A2 (0.1) and B1 (0.1) with a slight

increase at B2 (0.2) and C1 (0.2). Learners did not use alternation at C2. The CEDEL2/CAES showed an increase from A2 (0.4) to C2 (1.3). The trajectory in the frequency curve for alternation through extension in hypotaxis indicates as frequency increases, so does proficiency until the advanced level. Findings from the EFCAMDAT2 and the CEDEL2/CAES do not confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). In a different trend from addition and variation, alternation saw L2 Spanish learners using the single MBU at a higher frequency than English L2 learners.

4.5.2.4 Hypotaxis and Enhancement Through Temporal Markers

The first MBC in enhancement consists of temporal subordinators with a total of 12 measurements. The EFCAMDAT2 saw a sharp increase in temporal subordination from A1 (6.0) to A2 (29.7) which was largely driven by the use of the single MBU *when*. This suggests a possible influence of task-based formulaic sequences in which learners recall a specific word or sequence of words based on communicative demands (Alexopoulou et al., 2015; Chen et al., 2021).

Frequency in the temporal MBC in the EFCAMDAT2 decreased from A2 (29.7) to B1 (22.9) then increased at B2 (24.0) with another decrease at C1 (16.1) followed by a final increase at C2 (25.9). Although the overall relative frequency remained elevated, temporal subordination was used less at higher levels in the EFCAMDAT2. The decrease in frequency at the C1 level corresponds with an increase of frequency in the categories of spatial, manner and causal conditional hypotaxis through enhancement thus indicating learners might use other forms of subordination causing a decrease in frequency at the lower advanced level for temporal indicators.

The uptick in frequency from C1 (16.1) to C2 (25.9) in the EFCAMDAT2 was due to an increase of overall use of temporal subordinators. Unlike at the A2 level, which was driven by the MBU *when*, English L2 writers used a larger variety of subordinators at a higher frequency at advanced levels. Moreover, this would correspond with a decrease in subordination use in spatial, manner and causal-conditional MBCs thus suggesting learners

returned to using this linguistic resource at the C2 level while frequency dropped in every other MBC in enhancement.

The erratic nature of frequency through proficiency levels for the temporal MBC in the EFCAMDAT2 lead to a variety of conclusions. Apart from C1, learners used temporal indicators in levels A2 through C2 at a frequency between 29.7 and 24.0 thus indicating a variation of 5.7 words per 5k. Temporal indicators seemed to plateau between A2 (29.7) and C2 (25.9) which is thought to be due to formulaic sequencing. This study cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) for temporal indicators in hypotaxis and enhancement since the highest frequency occurred at A2 with the second peak at C2. With that said, results remain more inconclusive than clear since there was not a clear pattern of lower subordination at beginner levels or higher subordination at advanced levels. It is suggested by the current study that temporal indicators are used as a stable linguistic tool throughout A2, B1, B2 and C2. Furthermore, further research investigating how genre affects the frequency at proficiency levels might account for fluctuations.

The results from the CEDEL2/CAES suggested hypotaxis and enhancement temporal markers were used at an ever-increasing frequency from A1 (9.3) to B1 (29.7). The plateau in the curve from B1 (29.7) to C1 (24.9) suggests temporal subordination stabilized at the intermediate and lower advanced levels. From the C1 (24.9) to C2 (19.1) there was an overall decrease in subordination which might imply other linguistic resources were being used. Findings from the CEDEL2/CAES do not support the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) for temporal indicators in hypotaxis and enhancement as subordination remained stable between B1 and C1.

The average frequency across all six proficiency levels for learner data in the EFCAMDAT2 showed a slightly higher frequency of temporal subordination use (20.7) versus learners from the CEDEL2/CAES (19.9). Nonetheless, temporal subordination is a linguistic resource favored by learners in either data set. A trend that appeared was a plateau effect in subordination in both data sets. While EFCAMDAT2 data showed higher frequencies at A2 (29.7) and C2 (25.9), there was only a 1.1 variation between B1 (22.9) and B2 (24.0) thus indicating a stabilization of frequency. The CEDEL2/CAES saw a frequency

plateau which appeared with only a 0.9 variation between B1 (24.8), B2 (24.0) and C1 (24.9) thus indicating that temporal subordination is a staple of intermediate and lower advanced writers. This study introduces the idea in which this group of subordinators is used as a constant linguistic resource by learners and new parameters for this type of high frequency meaning-based subordination need to be established by future research.

The only comparable study to be found considering hypotaxis and enhancement through temporal markers was carried out by Rasool and Mahmood (2023). While they didn't incorporate CEFR proficiency levels, their findings showed the frequency of temporal indicators used by L1 Pakistani learners of English L2 when compared to other MBCs. Results from their study and the current study indicate that hypotaxis through temporal enhancement is the second most often used linguistic resource by Pakistani, Spanish and English L1s in written L2 texts. The question which comes to mind is what causes the frequency correlation between use of temporal indicators in three separate L1s? Further research should focus on the root causes of this phenomenon.

4.5.2.5 Hypotaxis and Enhancement Through Spatial Markers

Spatial MBUs in hypotaxis and enhancement incorporating one MBU showed an increase from A1 (1.5) to C1 (4.2) and a sharp decline to C2 (0.8) in the EFCAMDAT2. Data observing continued frequency increases through the C1 level confirms that Spanish L1 writing in English L2 use spatial subordination as an ongoing resource and does not support the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

The CEDEL2/CAES showed a trend resulting in increased subordination from A1 (2.0) to B2 (6.1) and then a following decrease to C2 (1.9) in spatial enhancement. The phenomena suggests that advanced levels use different linguistic resources. Therefore, spatial enhancement results support the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). An explanation for this phenomenon might be higher subordination density of MBUs in temporal and causal-conditional enhancement. The frequency of subordination increases substantially from the B2 to C2 level in the causal-conditional MBC thus suggesting learners

are using this type of subordination instead of other resources such as spatial subordination. It is also worth mentioning there is only one measure within the spatial MBC which could affect results.

EFCAMDAT2 findings concur with the Rasool and Mahmood (2023) study in that spatial indicators were used at a lower frequency than any other MBC in enhancement. In contrast, texts in the CEDEL2/CAES showed the single spatial indicator to be the second least used MBC before indicators of manner. It was observed that L2 Spanish texts contained more spatial subordination at a higher frequency than L2 English texts.

4.5.2.6 Hypotaxis and Enhancement Through Manner Markers

Manner through hypotaxis and enhancement incorporated three measures. The EFCAMDAT2 saw an overall increase from A1 (0.2) to C1 (7.1) with a decrease to C2 (5.0). The overall increase from beginner to the lower advanced level shows that, in the case of manner, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) is not applicable.

The CEDEL2/CAES showed a continual increase from A1 (0.4) to C2 (4.3) with the highest frequency of hypotactic enhancement through manner seen at C2 thus indicating findings are not consistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). Results show this type of subordination doesn't decrease yet is used more often by learners with a higher proficiency level.

EFCAMDAT2 learners used manner subordinators at a higher frequency than learners in the CEDEL2/CAES. Both data sets saw continued increases in manner subordination from beginner to advanced levels. This finding is inconsistent with Lorenzo et al. (2019) as they found learners ranging from puberty to late adolescence start to produce manner adverbials first and at an older age grammaticalization occurs in the form of concessive subordination (p. 2). While the present study contains a broader category of learners, neither data set showed a high use frequency of manner adverbials at lower proficiency levels. Moreover, as learners' proficiency increased, use of manner subordination also increased.

4.5.2.7 Hypotaxis and Enhancement Through Causal-Conditional Markers

The causal-conditional MBC contains the largest representative sample of MBUs in enhancement with a total of 29 measures thus having an influence on the overall frequency of hypotaxis. The EFCAMDAT2 displayed a significant increase from A1 (12.8) to B1 (45.5) which showed as learners gained more proficiency they used more subordination through the lower intermediate levels. There was a small decrease in subordination from B1 (45.5) to B2 (41.8). A sharp increase was seen from B2 (41.8) to C1 (79.2) thus indicating frequency of use peaked at the lower advanced level and did not subside, thus not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). Frequency sharply decreased from C1 (79.2) to C2 (17.5) which is thought to be due to the lack of the representative sample at the C2 level and as also was found in the Chen et al. (2021) study.

The frequency for causal-conditional subordinators in the CEDEL2/CAES was much higher from A1 (25.4) to B1 (54.2). There was a resulting decrease from B1 (54.2) to B2 (49.8) followed by a sharp increase at C1 (65.9) and a final small increase at C2 (68.7). The overall trend shows an increase from A1 through to C2. The frequency curve indicated that as frequency increased so did proficiency. Findings for the causal-conditional MBC in the CEDEL2/CAES are not consistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

There are two aspects that need to be highlighted with causal-conditional MBUs in hypotaxis and enhancement. First, in both EFCAMDAT2 and the CEDEL2/CAES there was a decrease in subordination from B1 to B2 of less than 5.0 per 5k. A possible explanation might be the sharp increase in use of manner MBUs and mild increase in spatial MBU use. Moreover, the high frequency of temporal subordination might have influenced results. Findings coincided with Rasool and Mahmood (2023) in that hypotaxis and enhancement through causal-conditional indicators evidenced a higher use frequency than temporal, manner and spatial MBCs.

The second aspect under discussion is when subordination at the C1 level in the EFCAMDAT2 (79.2) spiked to a peak over the CEDEL2/CAES (65.9). It is possible that the tasks' parameters allowed texts to be longer in the CEDEL2/CAES resulting in a higher frequency of adverbial subordination. This is not the case at the C1 level in the EFCAMDAT2 as learners produced an average of 189 words per task in a total of 340 texts. In contrast, the same level in the CEDEL2/CAES contained an average of 545 words per task in a total of 455 texts (Table 75). L2 Spanish writers should have used causal-conditional subordination at much higher frequency than L2 English writers based on corpus construct according to the increased number of words and texts produced. Data suggests that L1 Spanish writers reach a categorical peak at C1. This pattern holds true in manner and spatial subordination under hypotaxis but not the temporal MBC. Further research is necessary at the C2 level with a larger representative sample to clarify at what frequency and which type of MBUs learners employ in written texts, and to see if results would mirror those of the CEDEL2/CAES.

4.5.2.8 Parataxis and Extension Through Addition

Addition through parataxis and extension is the first and most extensive classification under parataxis with a total of five measures. The additive MBC frequency was categorically the highest out of every single MBC in the EFCAMDAT2 for parataxis. The data set showed an increase in coordination from A1 (64.9) to B1 (117.7) and a decrease from B1 (117.7) to C2 (85.1). These findings are not consistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) since coordination through addition peaks at the lower intermediate level.

The frequency increase in the EFCAMDAT2 from A1 (64.9) to B1 (117.7) is worth noting since it is consistent with studies carried out by Lahuerta Martínez (2018b) and Zarco-Tejada et al. (2016). The current study suggests that parataxis through additive extension at B1 might be a distinguishing proficiency level for L1 Spanish writing in L2 English that separates the beginning stage of acquisition from more intermediate and advanced stages. It is also worth pointing out that the main driver of high frequency in either previously mentioned study, as well as in the current study, is the coordinator *and*. Another aspect that stands out is

that findings from the EFCAMDAT2 are consistent with the Zarco-Tejada et al. (2016) in which a decrease in coordination from B1 to B2 is thought to be a possible sign that learners are replacing parataxis with other linguistic resources.

Findings from the CEDEL2/CAES in addition through parataxis and extension contained the highest frequency of MBU use in every category in the present study. While there was a decrease from A1 (138.1) to A2 (126.6), a sharp increase was seen from A2 (126.6) to B1 (143.0). The following pattern between B1 (143.0) and C2 (139.1) showed use of parataxis and extension remained relatively stable with less than a 3.9 word variation per 5k. Bearing this in mind, this study cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) with addition through paratactic extension in the CEDEL2/CAES.

Restrepo-Ramos (2021) studied the frequency of a larger variety of coordination measures for additive extension than the current study, yet the commonality encountered was an increase from A1 to B1 with a slight decrease from B1 to B2. If A2 is removed from the CEDEL2/CAES data, the frequency curve remains fairly level between the A1 (138.1), B1 (143.0), B2 (140.4), C1 (143.3) and C2 (139.1) in the CEDEL2/CAES. Findings are consistent with previous studies which found no significant differences between L2 proficiency levels (Lu, 2011; Ai & Lu, 2013) as well as no significant differences in the rate of coordination between intermediate and advanced L2 learners (Neary-Sundquist; 2016). When looking at stabilization of additive MBUs, the CEDEL2/CAES sets a pattern that would indicate coordination is a baseline tool used by L1 English speakers in L2 Spanish written texts.

Learner data extracted from the CEDEL2/CAES indicated learners used addition in parataxis and extension at a much higher rate than EFCAMDAT2 learners. Empirical findings from the EFCAMDAT2 and CEDEL2/CAES were consistent with previous studies (Rasool & Mahmood, 2023; Zarco-Tejada et al., 2016) which revealed that out of all the categories in paratactic extension, additive markers were used at the highest frequency.

4.5.2.9 Parataxis and Extension Through Variation

Results for variation under parataxis and extension exhibited interesting patterns in both sets of data incorporating two MBUs. From A1 (1.5) to A2 (2.3) there was a slight increase followed by a decrease from A2 (2.3) to B1 (1.4). This decrease may be due to a high use frequency in the addition MBC between these levels. Likewise, there is the possibility the increase from B1(1.4) to C2 (5.8) is the result of a general decrease in addition coordinators in said proficiency levels. The overall increase in frequency at advanced levels would not confirm DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in EFCAMDAT2 data with variation through paratactic extension.

EFCAMDAT2 texts demonstrated a much higher use frequency in variation than data from the CEDEL2/CAES. When trying to explain results for variation, Zarco-Tejada et al. (2016) found an overall decrease from A2 to B2 using three variation subordinators which included *instead*, *apart from that* and *or*. Data extracted from the EFCAMDAT2 directly showed the opposite tendency as the curve fell from A2 (2.3) to B1 (1.4) yet increased at B2 (3.3). The difference in measures as well as representative sample size may explain this divergence.

Unlike addition through parataxis and extension, results from variation in the CEDEL2/CAES occurred at a much lower frequency. There was a slight increase from A1 (0.0) to the peak at B1 (1.2.). B2 (0.8) and C1 (0.6) demonstrated a slight decrease followed by a slight increase at C2 (1.0). Since there were frequency peaks at B1 (1.2) and C2 (1.0), the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) with variation through paratactic extension in the CEDEL2/CAES cannot be confirmed. With that said, a larger representative sample is required to better gauge this MBC.

4.5.2.10 Parataxis and Extension Through Alternation

Results from the EFCAMDAT2 for alteration through parataxis and extension showed a decrease in frequency from A1 (8.6) to B2 (4.2). From B2 (4.2) to C2 (11.7) there was an

overall frequency increase in variation which is not consistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

The CEDEL2/CAES showed a significant drop in frequency from A1 (7.4) to A2 (2.5) for alternation through parataxis and extension. The decrease between these levels, although not as significant, mirrored the same findings denoted for addition through parataxis and extension. The use of other types of parataxis might explain this occurrence as there is an increase in temporal and manner enhancement as well as clarification and exemplification in elaboration from A1 to A2. There was an increase from A2 (2.5) to B2 (5.7) and a slight decrease to C1 (5.0) followed by an increase at C2 (6.6).

Data from alternation through parataxis extracted from the CEDEL2/CAES cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) with alternation through paratactic extension due to the irregular nature of frequency distribution. The question that needs to be answered is what circumstances lead to the drop in frequency from the A1 to A2 level and a rise in frequency at advanced levels. This study suggests further research is needed into alternative coordination in parataxis and extension as well as increasing number of measurements to obtain a better perspective.

When compared to the CEDEL2/CAES, the EFCAMDAT2 showed a much higher MBU frequency per 5k. However, the frequency sequence between proficiency levels in both data sets appears slightly erratic. Findings from Rasool and Mahmood (2023) concur with EFCAMDAT2 and CEDEL2/CAES data in that alternation is used more than variation, but less than extension.

4.5.2.11 Parataxis and Enhancement Through Temporal Markers

The temporal MBC is the first category in parataxis and enhancement with five measurement units. The trends in both data sets exhibited different patterns. Firstly, frequency in the EFCAMDAT2 saw a slight increase from A1 (1.4) to an overall peak at A2 (1.5). The fact that the category culminates at A2, as well as a higher frequency at beginner

levels, might be due to formulaic sequences in which the learner recalls a particular sequence of words and incorporates it into a communicative task designed to extrapolate said sequence (Alexopoulou et al., 2015; Chen et al., 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002).

There was a continuous decrease from A2 (1.5) in frequency of use until C2 (0.0). The increased use of temporal parataxis at A1 and A2 with the decrease through advanced levels confirms the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) with temporal indicators through paratactic extension. Findings are consistent with previous studies which found coordination to peak at beginner levels and subside at higher levels of proficiency (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012).

Data from the CEDEL2/CAES manifested vastly different results in that temporal indicators through parataxis and enhancement appeared at a relatively low frequency at A1 (0.4) with a gradual increase to C1 (1.8) and C2 (1.8). Temporal coordination set a pattern in which there were increases from beginning to advanced levels thus appearing more like patterns seen by the current study in hypotaxis. With that said, this study cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in the CEDEL2/CAES with paratactic temporal enhancement.

CEDEL2/CAES findings concur with previous research which found temporal coordination increased as proficiency increased from A2 to B1 (Giagkou et al., 2015; Zarco-Tejada et al., 2016). Spanish L2 writers used paratactic temporal coordinators at a higher frequency than L2 English writers. Results from Rasool and Mahmood (2023) and the EFCAMDAT2 both indicate that temporal indicators are the second most-used form of coordination in paratactic enhancement. In contrast, data from the CEDEL2/CAES show that temporal MBC is the most often used form of paratactic enhancement. Comparison between the three previously mentioned data sets indicate there are differences in L1 use of temporal coordination. While other studies (Wenhui Xuan, 2019; Yang et al., 2017; Zarco-Tejada et al., 2016) researched temporal coordination, their measurements were not adequate to cross-reference with the current study.

4.5.2.12 Parataxis and Enhancement Through Spatial Markers

Spatial enhancement through parataxis showed mixed results between either data set with a single measurement. The EFCAMDAT2 showed comparatively similar results to temporal enhancement. Frequency under spatial markers started with a relatively high frequency at A1 (4.3) which was due to the high frequency of use of *and there*. However, frequency plummeted at the A2 level (0.6). A slight increase at B1 (0.8) was observed with an overall decrease through to C2 (0.0). EFCAMDAT2 data confirms the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) with temporal indicators through paratactic enhancement.

Data extracted from the CEDEL2/CAES showed the only frequency occurrences happening at A1 (0.2), B1 (0.1) and C2 (0.2). With such a low frequency in every proficiency level, it cannot be confirmed that L1 English use *y allí* as a tool when writing in L2 Spanish at any level. Confirmation of the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) for spatial indicators through paratactic enhancement remains inconclusive for the CEDEL2/CAES. It is suggested that future research include more measurements to get an accurate representative sample.

Learners from the EFCAMDAT2 used the single paratactic spatial MBU more often than learners from the CEDEL2/CAES. The variance in frequency between A1 and A2 is thought to be the result of task-based formulaic sequencing (Alexopoulou et al., 2015; Alhassan & Wood, 2015; Chen et al., 2021; Lewis, 1997; Wray, 2002). Rasool and Mahmood (2023) incorporated the highest number of MBCs out of any known study and found zero use in parataxis through spatial indicators. While their findings are not consistent with the EFCAMDAT2, they tend to fall more in line with those from the CEDEL2/CAES as the overall frequency was both sporadic and low.

A possible insight to account for the difference in frequency may come from the specific grammatical functions and uses of *and there* versus *y allí*. In English *and there* is

formed by the coordinator *and* plus *there* operating as the existential subject (Quirk et al., 1985). In contrast, *y* is a coordinator in Spanish yet *allí* is considered as an adverb of place which in all certain terms wouldn't function as an existential subject (RAE, 2010). Within the pages of *Nueva Gramática de la Lengua Española* (2010) example (187) was found thus indicating the form *y allí* is a form used in Spanish, yet in terms of the example one could argue that it appears as an adverb of place.

- (187) ***Y allí** se desencadenó sobre nosotros un temporal que justifica el fracaso de los persas, no obstante su podrio temible.*
***And there** was unleashed upon us a storm that justifies the failure of the Persians, despite their fearsome power.*

Considering the fundamental differences between *and there* and *y allí*, this study proposes that due to having a function as an adverb of place, the Spanish form does not represent manner which would correlate as to why frequency was so low in the EFCAMDAT2. This finding indicated what can be termed as a negative Spanish L1 transfer to L2 English texts in that the coordinator was not used. The data extracted from the CEDEL2/CAES showed the opposite effect. The increase in use of *y allí* by English L1 writing in Spanish L2 does suggest L1 transfer to L2 as it is a commonly used form in the English language.

4.5.2.13 Parataxis and Enhancement Through Manner Markers

The results for the manner MBC with two measurements displayed different patterns in each data set. The EFCAMDAT2 displayed inconclusive results with a frequency of 0.1 taking place at the A1 and B2 level. It is observed that L1 Spanish learners of English do not use either MBU in this category as a linguistic resource. With such a low use frequency, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed or denied for the manner MBC in the EFCAMDAT2. It is recommended that future research include more measurements in paratactic enhancement through manner and increase the size of the data set to be able to clarify findings.

Manner through parataxis and enhancement in the CEDEL2/CAES showed an increase from A1 (0.2) to B1 (1.0) followed by a sharp decline from B1 (1.0) to B2 (0.5). From B2 (0.5) to C2 (1.5) there was a frequency increase in the parataxis and enhancement manner MBC. With the overall frequency increase peaking at the B1 and the C2 levels, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed for the manner MBC in the CEDEL2/CAES.

L1 English writers in L2 Spanish used manner coordination at a higher rate than L1 Spanish writing in L2 English. Findings from the CEDEL2/CAES are consistent with the Zarco-Tejada et al. (2016) study for manner through parataxis in that both studies found an increase in manner coordination from A2 to B2. Due to the low frequency from learners in the EFCAMDAT2, no obvious patterns can be taken from the information at hand and future research would benefit from more measurements in this MBC.

4.5.2.14 Parataxis and Enhancement Through Causal-Conditional Markers

The final category under parataxis and enhancement was the causal-conditional MBC containing four measurements with low use frequency in both data sets. Results from the EFCAMDAT2 showed a pattern more indicative of hypotaxis with an overall increase from A1 (0.0) to C1 (0.3) then a decrease at the C2 level (0.0). While the frequency was very low, the overall curve shows an increase from beginner levels to lower advanced levels thus not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Results from data extracted from CEDEL2/CAES for paratactic enhancement through causal-conditional markers showed zero frequency from A1 to B1 with frequency starting at B2 (0.2). A slight decrease appeared from B2 (0.2) to C1 (0.1) and with the same frequency continuing through to C2 (0.1). The curve set by Spanish L2 learners would suggest that this type of coordination increases from A2 to B2 which counters the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). Notwithstanding, a larger representative sample or an increase in measurements would better justify these findings.

Albeit at a low frequency, EFCAMDAT2 data showed that learners used the causal-conditional MBC with a higher frequency than CEDEL2/CAES learners. Rasool and Mahmood (2023) found causal-conditional coordination to be the most often linguistic resource used by learners in parataxis and enhancement. In contrast, this study found that causal-conditional coordination was used at a lower frequency than temporal and spatial coordination in the EFCAMDAT2. Moreover, causal-conditional indicators were the least-used resource in paratactic enhancement in data from the CEDEL2/CAES.

4.5.2.15 Parataxis and Elaboration Through Exposition

Exposition through parataxis is the first of three categories in elaboration incorporating three measurements. The EFCAMDAT2 showed a gradual increase from A1 (0.0) to B2 (0.8). From B2 (0.8) to C1 (0.4) there was a decrease followed by an increase from C1 (0.4) to C2 (0.8). Bearing in mind that coordination, although at low frequencies, increases through the intermediate levels and again increases from C1 to C2, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed for the exposition through elaboration in the EFCAMDAT2.

The CEDEL2/CAES showed a mix of results in paratactic elaborative exposition which started with a decrease from A1 (0.4) to A2 (0.3). Nevertheless, there was a sharp increase in frequency from A2 (0.3) to B1 (1.8) which might be a consequence of formulaic sequencing (Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002). From B1 (1.8) to B2 (0.5) there was a decrease which continued through C1 (0.6) with a final increase at C2 (1.1). Given the increase in elaborative coordination from A2 to B1 and the secondary increase from B2 to C2, it cannot be concluded the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) takes place in paratactic exposition through elaboration in the CEDEL2/CAES.

Exposition through paratactic elaboration was used at an overall higher frequency by learners in the CEDEL2/CAES than learners in data collected from the EFCAMDAT2.

Moreover, exposition was also used at a lower frequency than exemplification or clarification in both data sets which is consistent with Rasool and Mahmood (2023).

4.5.2.16 Parataxis and Elaboration Through Exemplification

Exemplification through parataxis and elaboration was less erratic than exposition in terms of the frequency of use distributed among the six proficiency levels using three measurements. The EFCAMDAT2 showed an increase from A1 (0.6) to B1 (3.2). There was a slight decrease from B1 (3.2) to B2 (2.4) which may correlate with an increase in use of exposition coordinators between those said levels. A notable phenomenon was the sharp increase in use between B2 (2.4) and C2 (6.7). Findings do not confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in exemplification through elaboration in the EFCAMDAT2. Findings suggest that as learners increase the frequency of exemplification through parataxis, they achieve a higher proficiency level.

The Zarco-Tejada et al. (2016) study incorporated three appositive measures incorporating the equivalent exemplification measure *for example*. Table 18 shows how Zarco-Tejada et al. (2016) found there was an increase from one occurrence at the B1 level to three occurrences at the B2 level. When compared to the current study, findings are not consistent with data from the EFCAMDAT2 as it showed the opposite trend appearing with a decrease from B1 to B2. If Zarco-Tejada et al. (2016) had included the advanced levels in their investigation, it would have been insightful to see if results matched those from the current study. Further research would clarify frequency patterns at upper intermediate and advanced levels.

While neither data set demonstrated an elevated use frequency, the CEDEL2/CAES showed an increase from A1 (0.2) to C2 (2.4). Results from this study demonstrated that L1 English writing in L2 Spanish used exemplification more often as their proficiency increased. Therefore, CEDEL2/CAES data for exemplification through parataxis and elaboration does not support the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). It is possible that exemplification increases with various types of subordination as phrasal level complexity is used more often at higher levels, thus necessitating examples as descriptions between clauses.

Further research with a larger representative sample and more measurements would be necessary to provide a better perspective into this matter.

Data showed L2 English writers clearly using paratactic exemplification more than L2 Spanish writers. Data from both the EFCAMDAT2 and CEDEL2/CAES concurred with the findings of the Rasool and Mahmood (2023) study in that paratactic elaboration through exemplification was used more than exposition, yet less than clarification.

4.5.2.17 Parataxis and Elaboration Through Clarification

The last category in parataxis and elaboration is composed of seven measurements representing clarification. Data from the EFCAMDAT2 showed an increase from A1 (1.1) to B2 (13.2) followed by a decrease to C1 (9.6) and a sizeable increase at C2 (16.7). Considering the pattern of an overall increase from A1 to C2, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed by EFCAMDAT2 data for parataxis and elaboration through clarification.

Finding from the CEDEL2/CAES showed continuous growth from A1 (0.2) to C2 (5.4) without any decreases. Data validates the proposal that English L1 learners of L2 Spanish increase their use of clarification through parataxis and elaboration as their proficiency increases. This study cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) for paratactic clarification in the representative sample taken from the CEDEL2/CAES.

The number of clarifying conjunctions increased from A2 to B2 in the study carried out by Zarco-Tejada et al. (2016) which is consistent with findings from both the EFCAMDAT2 as well as the CEDEL2/CAES. As with the current study, Rasool and Mahmood (2023) found learners to use clarification more than any other form of parataxis and elaboration.

4.5.3 Conclusions

In this section we investigate how L2 English and L2 Spanish writers used MBCs in written texts taken from the EFCAMDAT2 and the CEDEL2/CAES. Firstly, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) was tested across every MBC to address research question two. Secondly, research question three tracked MBC frequency trends used by learners in the six CEFR proficiency levels. Finally, research question four compared the differences in MBC frequency to explore L2 acquisition L1 to L2 transfer.

Findings from this study show that hypotaxis and extension is a phenomenon seen more frequently used by learners at the C1 and C2 proficiency level in the EFCAMDAT2 and the CEDEL2/CAES. While it couldn't be ruled out that grammatical metaphor takes place at these levels, no decrease in subordination was seen at advanced levels in addition, variation or alternation. Furthermore, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) could not be confirmed for any MBC in this grouping.

Hypotaxis and enhancement contained the highest number of MBUs in the entire study. Like extension, the trend in the EFCAMDAT2 as well as the CEDEL/CAES2 was that of increased use of hypotaxis and enhancement MBCs until advanced levels. Results showed that as learners increase their proficiency, they implement more adverbial subordination in their writing. One exception to the pattern was with the EFCAMDAT2 temporal MBC which showed high subordination use at A2, B1, B2 and C2. The second exception was with the CEDEL2/CAES with temporal subordination where the curve led to elevated frequency at B1, B2 and C1. This study suggests that temporal subordination is a steady linguistic resource used by L2 English and Spanish writers over a wide range of proficiency levels.

The DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) could not be confirmed for temporal, manner and causal-conditional conditional MBCs for hypotaxis and enhancement in either data set. The only MBC in this study to confirm said hypothesis was spatial enhancement in the CEDEL2/CAES which saw a clear tendency of increased

subordination at B1 and B2 followed by a clear decrease in frequency at beginner and advanced levels. In contrast, the previously mentioned hypothesis could not be confirmed in the spatial MBC in the EFCAMDAT2.

Research question three investigated cross-linguistic patterns in CEFR proficiency levels in the two data sets. There was a clear pattern of increased frequency from beginner to advanced levels in hypotaxis and enhancement in temporal, manner and causal-conditional MBCs in both data sets. The spatial MBC in the EFCAMDAT2 also demonstrated a pattern of increased use from beginner to advanced levels. Temporal subordination in the EFCAMDAT2 demonstrated a distinct pattern as subordination reached peak frequency at A2 which is thought to have occurred because of task-based formulaic sequencing with the elevated use of *when*. The resulting pattern of elevated frequency in A2, B1, B2 and C2 was thought to indicate a baseline stabilization in temporal subordination use, thus indicating learners use this MBC as a go to resource in their linguistic repertoire.

Parataxis and extension through addition achieved the highest frequencies seen in this study for any MBC. The trend for the EFCAMDAT2 saw a frequency increase through B1, and then small decreases through C2 thus resulting in a stable pattern with less than a 5.0 difference between the four proficiency levels. Peak frequencies in the CEDEL2/CAES occurred at B1 and C1. However, the overall trend, except for A2, was for a stable pattern of high additive coordination use with only a 5.0 deviation in frequency per 5k from A1 to C2. This suggests that additive coordination was a constant and steady linguistic resource used by L2 English and Spanish writers.

Addition, variation and alternation in parataxis through extension provided evidence in which the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) could not be confirmed in either the EFCAMDAT2 or the CEDEL2/CAES. Alternation in learner data from the EFCAMDAT2 and CEDEL2/CAES saw a peak frequency at A1 followed by decreases through intermediate levels. Advanced levels saw secondary increases, yet not to peak frequency at low beginner levels. While the said hypothesis could not be confirmed for this MBC due to increases at advanced levels, future research with a representative sample including more measures and a larger data set would help clarify findings.

As we move on to parataxis and enhancement, data from the EFCAMDAT2 confirms the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in temporal and spatial MBCs. In the case of the spatial MBC, the peak at A1 is a result of an increased use of *and* *there* which may be due to task-based formulaic sequencing. Results were inconclusive for manner in the EFCAMDAT2. Said hypothesis could not be confirmed in conditional-causal parataxis due to overall low frequencies in this category. Future research might consider a larger representative sample to measure this category and compare findings.

Parataxis through temporal coordination peaked at the C1 and C2 levels and not beginner levels in the CEDEL2/CAES thus providing results which are inconsistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). Although frequencies were low in manner and causal-conditional parataxis, said hypothesis was not confirmed in either MBC. Results for the spatial MBC were inconclusive. The low frequencies seen across proficiency levels in the causal-conditional data set and the lack of an overall curve would indicate that results would benefit from a larger representative sample with more measures to clear up this gray zone.

Parataxis through elaboration showed very clear results in that exposition in the EFCAMDAT2 and the CEDEL2/CAES saw frequency peaks at both intermediate and advanced levels. The pattern for exemplification and clarification showed that frequency peaked at the advanced level. The DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed in either data set in elaboration through exposition, exemplification and clarification. A compelling phenomenon observed in parataxis and clarification was that every single MBC in both data sets saw frequency rise from A1 to C2. A separate study including more fine-grained measures in elaborative MBCs would prove interesting to better study this trend.

While not every MBC saw outstanding frequency patterns, the most important trends are highlighted to answer research question three. Hypotaxis through extension in addition, variation and alternation for both the EFCAMDAT2 and the CEDEL2/CAES was generally

more often used by learners from the upper intermediate level to advanced levels, therefore indicating more subordination is used as proficiency increases.

An interesting trend in both the EFCAMDAT2 and the CEDEL2/CAES was seen in temporal enhancement in hypotaxis. Although there were slight peaks and dips in frequency, use was consistent in the EFCAMDAT2 at A2, B1, B2 and C2. A similar configuration was seen in the CEDEL2/CAES in proficiency levels B1, B2 and C1. Findings suggest that temporal subordination is a stable and important linguistic resource used by Spanish and English L2 learners.

Manner through hypotactic enhancement maintained a steady increase in frequency from beginner levels to advanced levels. Moreover, with frequency peaks at the advanced levels in both data sets, this study suggests that manner subordination increases with proficiency.

The causal-conditional MBC saw the highest frequencies in hypotaxis. Both data sets saw an overall increase from A1 to B1. The point to highlight is a decrease of less than 5 words per 5k was seen in both the EFCAMDAT2 and the CEDEL2/CAES from B1 to B2. The decrease might be due to an accumulation of rises in frequency in spatial and manner MBCs. While it might appear that subordination was leveling off at the upper intermediate level, there was a sharp increase in frequency from B2 to C1 in the EFCAMDAT2 and in the CEDEL2/CAES. Causal-conditional subordination appears to not only be used at a relatively high frequency at every level yet results indicate that it is also an indicator of proficiency with a peak at C1 in the EFCAMDAT2 and an apex at C2 in the CEDEL2/CAES thus separating intermediate from advanced proficiency levels.

The results for parataxis and extension through addition were compelling in the sense that both data sets aligned with previous studies (Restrepo-Ramos, 2021; Zarco-Tejada et al., 2016) which found an overall increase from A1 to B1 and a following decrease from B1 to B2. EFCAMDAT2 data showed that L2 English writers used additive coordination at a descending frequency rate from B2 to C2. Data from the CEDEL2/CAES painted an interesting picture in the sense that if the decrease in frequency at the A2 level is taken out, then the variation in frequency between A1 and C2 did not go over 5.2 words per 5k. Results from the CEDEL2/CAES are consistent with previous studies (Lu, 2011; Ai & Lu, 2013) as

no significant difference between proficiency levels was observed and there was not a noteworthy difference in the rate of coordination between intermediate and advanced L2 learners (Neary-Sundquist; 2016). However, what this study additionally introduces is in the CEDEL2/CAES there was not a frequency deviance between A1 and C2.

Parataxis and enhancement through temporal coordination was frequently used from A1 to B1 in data extracted from the EFCAMDAT2 which was the result of a high frequency in the use of the coordinator *and then* thus possibly indicating learners employed task-based formulaic sequencing (Alexopoulou et al., 2015; Alhassan & Wood, 2015; Chen et al., 2021; Lewis, 1997; Wray, 2002). L2 Spanish texts in the CEDEL2/CAES saw a spike in temporal coordination use from B2 to C2 which was primarily due to increased use in four of the five measurements in this MBC which shows these learners use a larger variety of coordination. Furthermore, findings from the CEDEL2/CAES concur with previous studies (Giagkou et al., 2015; Zarco-Tejada et al., 2016) as there was an increase in frequency from A2 to B1.

Spatial enhancement in parataxis saw an initial spike at the A1 level which was the result of increased frequency spurred by the coordinator *and there*. Task-based formulaic sequencing (Alexopoulou et al., 2015; Alhassan & Wood, 2015; Chen et al., 2021; Lewis, 1997; Wray, 2002) is thought to be the cause of this phenomenon as there is low use frequency in the remaining proficiency levels. Spatial coordination in the CEDEL2/CAES happened at such a low frequency that it was impossible to clearly distinguish proficiency levels. The findings for paratactic enhancement through manner and causal-conditional coordination showed learners using this type of parataxis at such a low frequency that it made it difficult to highlight a specific proficiency level. This study recommends that future research use more measures with a larger representative sample.

Parataxis and elaboration through exemplification and clarification showed one defining trend in both the EFCAMDAT2 as well as the CEDEL2/CAES. In said MBCs, the overall frequency curve started from its lowest frequency at A1 and reached a peak at advanced proficiency levels. Moreover, findings for exposition additionally exhibited this pattern. This study suggests that the advanced levels distinguish learners using elaborative coordination and is a marker for proficiency. While paratactic exposition through elaboration in the EFCAMDAT2 did not precisely follow the same trend, peak frequencies were observed in both intermediate and advanced levels yet a clear distinction between levels

could not be made as frequency fluctuated between 0.1 and 0.8. With that said, more studies on elaborative parataxis are needed throughout the spectrum of CEFR proficiency levels to clarify this matter.

The goal of research question four is to illuminate the ways in which there is L1 to L2 transfer by highlighting frequency trends in MBCs. EFCAMDAT2 results indicated higher frequency of hypotactic addition and variation in extension in texts. In contrast, CEDEL2/CAES data showed learners used hypotactic alteration at a much higher frequency than L2 English writers.

When looking at hypotaxis and enhancement, data revealed that L2 English writers used the temporal, spatial and manner enhancement at a higher frequency than L2 Spanish writers. Learners in both data sets used the causal-conditional MBC at the highest frequency out of all four MBCs in enhancement. Data from the CEDEL2/CAES demonstrated that Spanish L2 learners used causal-conditional subordination with an overall higher frequency than L2 English learners.

MBCs in parataxis through extension highlight several interesting points. First, addition proved to have the highest frequencies observed in this study with the main frequency drivers being *and* and *y*. L1 transfer to the respective L2 appeared to be consistent in both data sets. Secondly, data from the CEDEL2/CAES revealed that L1 English writing in L2 Spanish used additive coordination at a higher frequency than learner data extracted from the EFCAMDAT2. This could possibly be due to the higher word count permitted in the CEDEL2/CAES texts. While addition in the CEDEL2/CAES does not distinguish proficiency levels, this study suggests that paratactic addition is a baseline tool for L2 Spanish writers and that such high frequencies over the spectrum of proficiency levels is a sign of L1 to L2 transfer. Lastly, EFCAMDAT2 data showed that English L2 writers used paratactic variation and alternation at a higher frequency than Spanish L2 learners which might indicate L1 to L2 transfer.

EFCAMDAT2 learners used parataxis and enhancement in spatial and causal-conditional MBCs at a higher frequency than learner data in the CEDEL2/CAES. On the other hand, temporal and manner coordination was used at a higher frequency with L2 Spanish writers. In general, every MBC in parataxis and enhancement would benefit from a

larger representative sample with a higher number of measurements to clarify results from the present study.

Parataxis through elaboration is the final classification which is composed of the MBCs including exposition, exemplification and clarification. The only MBC in which CEDEL2/CAES data saw the highest frequency was in exposition. EFCAMDAT2 learners used exemplification and clarification at a substantially higher frequency thus suggesting L1 to L2 transfer in this group of learners.

4.6 Experiment 4: Level 4 Meaning-Based Units

Level 4 is the final phase which analyzes the patterns of individual MBUs to investigate the underlying trends which build the foundation of the current study. The basis for the fine-grained meaning-based approach is to understand which types of hypotaxis and parataxis are used in English and Spanish L2 written tasks and how this affects each level of analysis.

Within the parameters of Level 4, the analysis is broken down into meaning-based subcategories (MBSC) of MBCs which include indicators such as temporal same point, manner through means, etc. This information is included to give a more profound perspective of the different classifications of MBUs. While each subcategory could warrant its own analysis, there is enough overlap that repetition would be a factor, since there are MBCs categories in which there is one MBU. Each subcategory will undergo an analysis within the MBU framework, yet will not have a level of its own. Further research may expand MBUs in subcategories and use a larger representative sample to gain a better perspective.

Included within this final stage of analysis are all the subordinators and coordinators employed in this study. In the results section, MBU findings will be firstly listed in order of Levels 1, 2 and 3 followed by subcategories in Level 4. At the tail end of this sequence, each MBU will be detailed to show how they fit in the overall meaning-based framework. The idea behind how this information is organized is to allow for better interpretation of the data and a better analysis of the information in the discussion section. The findings for each MBU are detailed separately in the results section. To give a global perspective, MBUs are grouped

together in the MBF through Level 2 expansion categories under the corresponding taxis classification.

In the discussion for Level 4, frequency will be looked at to see how learners use MBUs and which MBUs are most used to identify L2 acquisition and L1 to L2 transfer patterns. MBUs are analyzed according to CEFR levels to see which type of subordination or coordination is most often used at a given proficiency. This will indicate how use patterns emerge as students acquire an L2. When applicable in terms of having reliable data from comparable studies, specific MBUs are detailed and analyzed to answer why a given pattern may or may not be happening.

Due to the large spectrum of frequency within the numerous MBUs, this study implemented a frequency classification system based on the relative density and frequency of MBUs. The frequency classification system is divided across four categories. The first category consists of MBUs which appear at zero (0.0) frequency. The second category is composed of high density low frequency (HDLF) MBUs in the use range of 0.1-0.9. The third category is low density medium frequency (LDMF) MBUs encompassing the range of 1.0-9.9. The final category is low density high frequency (LDHF) MBUs that have a range starting at 10.0 and continuing through the highest frequencies observed in this study. It should be noted the previously mentioned system is relative to the amount of MBUs and their relative frequency in this study as well as the size of the data set. Studies looking to replicate findings would have to devise a system based on the parameters of their representative sample and measurements. Furthermore, once the concept of the meaning-based framework is applied to other studies, the accumulation of knowledge may pave the way for a standardized system for measuring the frequency of meaning-based complexity.

4.6.1 Results

4.6.1.1 Hypotaxis and Extension

The first set of results for Level 4 begins with hypotaxis seen through the focal lens of extension. There are five MBSCs in extension which include positive additive, adversative addition, replacive variation, subtractive variation and alternation. MBUs are listed according to previously mentioned classifications.

Table 84 shows the first MBU for the positive additive MBSC from the EFCAMDAT2 is *whereas* which had a zero frequency at A1 and slightly increased at A2 (0.1). There was a decrease from A2 (0.1) to B1 (0.0) followed by a series of increases at B1 (0.0), B2 (0.3) and C1 (0.4). A decrease appeared from C1 (0.4) to C2 (0.0). As for *mientras que* from the CEDEL2/CAES, there was a lack of use of this MBU at A1, A2 and B1. The sequence changed with a 0.1 frequency at the B2, C1 and C2 levels.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2			<i>whereas</i>	0.0	0.1	0.0	0.3	0.4	0.0
CEDEL2/ CAES			<i>mientras que</i>	0.0	0.0	0.0	0.1	0.1	0.1
EFCAMDAT2	Hypotaxis Extension Addition	Positive Additive	<i>besides</i>	0.1	0.1	0.0	0.1	0.1	0.0
CEDEL2/ CAES			<i>además de que</i>	0.0	0.0	0.0	0.1	0.3	0.5
EFCAMDAT2			<i>apart from</i>	0.0	0.0	0.0	0.0	0.0	0.8
CEDEL2/ CAES			<i>aparte de que</i>	0.0	0.0	0.0	0.0	0.0	0.1

Table 84 Level 4 Results Hypotaxis Extension and Addition through Positive Additive.

The positive additive MBU *besides* had a null frequency use at B1 and C2. At A1, A2, B2 and C1, EFCAMDAT2 learners used this form at a frequency of 0.1 per 5k. The results for *además de que* demonstrated a different pattern. At A1, A2 and B1 there was zero use of the MBU. The trend changed with an increase at B2 (0.1) followed by two more increases at C1 (0.3) and C2 (0.5).

The final MBUs under the positive addition MBSC are *apart from* and the equivalent *aparte de que*. MBUs from both data sets show similar tendencies. There was zero use of *apart from* or *aparte de que* in A1 to C1. The pattern changed with low frequency at the C2 (0.8) with *apart from* as well as *aparte de que* (0.1).

The second MBSC under hypotaxis extension and addition (Table 85) is adversative addition which contains a single MBU. *Without* was analyzed from the EFCAMDAT2 and saw an increases from A1 (0.1) to a peak at C2 (1.0) and a final decrease at the C2 (0.0). The equivalent MBU used in the CEDEL2/CAES was *sin que* which saw a decrease from A1 (0.2) to A2 (0.1). From A2 (0.1) there was a slight increase to B1 (0.3) and another increase to B2 (0.4) with no change at C1 (0.4). There was a decrease at C2 (0.2) which is the same as the initial starting point at A1 (0.2).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis	Adversative Addition	<i>without</i>	0.1	0.1	0.2	0.5	1.0	0.0
CEDEL2/ CAES	Extension Addition		<i>sin que</i>	0.2	0.1	0.3	0.4	0.4	0.2

Table 85 Level 4 Results Hypotaxis Extension and Addition through Adversative Addition.

The third MBSC is replacive variation (Table 86) which contained one MBU. Data extracted from the EFCAMDAT2 saw no use of *instead of* in the A1 and A2 levels. A gradual increase was seen from B1 (0.1) to B2 (0.3) in the intermediate levels. The trend continued at advanced levels with an increase at C1 (0.4) and a final increase at C2 (0.8). *En vez de* was used in the CEDEL2/CAES and showed a similar trend. In levels A1, A2, and B1 there was not an instance of use by learners. B2 (0.1) saw the first occurrence with a gradually rising trend to C1 (0.3) and a peak at C2 (0.5).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis	Replacive	<i>instead of</i>	0.0	0.0	0.1	0.3	0.4	0.8
CEDEL2/ CAES	Extension Variation		<i>en vez de</i>	0.0	0.0	0.0	0.1	0.3	0.5

Table 86 Level 4 results Hypotaxis Extension and Addition through Replacive Variation.

Subtractive variation (Table 87) is the fourth MBSC in hypotaxis and extension through addition and includes one measurement. *Except that* from the EFCAMDAT2 and *salvo que* from the CEDEL2/CAES did not exhibit any use from A1 to C2.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis	Subtractive	<i>except that</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES	Extension Variation		<i>salvo</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 87 Level 4 results Hypotaxis Extension and Addition through Subtractive Variation.

Alternation under hypotaxis through extension (Table 88) is an MBC, not a MBSC. While the results are the same as Level 3 alternation, to paint a complete picture, they will be repeated under Level 4 for the MBU. There was one measure in this category encompassing *if not ... then* from the EFCAMDAT2. Overall frequency was low with zero frequency at A1 then there was a slight uptick at A2 (0.1) and B1 (0.1). Another slight increase was observed from B1 (0.1) to B2 (0.2). At C1 (0.2) the frequency leveled out with a decrease to zero use at C2. Data extracted from CEDEL2/CAES incorporating *si no* showed no use at A1 with an

increase at A2 (0.4), B1 (1.0) and B2 (1.2). C1 (1.0) saw a small decrease while there was an overall peak at C2 (1.3).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis		<i>if not ... then</i>	0.0	0.1	0.1	0.2	0.2	0.0
CEDEL2/ CAES	Extension Alternation	n/a	<i>si no</i>	0.0	0.4	1.0	1.2	1.0	1.3

Table 88 Level 4 results Hypotaxis Extension and Alternation.

4.6.1.2 Hypotaxis and Enhancement

Hypotaxis and enhancement is the most sizeable category containing the largest representative sample in this study which incorporates 12 MBSCs. The first hypotaxis MBSC is through same time temporal enhancement and incorporates seven measurements (Table 89). The first MBU for the EFCAMDAT2 was *as* and it showed an increase from A1 (0.0) to A2 (0.1) and remained the same at B1 (0.1). A secondary increase appeared at B2 (0.2) followed by a leveling off at C1 (0.2) and a final increase at C2 (0.8). Results from the CEDEL2/CAES saw a zero use of *a medida que* in levels A1 to C2.

The second MBU for same time temporal enhancement was *while*. EFCAMDAT2 saw frequency increases at A1 (0.1), A2 (1.0) and B1 (1.8). There was a decrease at B2 (1.4) then there was a sharp increase at C1 (2.6) followed by a slight decrease at C2 (2.5). Results for *mientras* from the CEDEL2/CAES saw increases at A1 (1.1), A2 (1.3) and B1 (1.6). Like the EFCAMDAT2, there was a decrease at B2 (1.5). Frequency increased at C1 (2.4) and C2 (3.0).

The most significant sample with the highest frequency in same time temporal MBUs was with *when* and *cuando*. With EFCAMDAT2 learners, there was a low frequency at A1 (2.9). A sharp increase appeared at A2 (21.9) followed by a decrease at B1 (14.4). There was a relatively small increase at B2 (14.7) followed by a decrease at C1 (9.1) and a slight increase at C2 (9.2). Data from the CEDEL2/CAES with *cuando* saw a sharp increase from A1 (5.9) to A2 (14.0). There was another increase at B1 (19.7) yet a decrease in frequency at B2 (18.8). C1 (19.1) experienced a slight increase followed by another decrease at C2 (14.6).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis Enhancement Temporal	Same time	<i>as</i>	0.0	0.1	0.1	0.2	0.2	0.8
CEDEL2/ CAES			<i>a medida que</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>while</i>	0.1	1.0	1.8	1.4	2.6	2.5
CEDEL2/ CAES			<i>mientras</i>	1.1	1.3	1.6	1.5	2.4	3.0
EFCAMDAT2			<i>when</i>	2.9	21.9	14.4	14.7	9.1	9.2
CEDEL2/ CAES			<i>cuando</i>	5.9	14.0	19.7	18.8	19.1	14.6
EFCAMDAT2			<i>as soon as</i>	0.0	0.2	0.7	0.5	0.1	0.8
CEDEL2/ CAES			<i>tan pronto como</i>	0.2	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>the moment en el</i>	0.1	0.0	0.2	0.2	0.1	0.0
CEDEL2/ CAES			<i>momento en que</i>	0.0	0.0	0.0	0.0	0.1	0.0
EFCAMDAT2			<i>upon</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>al</i>	0.0	0.0	0.0	0.0	0.2	0.3
EFCAMDAT2			<i>every time</i>	0.0	0.0	0.0	0.0	0.1	0.0
CEDEL2/ CAES			<i>cada vez de</i>	0.0	0.0	0.1	0.0	0.1	0.0

Table 89 Level 4 results for Hypotaxis through Same Time Enhancement.

The fourth MBU in the temporal same time category for the EFCAMDAT2 is *as soon as* which had zero use at A1. At A2 (0.2) there was an increase that was followed by a second increase at B1 (0.7). There were decreases at B2 (0.5) and C1 (0.1). The peak in frequency appeared at C2 (0.8). Results from the CEDEL2/CAES using *tan pronto como* showed a low frequency A1 (0.2) and then null use in A2 through C2.

The fifth MBU for same time temporal enhancement in the EFCAMDAT2 utilized *the moment* which started with a low frequency at A1 (0.1) and then a decrease at A2 (0.0). There was an increase at B1 (0.2) with a plateau at B2 (0.2). A decrease started with C1 (0.1) finishing with a lack of use at C2 (0.0). Data from the CEDEL2/CAES using *en el momento en que* showed zero frequency in A1, A2, B1, B2 and C2. The only instance of frequency appeared at C1 (0.1).

The sixth MBU was *upon* for the EFCAMDAT2 and showed zero frequency in all CEFR proficiency levels. Results for *al* from the CEDEL2/CAES revealed zero use from A1 to B2. There was a slight increase at C1 (0.2) and an increase at C2 (0.3).

The final MBU for the temporal same time was *every time* for the EFCAMDAT2 with the only occurrence at the C1 level (0.1). *Cada vez de* was the MBU for the CEDEL2/CAES and displayed zero use in A1, A2, B2 and C2. Low frequency was observed at B1 (0.1) and C1 (0.1).

Hypotaxis and enhancement through different time temporal indicators (Table 90) included five separate measures. The first MBU for this classification incorporated *after* for the EFCAMDAT2. From A1 (1.0) to A2 (5.2) there was a sharp increase followed by a sharp decrease at B1 (1.8). Another increase occurred at B2 (3.0) and followed by yet another decrease at C1 (1.6). Finally, C2 (5.0) saw a sharp increase with the second highest frequency seen with the MBU. *Después de que* was used for data extraction in the CEDEL2/CAES and registered a continuous increase from A1 (0.4) through A2 (1.5) and B1 (1.5) reaching a peak frequency at B2 (1.7). From B2 (1.7), there was a decrease at C1 (1.3) followed by a further decrease at C2 (0.7).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis Enhancement Temporal	Different time	<i>after</i>	1.0	5.2	1.8	3.0	1.6	5.0
CEDEL2/ CAES			<i>después de que</i>	0.4	1.5	1.5	1.7	1.3	0.7
EFCAMDAT2			<i>since</i>	0.2	0.2	0.7	0.7	0.3	1.7
CEDEL2/ CAES			<i>desde que</i>	0.4	0.1	0.2	0.2	0.2	0.0
EFCAMDAT2			<i>once</i>	0.7	0.0	0.2	0.8	0.5	1.7
CEDEL2/ CAES			<i>una vez que</i>	0.0	0.1	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>before</i>	1.0	0.8	2.1	1.5	1.2	1.7
CEDEL2/ CAES			<i>antes de que</i>	0.2	0.6	1.3	1.4	1.1	0.3
EFCAMDAT2			<i>until</i>	0.2	0.3	0.9	0.9	0.2	2.5
CEDEL2/ CAES			<i>hasta que</i>	1.1	0.1	0.4	0.3	0.3	0.1

Table 90 Level 4 results for Hypotaxis through Different Time Enhancement.

The second MBU in the temporal same time class incorporated *since* from the EFCAMDAT2 which saw the same use frequency at A1 (0.2) and A2 (0.2) with an increase at the intermediate level with B1 (0.7) and B2 (0.7) showing the same frequency. A decrease in frequency happened at C1 (0.3) with a sharp increase appearing at the C2 level (1.7). Data from the CEDEL2/CAES for *desde que* showed, albeit at a low frequency, the highest use case at A1 (0.4). A2 (0.1) saw a decrease followed by a slight increase to a plateau at B1 (0.2), B2 (0.2) and C1 (0.2) with all showing the same frequency. At the C2 (0.0) there was

no use of this MBU.

The third EFCAMDAT2 MBU in different time temporal indicators is the subordinator *once* which started at a low frequency at A1 (0.7) and decreased to zero use at A2 (0.0). There was an increase in intermediate levels at B1 (0.2) and B2 (0.8) followed by a decrease at C1 (0.5). The highest frequency for *once* occurred at C2 (1.7). Results using *una vez que* in the CEDEL2/CAES showed use at A2 (0.1) and zero use in the five remaining proficiency levels.

The fourth MBU encompassed *before* in the EFCAMDAT2 and began with a higher frequency at A1 (1.0) with a decrease at A2 (0.8). There was a sharp increase to the peak frequency at B1 (2.1). A decrease to B2 (1.5) was seen followed by a second one at C1 (1.2). C2 (1.7) saw an increase, yet not as high as the peak at B1 (2.1). Findings from the CEDEL2/CAES for *antes de que* showed increases at A1 (0.2), A2 (0.6), B1 (1.3) and B2 (1.4). The pattern decreased at C1 (1.1) and then again at C2 (0.3).

The fifth and final MBU for different time temporal subordinators was *until* in the EFCAMDAT2. Results showed increases at A1 (0.2), A2 (0.3) and B1 (0.9). At B1 (0.9) and B2 (0.9) the frequency remained the same, yet it decreased at C1 (0.2). In contrast to the previously mentioned decrease, the C2 (2.5) level demonstrated the overall highest frequency of *until*. The opposite trend appeared in the findings for *hasta que* in the CEDEL2/CAES with peak frequency appearing at A1 (1.1) with a downward trend to A2 (0.1). There was a slight increase at B1 (0.4) followed a decrease at B2 (0.3) with C1 (0.3) having the same frequency. C2 (0.1) saw a final decrease.

Table 91 shows same place spatial enhancement in hypotaxis is a MBSC with only one measurement. Despite being covered in Level 3 and for the sake of providing the full spectrum of MBUs, it was decided to include data in the results. *Where* showed an overall increasing trend from lower beginner to lower advanced. From A1 (1.5) frequency increased at A2 (2.2), B1 (3.0), B2 (3.3) and C1 (4.2). Contrary to the other five levels, there was a sharp decrease at C2 (0.8). The findings using *donde* from the CEDEL2/CAES showed a similar pattern with an increase from A1 (2.0) to A2 (4.2) and then additional increases at B1 (5.3) and B2 (6.1). The sequence changed with a decrease from B2 (6.1) to C1 (3.8) and a second decrease at C2 (1.9).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis		<i>where</i>	1.5	2.2	3.0	3.3	4.2	0.8
CEDEL2/ CAES	Enhancement Spatial	Same place	<i>donde</i>	2.0	4.2	5.3	6.1	3.8	1.9

Table 91 Level 4 results for Hypotaxis through Same Place Spatial Enhancement.

The manner and means MBSC (Table 92) through hypotaxis and enhancement employed one measure for both data sets. *By* was used to extract data in this classification in the EFCAMDAT2 and showed zero frequency at A1 and A2. Frequency climbed at B1 (0.1), B2 (0.9) and C1 (2.1) with a final decline at C2 (1.7). Data extracted from the CEDEL2/CAES used *por* which displayed zero use at A1 followed by an increase at A2 (0.2) and a leveling off at B1 (0.2). A second frequency increase came about at B2 (0.3) with a decrease at C1 (0.1) and a third increase at C2 (0.2).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis		<i>by</i>	0.0	0.0	0.1	0.9	2.1	1.7
CEDEL2/ CAES	Enhancement Manner	Means	<i>por</i>	0.0	0.2	0.2	0.3	0.1	0.2

Table 92 Level 4 results for Hypotaxis through Manner and Means Enhancement.

Findings for manner and comparison enhancement in hypotaxis (Table 93) incorporated two separate measures. Results for *as if* in the EFCAMDAT2 exhibited no use at A1 and A2. Use started at B1 (0.1) and increased at B2 (0.3) and C1 (0.6) with a final decrease at C2 (0.0). *Como si* was used in the CEDEL2/CAES and showed null use in A1, A2 and B1. However, increases were seen at B2 (0.1), C1 (0.4) and C2 (0.8).

The second measure for the manner and comparison MBSC used *as* with the EFCAMDAT2. Results indicated a steady increase from A1 (0.2) to A2 (1.0), B1 (1.9), B2 (2.8) and C1 (4.4) with a decrease at the C2 (3.3). CEDEL2/CAES results using *como* showed the same gradual increase in frequency starting at A1 (0.2) and continuing to A2 (0.9). At B1 (0.9) the use frequency leveled off. There were sharp frequency increases at B2 (3.0), C1 (4.1) and C2 (5.9).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis Enhancement Manner	Comparison	<i>as if</i>	0.0	0.0	0.1	0.3	0.6	0.0
CEDEL2/ CAES			<i>como si</i>	0.0	0.0	0.0	0.1	0.4	0.8
EFCAMDAT2			<i>as</i>	0.2	1.0	1.9	2.8	4.4	3.3
CEDEL2/ CAES			<i>como</i>	0.2	0.9	0.9	3.0	4.1	5.9

Table 93 Level 4 results for Hypotaxis through Manner and Comparison Enhancement.

Cause and reason through causal-conditional enhancement (Table 94) is one of the largest MBSC in this study with a total of 13 measures. The first measure from the EFCAMDAT2 utilizes *because* which starts with a relatively high frequency at A1 (8.5) and continues increasing at A2 (18.4) and peaked at B1 (18.8). There was a decrease at B2 (14.4) followed by an increase at C1 (14.6) then a drop in frequency to zero at C2. *Porque* was used for CEDEL2/CAES data extraction and showed an increase from A1 (13.0) to A2 (21.7) with a peak at B1 (25.4). A decreasing frequency trend was seen for the remainder of proficiency levels starting at B2 (24.8) then continuing through to C1 (19.3) and ending at C2 (16.5).

The second indicator for the cause and reason MBSC in hypotaxis is *as* in the EFCAMDAT2 which started at a low frequency at A1 (0.1). The trend showed increases at A2 (0.6), B1 (1.1), B2 (1.4) and apexed at C1 (2.4). There did not appear any use of *as* at C2 (0.0). Results for *ya que* in the CEDEL2/CAES did not show any use at A1, A2 or B1. B2 (0.1) was the first instance of use with an increase at C1 (0.6) and a peak at C2 (1.8).

In case was the fourth MBU in the cause reason MBSC for the EFCAMDAT2 and had zero use in A1, A2 and C2. The first use was recorded at B1 (0.2) with an increase at B2 (0.4) and C1 (0.5). The MBU used for the CEDEL2/CAES *en caso de que* showed zero use at any proficiency level.

The fifth MBU was *seeing that* in the EFCAMDAT2 and *visto que* in the CEDEL2/CAES in the cause and reason MBSC. The EFCAMDAT2 MBU had zero use. The only recorded occurrence for *visto que* appeared at C2 (0.1) in the CEDEL2/CAES.

Given that was the sixth MBU for cause and reason in the EFCAMDAT2 which showed the only use at C1 (0.2). *Dado que* in the CEDEL2/CAES appeared with only two instances of use at C1 (0.1) and C2 (0.2).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2			<i>because</i>	8.5	18.4	18.8	14.4	14.6	0.0
CEDEL2/ CAES			<i>porque</i>	13.0	21.7	25.4	24.8	19.3	16.5
EFCAMDAT2			<i>as</i>	0.1	0.6	1.1	1.4	2.4	0.0
CEDEL2/ CAES			<i>ya que</i>	0.0	0.0	0.0	0.1	0.6	1.8
EFCAMDAT2			<i>since</i>	0.1	0.1	0.3	0.4	0.3	0.0
CEDEL2/ CAES			<i>puesto que</i>	0.0	0.0	0.0	0.0	0.1	0.4
EFCAMDAT2			<i>in case</i>	0.0	0.0	0.2	0.4	0.5	0.0
CEDEL2/ CAES			<i>en caso de que</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>seeing that</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>visto que</i>	0.0	0.0	0.0	0.0	0.0	0.1
EFCAMDAT2			<i>given that</i>	0.0	0.0	0.0	0.0	0.2	0.0
CEDEL2/ CAES			<i>dado que</i>	0.0	0.0	0.0	0.0	0.1	0.2
EFCAMDAT2			<i>considering</i>	0.0	0.0	0.1	0.3	0.6	0.0
CEDEL2/ CAES			<i>considerando</i>	0.0	0.0	0.0	0.0	0.1	0.0
EFCAMDAT2			<i>with</i>	0.0	0.0	0.0	0.1	0.0	0.0
CEDEL2/ CAES			<i>con que</i>	0.0	0.1	0.0	0.0	0.2	0.2
EFCAMDAT2			<i>as a result of</i>	0.0	0.1	0.1	0.0	0.0	0.0
CEDEL2/ CAES			<i>como</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>consecuencia</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>de</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>because of</i>	0.0	0.0	1.9	0.0	0.0	0.0
CEDEL2/ CAES			<i>a causa de</i>	0.4	0.0	0.0	0.2	0.1	0.4
EFCAMDAT2			<i>in case of</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>en caso de</i> ³¹	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>due to</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>debido a que</i>	0.0	0.0	0.0	0.0	0.0	0.1
EFCAMDAT2			<i>thanks to</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>gracias a que</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 94 Level 4 results for Hypotaxis through Cause and Reason Causal-conditional Enhancement.

The third MBU in the cause and reason classification for the EFCAMDAT2 is *since* which had a low frequency with for A1 (0.1) and A2 (0.1). A slight rise in frequency occurred at B1 (0.3) and B2 (0.4). C1 (0.3) had a decrease and then zero use at C2. Findings for *puesto que* in the CEDEL2/CAES showed zero use at A1 to B2. C1 (0.1) was the first occurrence of low frequency followed by an increase at C2 (0.4).

³¹ *En caso de* was used, alongside *in case of*, to extract infinitive forms and should not be confused with *in case* and *en caso de que* which was used to extract finite and subjunctive forms.

The seventh MBU in the cause and reason MBSC included *considering* for the EFCAMDAT2 with zero use in A1, A2 and C2. However, an upward trend was seen starting at B1 (0.1), B2 (0.3) and ending at C1 (0.6). Results using *considerando que* were extracted from data in the CEDEL2/CAES with the only occurrence taking place at C1 (0.1).

The eighth MBU for the cause and reason classification included *with* in the EFCAMDAT2 with only one occurrence at B2 (0.1). *Con que* was the MBU used in the CEDEL2/CAES and showed zero frequency at A1, B1 and B2. A2 (0.1) saw a low use frequency with a minimal increase at C1 (0.2) and C2 (0.2).

The ninth cause and reason MBU for EFCAMDAT2 was *as a result of*. Use of this MBU was seen at A2 and B1 with a frequency of 0.1. The remaining proficiency levels saw zero use of said MBU. *Como consecuencia de* was used for data extraction in the CEDEL2/CAES and saw no use in all six CEFR proficiency levels.

The tenth MBU in cause and reason is *because of* for the EFCAMDAT2 data set. There was only one occurrence at the B1 (1.9) level. *A causa de* in the CEDEL2/CAES saw zero use at A1 through B1. Learners of L2 Spanish first used this MBU at B2 (0.2) with a slight decrease at C1 (0.1) followed by a final increase at C2 (0.4).

The eleventh MBU in the cause and reason MBSC used *in case of* and *en caso de*. There were no instances of use in either data set by both L2 English and L2 Spanish learners.

The twelfth MBU included *due to* for the EFCAMDAT2 and did not see any use in all six proficiency levels. The only recorded use for *debido a que* in the CEDEL2/CAES occurred at C2 (0.1).

The thirteenth and final MBU for the cause and reason MBSC included *thanks to* for the EFCAMDAT2 and *gracias a que* for the CEDEL2/CAES. There was zero use across the board in either data set.

The MBSC of purpose through causal-conditional enhancement in hypotaxis contained five measures (Table 95). *So that* is the first MBU used in the EFCAMDAT2. A1

saw zero use while a steady increase started at A2 (0.5) and continued through B1 (0.6). There was a decline in use at B2 (0.1) followed by an upswing at C1 (0.5) and finally a return to zero at C2. The MBU *para que* started at A1 (0.4) then decreased at A2 (0.1). There were increases in the remaining levels B1 (0.5), B2 (0.9), C1 (2.2) and C2 (2.8).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis Enhancement Causal- conditional	Purpose	<i>so that</i>	0.0	0.5	0.6	0.1	0.5	0.0
CEDEL2/ CAES			<i>para que</i>	0.4	0.1	0.5	0.9	2.2	2.8
EFCAMDAT2			<i>to</i> (in order to)	0.3	1.0	1.6	2.9	5.1	7.5
CEDEL2/ CAES			<i>para</i>	6.5	11.8	18.4	8.3	28.3	28.3
EFCAMDAT2			<i>with the aim of</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>con el objetivo de</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>in exchange for</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>a cambio de que</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>with the intention of</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>con la intención de que</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 95 Level 4 results for Hypotaxis through Purpose Causal-conditional Enhancement.

The second measurement for purpose indicators in causal-conditional enhancement was *to* which makes the implication of <<in order to>>. EFCAMDAT2 findings demonstrated an increase in use starting at A1 (0.3) and continuing through A2 (1.0), B1 (1.6), B2 (2.9), C1 (5.1) and peaking at C2 (7.5). Results using the MBU *para* in the CEDEL2/CAES displayed an overall high frequency at A1 (6.5) which continued through A2 (11.8) and B1 (18.4). There was a decrease at the B2 (8.3) level with a rebound at C1 (28.3) and a leveling of frequency at C2 (28.3).

The third, fourth and fifth MBUs for the purpose MBSC constituting *with the aim of*, *in exchange for* and *with the intention of* from the EFCAMDAT2 showed zero frequency. Data from the CEDEL2/CAES showed zero frequency for the final three MBUs including *con el objetivo de*, *a cambio que* and *con la intención de*.

The positive condition is a MBSC in causal-conditional enhancement through hypotaxis is composed of five measures (Table 96). The first measure for the EFCAMDAT2

is *if* with frequency increasing from A1 (3.1) to A2 (6.0) and another exceptionally large increase at B1 (17.6). B2 (15.2) saw a decrease in frequency yet there was another exceptional peak at C1 (45.0). A significant decrease appeared at C2 (6.7).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2			<i>if</i>	3.1	6.0	17.6	15.2	45.0	6.7
CEDEL2/ CAES			<i>si</i>	3.9	5.4	8.3	11.9	9.5	11.8
EFCAMDAT2			<i>provided that</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>siempre que</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2	Hypotaxis	Positive Condition	<i>as long as</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES	Enhancement		<i>siempre y cuando</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2	Causal- conditional		<i>so long as</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>con tal de que</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>on the condition that</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>a condición de que</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 96 Level 4 results for Hypotaxis through Positive Condition Causal-conditional Enhancement.

Si in the CEDEL2/CAES was the equivalent MBU to *if* in causal-conditional enhancement through hypotaxis in positive condition. There were frequency increases starting at A1 (3.9) and continuing through A2 (5.4), B1 (8.3), B2 (11.9). At the C1 (9.5) level there was a decrease followed by secondary increase at C2 (11.8) to reach just shy of peak frequency for this category.

The remaining four MBUs for EFCAMDAT2 including *provided that*, *as long as*, *so long as* and *on the condition that* did not show any use. Likewise, the final four MBUs composing of *siempre que*, *siempre y cuando*, *con tal de que* and *a condición de que* failed to see any frequency occurrences.

The negative condition MBSC under causal-conditional hypotaxis is comprised of two MBUs (Table 97). *Unless* was used for data extraction in the EFCAMDAT2 and demonstrated zero results in A1, A2 and C2. The first occurrence of use was at B1 (0.4) with a decrease to B2 (0.1) and a final increase at C1 (0.3). Findings for *a menos que* in the

CEDEL2/CAES saw two frequency occurrences at A1 (0.2) and C2 (0.1). Proficiency levels A2, B1, B2 and C2 showed zero use.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis Enhancement Causal- conditional	Negative Condition	<i>unless</i>	0.0	0.0	0.4	0.1	0.3	0.0
CEDEL2/ CAES			<i>a menos que</i>	0.2	0.0	0.0	0.0	0.0	0.1
EFCAMDAT2			<i>without</i>	0.0	0.2	0.0	0.7	0.0	0.0
CEDEL2/ CAES			<i>sin</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 97 Level 4 results for Hypotaxis through Negative Condition Causal-conditional Enhancement.

The second MBU in negative condition is *without* in the EFCAMDAT2 which showed results at A2 (0.2) and B2 (0.7). The remaining four levels had zero use. The results from the CEDEL2/CAES used *sin* and did not produce any examples of use by L2 Spanish learners.

Concessive condition is the last MBSC in causal-conditional enhancement in hypotaxis and contains four measures (Table 98). The first MBU used for the EFCAMDAT2 encompassed *even if* while the CEDEL2/CAES factored in the use of *aun si*. Zero use was observed across all six proficiency levels in both data sets.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis Enhancement Causal- conditional	Concessive Condition	<i>even if</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>aun si</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>even though</i>	0.0	0.0	0.6	0.4	1.2	0.0
CEDEL2/ CAES			<i>a pesar de que</i>	0.0	0.1	0.0	0.0	0.2	0.0
EFCAMDAT2			<i>although</i>	0.6	0.5	2.1	5.3	8.5	3.3
CEDEL2/ CAES			<i>aunque</i>	0.9	0.5	1.6	3.5	5.1	5.8
EFCAMDAT2			<i>despite</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>pese a que</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 98 Level 4 results for Hypotaxis through Concessive Condition Causal-conditional Enhancement.

The second set of MBUs for the concessive condition MBSC saw slightly more activity with *even though* in the EFCAMDAT2 showing zero use at the A1, A2 and C2 levels with frequency starting at B1 (0.6). A small decrease was seen at B2 (0.4) followed by an

increase at C1 (1.2). The results from the CEDEL2/CAES for *a pesar de que* showed use at A2 (0.1) and C1 (0.2) with the remaining four proficiency levels registering zero use.

The third set of MBUs for the concessive condition MBSC showed the highest frequency for the classification. Findings for *although* from the EFCAMDAT2 started with a low frequency at A1 (0.6) and A2 (0.5). There was an uptick in use at B1 (2.1) followed by relatively significant increases at B2 (5.3) and C1 (8.5). The C2 (3.3) level displayed a sharp drop in frequency. Results from the CEDEL2/CAES using *aunque* started off with a slightly higher frequency at A1 (0.9) with a decrease at A2 (0.5). There was an increase starting at B1 (1.6) and continuing through B2 (3.5), C1 (5.1) and C2 (5.8).

Findings for the final MBU in concessive condition encompassed *despite* in the EFCAMDAT2 and *pese a que* for the CEDEL2/CAES. In all six of the CEFR proficiency levels, zero use occurred in L2 learners of both Spanish and English.

4.6.1.3 Hypotaxis and Elaboration

The final category in hypotaxis is elaboration with five measurements. Elaboration first appeared in Level 2, but it could not be classified in Levels 3 for lack of an MBC. Nonetheless, this general category emerges again in Level 4 to compare the differences in MBUs.

The first MBU in hypotaxis and elaboration (Table 99) for the EFCAMDAT2 incorporated *which*. Findings showed low frequencies at A1 (0.3) and A2 (0.3). A gradual increase appeared starting at B1 (0.6) and continuing through B2 (1.4), C1 (2.3) and C2 (3.3). Data from the CEDEL2/CAES used *que*, with the same meaning of the English *which*. Results started at A1 (0.2) and showed an increase at A2 (0.4) followed by a jump in frequency at B1 (3.0). There was a decrease at B2 (1.1) followed by slight increases at C1 (1.9) and C2 (2.2).

Corpus	Levels 1 & 2	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis Elaboration	n/a	<i>which</i>	0.3	0.3	0.6	1.4	2.3	3.3
CEDEL2/ CAES			<i>que</i>	0.2	0.4	3.0	1.1	1.9	2.2
EFCAMDAT2			<i>that</i>	0.0	0.2	0.8	1.2	1.8	0.0
CEDEL2/ CAES			<i>que</i>	0.0	0.0	1.1	0.4	0.7	1.2
EFCAMDAT2			<i>who</i>	0.1	0.5	0.5	0.7	0.5	0.0
CEDEL2/ CAES			<i>quien</i>	0.0	0.1	0.2	0.5	0.7	0.2
EFCAMDAT2			<i>when</i>	0.0	0.2	0.6	0.1	0.0	0.0
CEDEL2/ CAES			<i>cuando</i>	0.0	0.0	0.1	0.1	0.1	0.2
EFCAMDAT2			<i>where</i>	0.0	0.0	0.1	0.1	0.0	0.0
CEDEL2/ CAES			<i>donde</i>	0.0	0.0	0.0	0.1	0.1	0.0

Table 99 Level 4 results for Hypotaxis through Elaboration.

The second MBU for hypotaxis and elaboration included *that* for data extracted from the EFCAMDAT2. Results showed zero use at A1 with an increase in use starting at A2 (0.2) and continuing at B1 (0.8), B2 (1.2) and C1 (1.8). There was a decrease to zero use at C2. *Que*, with the meaning of *that*, was used for data extraction in the CEDEL2/CAES. Zero use appeared at A1 and A2. At B1 (1.1) the frequency picked up then decreased at B2 (0.4). A slight increase was seen at C1 (0.7) followed by another increase at C2 (1.2).

The third MBU in the sequence used for EFCAMDAT2 data extraction was *who*. Overall use frequency for this MBU was low from A1 (0.1) with a slight increase at A2 (0.5) and a leveling off at B1 (0.5). A second small increase happened at B2 (0.7) followed by a downward turn at C1 (0.5) and C2 (0.0). Findings from the CEDEL2/CAES used *quien* as the MBU. While there was zero use at the A1 level, a gradual increase occurred starting at A2 (0.1) and moving through B1 (0.2), B2 (0.5) and peaking at C1 (0.7). A decrease appeared at C2 (0.2).

The fourth MBU in elaboration under hypotaxis is symbolized by *when* for data taken from the EFCAMDAT2. At the A1 level there was no use of this MBU. A2 (0.2) saw the start of an increase which lasted until B1 (0.6). Following the B1 (0.6) peak level, there was a decrease at B2 (0.1) and zero use of *when* at C1 and C2. It is important to note that *when* was analyzed in the form of a relative clause and not a temporal subordinator. *Cuando* was used

as the MBU from data taken from the CEDEL2/CAES and results showed zero use in A1 and A2. Low frequency rates were seen in B1 (0.1), B2 (0.1) and C1 (0.1) and C2 (0.2).

The final MBU for hypotaxis and elaboration is *where* for the EFCAMDAT2. It is important to note *where* was analyzed as a relative clause form and not a spatial indicator as seen in hypotaxis and enhancement. The only noted use was at B1 (0.1) and B2 (0.1). Data from the CEDEL2/CAES exhibited two low frequency occurrences for *donde* at the B2 (0.1) and C1 (0.1) levels. The remaining levels including A1, A2, B2 and C2 saw zero use.

4.6.1.4 Parataxis and Extension

Parataxis is the second taxis category in this study. The first category of extension is composed of the MBSC position additive in the MBC of addition (Table 100). There are two measurements with the first one being *and* for the EFCAMDAT2. The overall frequency of use was elevated when compared to the majority of MBUs in this study with increases seen from A1 (40.8) to A2 (73.7) and B1 (97.5). After the peak at B1 (97.5), there was a decrease starting at B2 (88.9) which lasted through C1 (85.8) and C2 (74.2). Data extracted for *y* from the CEDEL2/CAES showed a much higher use frequency with an apex at A1 (116.6). A decrease appeared at A2 (106.4) with an increase at B1 (116.4) to almost the same height of frequency as A1 (116.6). B2 (114.6) saw a decrease. A relatively small increase appeared at C1 (115.7) followed by a final decline in use at C2 (112.9).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2			<i>and</i>	40.8	73.7	95.7	88.9	85.8	74.2
CEDEL2/ CAES	Parataxis Extension Addition	Positive Additive	<i>y</i>	116.6	106.4	116.4	114.6	115.7	112.9
EFCAMDAT2			<i>not only ... but also</i>	0.0	0.0	0.1	0.6	0.6	1.7
CEDEL2/ CAES			<i>no solo ... sino también</i>	0.0	0.0	0.0	0.2	0.2	0.4

Table 100 Level 4 Results Parataxis Extension and Addition through Positive Additive.

The second MBU for positive additive under addition for the EFCAMDAT2 incorporated *not only ... but also* with results displaying a tendency for much lower use than the previous MBU. The A1 and A2 levels showed zero use. There was an increase from B1 (0.1) to B2 (0.6) with C1 (0.6) maintaining the same frequency as the upper intermediate

level. A final increase to peak use was seen at C2 (1.7). The MBU used in the CEDEL2/CAES incorporated *no solo ... sino también* which showed zero frequency for A1, A2 and B1. The first instance of use appeared at B2 (0.2) and C1 (0.2) with an increase at C2 (0.4).

The second MBSC for parataxis and extension through addition is negative additive which is made up of one MBU (Table 101). Findings from the EFCAMDAT2 used *neither ... nor* with there being a low frequency at A1 (0.2), A2 (0.2), B1 (0.2) and C1 (0.2). B2 (0.1) had lower use than previously mentioned levels followed by zero use at C2. CEDEL2/CAES used the MBU *ni ... ni* which saw a decrease in use from A1 (0.7) to A2 (0.1). The proceeding proficiency levels showed an increment in use starting at B1 (0.3) and moving on to B2 (0.8), C1 (2.1) and C2 (3.1).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis Extension	Negative	<i>neither ... nor</i>	0.2	0.2	0.2	0.1	0.2	0.0
CEDEL2/ CAES	Addition	Additive	<i>ni ... ni</i>	0.7	0.1	0.3	0.8	2.1	3.1

Table 101 Level 4 Results Parataxis Extension and Addition through Negative Additive.

The adversative MBSC in parataxis and extension through the MBSC of addition is composed of two measures (Table 102) with the first one employing *and yet* in the EFCAMDAT2 which showed zero results in all six proficiency levels. Findings from the CEDEL2/CAES showed zero use for *y aun así* from A1 to B1 and a low frequency at B2 (0.1), C1 (0.1) and C2 (0.1).

The second MBU included in this study was the adversative *but* for the EFCAMDAT2 and showed a peak in use at A1 (23.9) followed by a decrease at A2 (17.4) and an increase at B1 (21.7). The remaining levels saw a continuous decline in use beginning at B2 (17.2) and continuing to C1 (12.4) and C2 (9.2). Findings from the CEDEL2/CAES observed an undulating pattern with the adversative *pero*. There was a relatively minor decrease from A1 (20.9) to A2 (20.0). The intermediate B1 (26.3) level saw an increase followed by a decrease at B2 (24.6). The same pattern repeated itself with an increase at C1 (25.0) and then a decrease at C2 (22.7).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2			<i>and yet</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES	Parataxis	Adversative	<i>y aun así</i>	0.0	0.0	0.0	0.1	0.1	0.1
EFCAMDAT2	Extension		<i>but</i>	23.9	17.4	21.7	17.2	12.4	9.2
CEDEL2/ CAES	Addition		<i>pero</i>	20.9	20.0	26.3	24.6	25.0	22.7

Table 102 Level 4 Results Parataxis Extension and Addition through Adversative.

Variation is the next MBC in parataxis and extension with a total of two measures (Table 103). It is important to note there is no MBSC in this classification, therefore MBUs will be the central focus. The first MBU for the EFCAMDAT2 was *except* which demonstrated the highest use frequency at A1 (0.2) and A2 (0.2) with a decline at B1 (0.1) and B2 (0.1). At the two advanced levels there was zero use frequency. *Excepto* was used to extract data from the CEDEL2/CAES with zero use at A1, B1 and C2. The remaining A2 (0.1), B1 (0.1) and C1 (0.1) exhibited low frequencies.

The second MBU is the variative *but*, which entails a different meaning-based parameter than the adversative *but*. EFCAMDAT2 data showed an increase from A1 (1.3) to A2 (2.1). There was a decrease at B1 (1.3) followed by an increase starting at B2 (3.2) and continuing through C1 (3.3) and C2 (5.8). The variative *pero* taken from the CEDEL2/CAES also showed an increase, yet at a lower frequency from A1 (0.0) to A2 (0.4) through to B1 (1.2). From the peak at B1 (1.2), there was a decrease to B2 (0.7) and C1 (0.5). A final increase was seen at C2 (1.0).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2			<i>except</i>	0.2	0.2	0.1	0.1	0.0	0.0
CEDEL2/ CAES	Parataxis	n/a	<i>excepto</i>	0.0	0.1	0.0	0.1	0.1	0.0
EFCAMDAT2	Extension		<i>but</i>	1.3	2.1	1.3	3.2	3.3	5.8
CEDEL2/ CAES	Variation		<i>pero</i>	0.0	0.4	1.2	0.7	0.5	1.0

Table 103 Level 4 Results Parataxis Extension through Variation.

Alternation through parataxis and extension does not include a MBSC (Table 104). *Or* was used in the findings taken from the EFCAMDAT2 with a u-shaped curve being present. From A1 (8.6) to A2 (7.9) there was a decrease. The decline in use continued

through B1 (6.7) and B2 (4.2) with an upswing in frequency at C1 (7.1) and C2 (11.7). A more erratic pattern appeared in results using *o* from the CEDEL2/CAES. The highest rate of use appeared at A1 (7.4) with a decline at A2 (2.5). At the intermediate levels there were increases at B1 (4.3) and B2 (5.7). The lower advanced level saw a decrease in use at C1 (5.0) followed by an increase at C2 (6.6).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis	n/a	<i>or</i>	8.6	7.9	6.7	4.2	7.1	11.7
CEDEL2/ CAES	Extension Alternation		<i>o</i>	7.4	2.5	4.3	5.7	5.0	6.6

Table 104 Level 4 Results Parataxis Extension through Alternation.

4.6.1.5 Parataxis and Enhancement

Enhancement has the highest number of MBSCs in parataxis with the first one being same time temporal enhancement (Table 105). The first MBU for EFCAMDAT2 involved *and meanwhile* while the CEDEL2/CAES used *y mientras tanto*. Finding is both data sets showed zero use.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis	Same time	<i>and meanwhile</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES	Enhancement Temporal		<i>y mientras tanto</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 105 Level 4 results for Parataxis through Temporal Same Time Enhancement.

Four measurements were included for different time temporal enhancement under parataxis (Table 106). Findings for EFCAMDAT2 employed *and then* which from A1 (1.3) to A2 (1.4) showed a slight increase followed by a continuous decrease beginning at B1 (1.1) and going through B2 (0.8), C1 (0.5) and C2 (0.0). CEDEL2/CAES2 data showed Spanish L2 learners used *y luego* at a low frequency. The rate at A1 (0.4) and A2 (0.4) was the same with a decline at B1 (0.3) and B2 (0.2). The advanced levels saw an increase in use with both C1 (0.6) and C2 (0.6) having the same rate.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis Enhancement Temporal	Different time	<i>and then</i>	1.3	1.4	1.1	0.8	0.5	0.0
CEDEL2/ CAES			<i>y luego</i>	0.4	0.4	0.3	0.2	0.6	0.6
EFCAMDAT2			<i>and afterwards</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>y después</i>	0.0	0.3	0.5	0.9	1.0	1.0
EFCAMDAT2			<i>and before that</i>	0.0	0.1	0.0	0.1	0.0	0.0
CEDEL2/ CAES			<i>y antes</i>	0.0	0.1	0.1	0.1	0.1	0.1
EFCAMDAT2			<i>but before that</i>	0.0	0.1	0.1	0.0	0.1	0.0
CEDEL2/ CAES			<i>pero antes</i>	0.0	0.0	0.0	0.1	0.1	0.0

Table 106 Level 4 results for Parataxis through Temporal Different Time Enhancement.

The second MBU for different time temporal enhancement for the EFCAMDAT2 incorporated *and afterwards* which showed zero use at any level. Results for *y después* in the CEDEL2/CAES began with no use at A1 yet saw increases in the remaining five levels including A2 (0.3), B1 (0.5), B2 (0.9), C1 (1.0) and a plateau C2 (1.0).

Results for the different time temporal indicator *and before that* in the EFCAMDAT2 showed a low frequency at A2 (0.1) and B2 (0.1) with the remaining four levels showing null use. The CEDEL2/CAES used *y antes* which resulted in zero use at A1. Levels A2 through C2 all had a 0.1 frequency.

The final MBU for different time temporal indicators for enhancement under parataxis used *but before* for the EFCAMDAT2. Results showed low use at A2 (0.1), B1 (0.1) and C1 (0.1) with the remaining levels not showing any use. The CEDEL2/CAES demonstrated an equally low frequency at B2 (0.1) and C1 (0.1). The rest of the levels showed zero use.

Same place spatial enhancement under parataxis had one measure in the MBSC (Table 107). The MBU used in the EFCAMDAT2 was *and there*. Results showed the peak in frequency at A1 (4.3) with a much lower frequency occurring at A2 (0.6), B1 (0.8), B2 (0.5), C1 (0.4) and C2 (0.0). Results for *y allí* in the CEDEL2/CAES had a low frequency at A1

(0.2) and C2 (0.2) followed by a lower frequency at B1 (0.1). The remaining levels did not register any occurrences.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis		<i>and there</i>	4.3	0.6	0.8	0.5	0.4	0.0
CEDEL2/ CAES	Enhancement Spatial	Same place	<i>y allí</i>	0.2	0.0	0.1	0.0	0.0	0.2

Table 107 Level 4 results for Parataxis through Spatial Same Place Enhancement.

Manner and means enhancement in parataxis included one measurement. *And in that way* was used to extract data from the EFCAMDAT2 (Table 108). The only appearance of use appeared at A1 (0.1) with no recorded use at the other five proficiency levels. Findings taken from the CEDEL2/CAES for *y esa manera* displayed zero use in all six CEFR levels.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis		<i>and in that way</i>	0.1	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES	Enhancement Manner	Means	<i>y de esa manera</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 108 Level 4 results for Parataxis through Manner and Means Enhancement.

Manner and comparison through parataxis and enhancement incorporated one measurement (Table 109). *Thus* was the MBU used in the EFCAMDAT2 search and only showed use at B2 (0.1) with no use in the remaining five levels. *Así* extracted from the CEDEL2/CAES showed a different pattern with an increase from A1 (0.2) to A2 (0.4) and B1 (1.0). A decrease happened at B2 (0.5) with a resulting rising trend at C1 (0.6) and a peak at C2 (1.4).

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis		<i>thus</i>	0.0	0.0	0.0	0.1	0.0	0.0
CEDEL2/ CAES	Enhancement Manner	Comparison	<i>así</i>	0.2	0.4	1.0	0.5	0.6	1.4

Table 109 Level 4 results for Parataxis through Manner and Comparison Enhancement.

Enhancement through cause and reason in parataxis (Table 110) contains one measure for both data sets with *and so* used for the EFCAMDAT2. At the A1 and C2 levels there was zero use while A2 (0.2), B1 (0.2), B2 (0.2) and C1 (0.2) all had a constant low frequency.

Data from the CEDEL2/CAES with *y por lo tanto* displayed zero use at the A1, A2 and B1 levels while B2 (0.1), C1 (0.1) and C2 (0.1) exhibited low frequency.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis		<i>and so</i>	0.0	0.2	0.2	0.2	0.2	0.0
CEDEL2/ CAES	Enhancement Causal- conditional	Cause & reason	<i>y por lo tanto</i>	0.0	0.0	0.0	0.1	0.1	0.1

Table 110 Level 4 results for Parataxis through Cause and Reason Causal-conditional Enhancement.

Positive condition in causal-conditional enhancement under parataxis (Table 111) contains one measurement *and in that case* for the EFCAMDAT2. The indicator saw zero use. The findings for *y en este caso* in the CEDEL2/CAES had one appearance of frequency happening at B1 (0.1) with the remaining levels demonstrating zero use.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Hypotaxis		<i>and in that case</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES	Enhancement Causal- conditional	Positive Condition	<i>y en este caso</i>	0.0	0.0	0.0	0.1	0.0	0.0

Table 111 Level 4 results for Parataxis through Positive Condition Causal-conditional Enhancement.

The concessive condition is the final MBSC under causal-conditional enhancement in parataxis with two measurements (Table 112). The first set of MBUs including *and yet* in the EFCAMDAT2 and *y aun así* registered zero frequency in all six proficiency levels in both data sets.

The second MBU in the concessive conditional MBSC includes *but nevertheless* in the EFCAMDAT2. At the B2 and C1 levels there was 0.1 frequency. The remaining four levels registered zero use. Results for the CEDEL2/CAES employed *pero aun así* with no signs of use in all six CEFR levels.

Corpus	Levels 1, 2, 3	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis Enhancement Causal- conditional	Concessive Condition	<i>and yet</i>	0.0	0.0	0.0	0.0	0.0	0.0
CEDEL2/ CAES			<i>y aun así</i>	0.0	0.0	0.0	0.0	0.0	0.0
EFCAMDAT2			<i>but nevertheless</i>	0.0	0.0	0.0	0.1	0.1	0.0
CEDEL2/ CAES			<i>pero aun así</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table 112 Level 4 results for Parataxis through Positive Condition Causal-conditional Enhancement.

4.6.1.6 Parataxis and Elaboration

Elaboration is the final category in parataxis containing three MBSCs (Table 113) with exposition being the first. The EFCAMDAT2 used *in other words* which registered zero frequency at A1, A2 and C1. B1 (0.1) and B2 (0.1) saw low frequency while C1 (0.2) saw a slight increase. Results for *en otras palabras* in the CEDEL2/CAES registered zero use at A1, A2 and B1 and C2. The only recorded occurrences for this MBU happened at a low frequency at B2 (0.1) and C1 (0.1).

The second MBU under elaboration and exposition incorporated *that is to say* in the EFCAMDAT2 which saw a low frequency at C1 (0.2) with the rest of the CEFR levels having no use. *Es decir* was used as the MBU in the CEDEL2/CAES with zero use at A1 and A2. At the intermediate level peak use happened at B1 (1.7) with a lower frequency happening at B2 (0.4) and C1 (0.4). C2 (0.7) saw a final increase yet it did not reach the height of B1 (1.7).

Corpus	Levels 1 & 2	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis Elaboration	Exposition	<i>in other words</i>	0.0	0.0	0.1	0.1	0.2	0.0
CEDEL2/ CAES			<i>en otras palabras</i>	0.0	0.0	0.0	0.1	0.1	0.0
EFCAMDAT2			<i>that is to say</i>	0.0	0.0	0.0	0.0	0.2	0.0
CEDEL2/ CAES			<i>es decir</i>	0.0	0.0	1.7	0.4	0.4	0.7
EFCAMDAT2			<i>I mean</i>	0.0	0.1	0.4	0.7	0.1	0.8
CEDEL2/ CAES			<i>quiero decir</i>	0.4	0.2	0.1	0.1	0.1	0.4

Table 113 Level 4 results for Parataxis Elaboration through Exposition.

The third and final MBU for exposition was *I mean* for the EFCAMDAT2 and while there was zero use at A1, frequency increased at A2 (0.1), B1 (0.4), and B2 (0.7). There was a decrease at C1 (0.1) followed by an increase at C2 (0.8). Results for *quiero decir* in the CEDEL2/CAES, although in general displaying low frequency, had an open u-shaped curve starting at A1 (0.4) then decreasing at A2 (0.2), B1 (0.1), B2 (0.1) and C1 (0.1). Frequency rebounded at C2 (0.4) to match the starting point at A1 (0.2).

Exemplification is the second MBSC in parataxis and elaboration and contains three measurements (Table 114) with the first one being *such as* in the EFCAMDAT2. Results show a steady increase starting from A1 (0.1) then continuing through the rest of the CEFR levels with A2 (0.4), B1 (0.6), B2 (1.1), C1 (2.1) and C2 (4.2). Findings from the CEDEL2/CAES displayed a different pattern for *tal como* with a decrease from A1 (0.2) to A2 (0.1) then a slight rise at B1 (0.2). B2 and C2 showed null frequency with a low frequency reported at C1 (0.1).

The second MBU for exemplification incorporated *for example* in the EFCAMDAT2 and showed an increase from A1 (0.5) through A2 (0.9) and peaked at B1 (2.6). Following B1 (2.6), frequency decreased at B2 (1.2) then saw an uptick at C1 (1.5) and C2 (2.5). The CEDEL2/CAES utilized *por ejemplo* to extract exemplification data with zero use at A1 followed by increases at A2 (0.9), B1 (1.3), B2 (2.1) and C1 (2.2). A final slight decrease appeared at C2 (2.1).

The last MBU in exemplification under parataxis and elaboration in the EFCAMDAT2 was *in particular* and showed use at B2 (0.1) and C1 (0.2). A1, A2, B1 and C2 exhibited zero frequency. The MBU for CEDEL2/CAES data involved *en particular* and registered zero frequency at A1, A2 and B1. B2 (0.2) saw the first occurrence of frequency with a slight decrease at C1 (0.1) then a slight rise at C2 (0.2)

Corpus	Levels 1	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2	Parataxis Elaboration	Exemplificat -ion	<i>such as</i>	0.1	0.4	0.6	1.1	2.1	4.2
CEDEL2/ CAES			<i>tal como</i>	0.2	0.1	0.2	0.0	0.1	0.0
EFCAMDAT2			<i>for example</i>	0.5	0.9	2.6	1.2	1.5	2.5
CEDEL2/ CAES			<i>por ejemplo</i>	0.0	0.9	1.3	2.1	2.2	2.1
EFCAMDAT2			<i>in particular</i>	0.0	0.0	0.0	0.1	0.2	0.0
CEDEL2/ CAES			<i>en particular</i>	0.0	0.0	0.0	0.2	0.1	0.2

Table 114 Level 4 results for Parataxis Elaboration through Exemplification.

The final MBSC for parataxis and elaboration is clarification which encompasses seven measures (Table 115). The first MBU for the EFCAMDAT2 is *how* which begins with a low use rate at A1 (0.2) only to rise at A2 (2.2) then to decrease at B1 (1.7). The same pattern appears to repeat itself with an increase at B2 (4.6) followed by a decrease at C1 (3.0) and yet another increase at C2 (5.0). A different sequence appeared with *cómo* in the CEDEL2/CAES with zero frequency at A1 (0.0) followed by increases at A2 (0.5), B1 (1.0), B2 (1.3), C1 (1.7) with the apex at C2 (2.3).

The second MBU for the clarification MBSC is *when* for the EFCAMDAT2. Results for A1 (0.2) start with a low frequency but generally increase through A2 (2.0), B1 (2.8) and apex at B2 (4.6). After the upper intermediate level, there is a decline in use at C1 (2.6) and C2 (1.7). Data for *cuándo* retrieved from the CEDEL2/CAES showed a pattern that started with zero use at A1 (0.0) with an increase at A2 (0.1) and B1 (0.3). There was a decrease at B2 (0.2) followed by an increase at C1 (0.3) and then a decrease to zero use at C2.

The third MBU in the sequences for clarification is *what*. Data taken from the EFCAMDAT2 showed no use at A1 with a slight increase and plateau at A2 (0.1) and B1 (0.1). At B2 (0.6) frequency peaked with a lower rate at C1 (0.5) followed by no use at C2. Results for *qué* retrieved from the CEDEL2/CAES showed a roller coaster like pattern with a small increase from A1 (0.2) to A2 (0.3) followed by a decrease at B1 (0.2) then at B2 (0.1). A second rise in frequency presented itself at C1 (0.2) followed by an increase to peak frequency at C2 (0.6).

Corpus	Levels 1	MBSC	Level 4 MBU	A1	A2	B1	B2	C1	C2
EFCAMDAT2			<i>how</i>	0.2	2.2	1.7	4.6	3.0	5.0
CEDEL2/ CAES			<i>cómo</i>	0.0	0.5	1.0	1.3	1.7	2.3
EFCAMDAT2			<i>when</i>	0.2	2.0	2.8	4.6	2.6	1.7
CEDEL2/ CAES			<i>cuándo</i>	0.0	0.1	0.3	0.2	0.3	0.0
EFCAMDAT2			<i>what</i>	0.0	0.1	0.1	0.6	0.5	0.0
CEDEL2/ CAES			<i>qué</i>	0.2	0.3	0.2	0.1	0.2	0.6
EFCAMDAT2			<i>in fact</i>	0.1	0.1	0.6	1.0	1.0	2.5
CEDEL2/ CAES	Parataxis Elaboration	Clarification	<i>de hecho</i>	0.0	0.0	0.0	0.4	0.3	0.7
EFCAMDAT2			<i>actually</i>	0.2	0.2	0.8	1.2	0.9	3.3
CEDEL2/ CAES			<i>de verdad</i>	0.0	0.0	0.1	0.2	0.6	0.5
EFCAMDAT2			<i>indeed</i>	0.0	0.1	0.1	0.2	0.2	0.0
CEDEL2/ CAES			<i>en efecto</i>	0.0	0.0	0.0	0.0	0.0	0.1
EFCAMDAT2			<i>at least</i>	0.3	0.2	0.7	1.0	1.5	4.2
CEDEL2/ CAES			<i>por lo menos</i>	0.0	0.1	0.0	0.3	0.6	1.0

Table 115 Level 4 results for Parataxis Elaboration through Clarification.

Results for *in fact* in the EFCAMDAT2 had a low use at A1 (0.1) and A2 (0.1). The sequence saw an increase starting at B1 (0.6) then continuing to B2 (1.0) and C1 (1.0) with a final peak at C2 (2.5). Results for *de hecho* in the CEDEL2/CAES had zero use in A1, A2 and B1. At B2 (0.4) frequency first appeared with a decrease at C1 (0.3) and a final rise at C2 (0.7).

Frequency for *actually* in A1 (0.2) and A2 (0.2) remained the same with an increase at B1 (0.8) and B2 (1.2). There was a decline at C1 (0.9) with a relative surge in use at C2 (3.3). *De verdad* was used as the MBU of clarification in the CEDEL2/CAES with zero use at A1 and A2. Frequency started at B1 (0.1) and increased in B2 (0.2) and C1 (0.6). A final decrease was seen at C2 (0.5).

The sixth MBU entails *indeed* for the EFCAMDAT2 which exhibited zero use at A1 followed by low frequency plateau at A2 (0.1) and B1 (0.1). Frequency slightly increased at B2 (0.2) and leveled off at C1 (0.2). Zero use of *indeed* was seen at C2. *En efecto* in the

CEDEL2/CAES only saw results at the C2 (0.1) level with no use being registered at any of the previous lower proficiency levels.

The seventh and final MBU in clarification is *at least* in the EFCAMDAT2. Frequency declined from A1 (0.3) to A2 (0.2). A gradual increase started at B1 (0.7) and continued through B2 (1.0), C1 (1.5) and C2 (4.2). *Por lo menos* in the CEDEL2/CAES displayed an irregular use pattern with zero frequency at A1 and B1. There was an instance of use at A2 (0.1) then a gradual upward trend starting at B2 (0.3), C1 (0.6) and C2 (1.0).

4.6.2 Discussion

Level 4 takes a profound look into an area of research which has not received much attention with its focal point including a vast array of MBUs that form the foundation of taxis, expansion, logico-semantic-relations and MBCs. It must be acknowledged that Level 4 MBUs are not a complete list of indicators for hypotaxis or parataxis. MBUs were compiled to pull together a comprehensive and equivalent list based on elements found in *Introduction to Functional Grammar* (2014) and *Método de los Relojes* (2018) with the aim of comparing Spanish and English in L2 written texts.

4.6.2.1 Hypotaxis and Extension

The results for MBUs in hypotaxis and extension demonstrated that five of the seven MBUs (*whereas*, *besides*, *apart from*, *instead of* and *if not ... then*) in the EFCAMDAT2 had a frequency of between 0.1 to 0.9 (Table 116) which constitute what this study is defining as the high density low frequency (HDLF) range. With the majority of MBUs appearing in the 0.1-0.9 range, albeit ruling out *besides* since it appeared in four CEFR levels, the greater number of subordinators reached peak frequency at C1 or C2. The only exception was *if not ... then* which peaked at B2 and C1. Findings showed that *without*, which also apexed at C1, was the only low density medium frequency (LDMF) MBU in the range between 1.0 to 1.9. *Except that* was the single MBU which exhibited zero use in L2 English texts. Learners from both data sets used extension MBUs at an overall low frequency.

Zero Frequency	High Density Low Frequency	Low Density Medium Frequency
0	0.1-0.9	1.0-1.9
<i>except that</i>	<i>whereas</i> (C1) <i>besides</i> (A1/A2/B2/C1) <i>apart from</i> (C2) <i>instead of</i> (C2) <i>if not ... then</i> (B2/C1)	<i>without</i> (C1)

Table 116 EFCAMDAT2 Hypotaxis and Extension MBU Peak Frequency Per 5k.

What proved interesting was that out of seven MBUs in Table 117, five (*whereas*, *besides*, *apart from*, *instead of* and *without*) demonstrated a peak frequency at C1 or C2, while a fifth MBU (*if not ... then*) was observed to have a peak frequency between B2 and C1. Empirical results from hypotaxis and extension do not confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

A1	A2	B1	B2	C1	C2
n/a	n/a	n/a	<i>if not ... then</i>	<i>if not ... then</i> <i>whereas</i> <i>without</i>	<i>apart from</i> <i>instead of</i>

Table 117 EFCAMDAT2 Hypotaxis and Extension MBU Peak Frequency CEFR Proficiency Level.

Results for *besides* remained inconclusive since there was a 0.1 frequency in A1, A2, B1 and C1. As use is spread through the beginner, intermediate and advanced levels, no clear pattern was observed. It is suggested that more data needs to be retrieved to clarify findings incorporating *besides*. There was zero use of *except that* which suggests this MBU falls outside the linguistic toolbox of L1 Spanish writers in L2 English.

Table 118 indicates MBUs as they distributed throughout three different frequency ranges. Hypotactic extension is a category with relatively few MBUs in which one falls into the zero-use category. Five MBUs are found in the HDLF range while only one is seen in the LDMF. When compared to Table 117, the majority of MBUs are HDLF and they are used at advanced levels. Data indicates that learners acquire hypotaxis and extension MBUs as their proficiency advances.

# MBUs	1	5	1
Frequency	0	0.1-0.9	1.0-1.9
	Zero	HDLF	LDMF

Table 118 EFCAMDAT2 Hypotaxis through extension MBU density across frequency ranges.

Finding from the CEDEL2/CAES for hypotaxis and extension in Table 119 were composed of a total of five MBUs (*mientras que*, *además de que*, *aparte de*, *sin que* and *en vez de*) which all registered between 0.1 to 0.9 frequency in HDLF. Three of the five HDLF MBUs (*además de que*, *aparte de* and *en vez de*) peaked at advanced levels. Two MBUs (*sin que* and *mientras que*) reached maximum frequency at B2, C1 and C2. *Salvo* was the only MBU to show zero use. *Si no*, which manifested at C2, was the only MBU for extension to be in the LDMF category.

Zero Frequency	High Density Low Frequency	Low Density Medium Frequency
0	0.1-0.9	1.0-1.9
<i>salvo</i>	<i>mientras que</i> (B2/C1/C2) <i>además de que</i> (C2) <i>aparte de</i> (C2) <i>sin que</i> (B2/C1) <i>en vez de</i> (C2)	<i>si no</i> (C2)

Table 119 CEDEL2/CAES Hypotaxis and Extension MBU Peak Frequency per 5k.

When looking at which proficiency level L2 Spanish writers used hypotaxis and extension in Table 120, *mientras que* plateaued at B2, C1 and C2. *Sin que* followed a similar pattern with the peak in frequency coming in at B2 and C1. The remaining MBUs including *además de que*, *aparte de*, *sin que*, *en vez de* and *si no* all saw the highest frequency at either C1 or C2.

A1	A2	B1	B2	C1	C2
			<i>mientras que</i>	<i>mientras que</i>	<i>mientras que</i>
			<i>sin que</i>	<i>sin que</i>	<i>además de que</i>
n/a	n/a	n/a			<i>aparte de</i>
					<i>en vez de</i>
					<i>si no</i>

Table 120 CEDEL2/CAES Hypotaxis and Extension MBU Peak Frequency CEFR Level.

The peak frequency results from the CEDEL2/CAES indicate that learners started to use these forms at the B2 with the majority of peak frequency MBUs being used at C1 and C2. The DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed for MBUs in hypotaxis and extension in the CEDEL2/CAES.

Table 121 shows the density and frequency spread of MBUs in hypotaxis and extension in the CEDEL2/CAES. The majority of MBUs are in HDLF. Only one MBU showed zero frequency and there was another occurrence of one MBU appearing in the LDMF. When cross-referenced with Table 120, the majority of MBUs are found at C1 and C2 in the HDLF. Findings indicate that L2 Spanish learners acquire hypotaxis through extension as their proficiency increases. However, learners only rely on this resource at a low frequency which suggests these MBUs are on the outer limits of learners' linguistic repertoire.

# MBUs	1	5	1
Frequency	0	0.1-0.9	1.0-1.9
	Zero	HDLF	LDMF

Table 121 CEDEL2/CAES Hypotaxis through extension MBU density across frequency ranges.

The two MBUs in the EFCAMDAT2 and CEDEL2/CAES that showed zero use (*except that* and *salvo*) shared an interesting correlation in the sense they were grouped together as meaning-based equivalents. When looking at the phenomenon through the focal point of hypotaxis and extension through subtractive variation, findings suggest learners do not use this MBSC. This study proposes there is no L1 transfer to L2 with subtractive extension. Furthermore, learners might not be exposed these two subordinators or taught how to use them in a subtractive manner.

The highest frequency MBU was *without* in EFCAMDAT2 data and appeared in the MBSC of adversative addition at C1 (1.0). There are two different trends which highlight findings for extension. Firstly, EFCAMDAT2 learners reached their peak use frequency at the C1 level and showed a decrease in frequency at C2 which may be due to the smaller representative sample at said level (Table 12) as seen in the Chen et al. (2021) study. These findings reflect trends found in Level 1 taxis (Figure 13) and Level 2 expansion (Figure 15). A different pattern emerged in the CEDEL2/CAES with *si no* showing up with the highest frequency in alternation at C2 (1.3). CEDEL2/CAES writers reached their frequency peak at C2 which does not seem to influence overall Level 1 taxis frequency (Figure 13). However, Level 2 extension (Figure 15) clearly shows how Level 4 MBUs impact the frequency curve at C2.

The second trend takes into consideration L1 to L2 transfer. L2 English writers used a higher frequency of adversative addition with a peak at C1 (1.0) in the EFCAMDAT2 which was double the peak frequency of CEDEL2/CAES writers who achieved a peak frequency at B2 (0.4) and C1 (0.4). Furthermore, the same logics persists with alteration where the highest recorded MBU is at C2 (1.3) in the CEDEL2/CAES while EFCAMDAT2 learners' highest frequency appeared at B2 (0.2) and C1 (0.2). Findings suggest that L1 transfer occurs in adversative addition and alteration through hypotaxis and extension in in both data sets at advanced levels.

The EFCAMDAT2 and the CEDEL2/CAES data sets are consistent with each other in the overall sense that hypotaxis and extension MBUs occur at less frequency than hypotactic enhancement and elaboration. Although using different L1s and different measures, results from the Sulistyaningrum and Rasyid (2015), Wenhui Xuan (2019) as well as Yaqub and Shakir (2019) studies are in line with the current study by showing hypotaxis and extension had a lower frequency than enhancement and elaboration.

This study concludes that hypotaxis and extension is an area in which L2 English and L2 Spanish learners do not use at a high frequency. Findings from the current study suggest that learners acquire MBUs in hypotaxis and extension at B2, C1 and C2, thus making it an indicator of upper intermediate and advanced proficiency. Further research is needed to see

why learners use extension at upper intermediate to advanced levels and why there is a marked decrease at beginner levels.

4.6.2.2 Hypotaxis and Enhancement

A pattern which emerged in the results from the EFCAMDAT2 (Table 122) was a cluster of subordinators in the zero-use HDLF categories. There were a total of 14 subordinators which L1 Spanish writers did not use in L2 English texts. Of the 14 zero-use MBUs, 13 came from the causal-conditional MBC with four subordinators (*seeing that*, *in case of*, *due to* and *thanks to*) appearing in the MBSC cause and reason, three subordinators (*with the aim of*, *in exchange for* and *with the intention of*) in the purpose MBSC and the remaining six subordinators (*provided that*, *as long as*, *so long as*, *on the condition that*, *even if* and *despite*) showing up in positive and concessive condition. *Upon* was the only temporal indicator with zero-use.

Zero Frequency	0.1-0.9
<i>upon</i>	<i>as</i> (temporal) (C2)
<i>seeing that</i>	<i>as soon as</i> (C2)
<i>in case of</i>	<i>the moment</i> (B1/B2)
<i>due to</i>	<i>every time</i> (C1)
<i>thanks to</i>	<i>as if</i> (C1)
<i>with the aim of</i>	<i>since</i> (cause/reason) (B2)
<i>in exchange for</i>	<i>in case</i> (C1)
<i>with the intention of</i>	<i>given that</i> (C1)
<i>provided that</i>	<i>considering</i> (C1)
<i>as long as</i>	<i>with</i> (B2)
<i>so long as</i>	<i>as a result of</i> (A2/B1)
<i>on the condition that</i>	<i>so that</i> (B1)
<i>even if</i>	<i>unless</i> (B1)
<i>despite</i>	

Table 122 EFCAMDAT2 Hypotaxis and Enhancement High Density Low Frequency MBUs.

If we look at EFCADAT2 MBUs with zero frequency in this study, 13 of 14 MBUs come from the causal-conditional MBC. Under the microscope of MBSCs, the 13 causal-conditional subordinators are classified under concession, purpose, cause and reason. A possible explanation for this phenomenon is that all the previously mentioned categories

require a high degree of cognitive complexity which will typically be the output of pragmatic enrichment (Kortmann, 1999). The application of this perspective would mean zero-use MBUs require a higher cognitive function than English L2 writers are able to fulfill.

There were 13 EFCAMDAT2 MBUs in the HDLF which were distributed over five MBCs including five in temporal (*as, as soon as, the moment, every time* and *since*), five in cause/reason (*in case, given that, considering, with* and *as a result of*) as well as one MBU in manner and comparison (*as if*), purpose (*so that*) and negative condition (*unless*). HDLF MBU findings were aligned with Wenhui Xuan (2019) as well as Rasool and Mahmood (2023) as both studies found temporal and causal-conditional reason subordinators were used at a higher frequency than manner/comparison and purpose.

Table 123 shows a summary of how MBUs are categorized in the HDLF category. In total, there were 14 subordinators which manifested null use by Spanish L1 writing in English L2. Thirteen MBUs were used at the 0.1-0.9 frequency in data extracted from the EFCAMDAT2.

# MBUs	14	13
Frequency	0	0.1-0.9
	Zero	HDLF

Table 123 EFCAMDAT2 Hypotaxis through Enhancement HDLF MBUs across Frequency Ranges.

In reference to EFCAMDAT2 HDLF MBUs per the CEFR (Table 124), subordinators were scattered across different proficiency levels. Data shows that there was one frequency peak at A1, seven instances in B1/B2 and seven instances in C1/C2 in which MBUs reached peak frequency. Looking at the phenomenon in a different perspective, if we remove the two subordinators that plateaued at two levels (*as a result of* and *the moment*), we see the majority of hypotaxis takes place at the C1 and C2 between 0.1 to 0.9 frequency. This conclusion leads us to believe that these subordinators are used at advanced levels when learners have acquired more knowledge and have a higher cognitive function (Kortmann, 1999).

A1	A2	B1	B2	C1	C2
	<i>as a result of</i>	<i>the moment</i>	<i>the moment</i>	<i>every time</i>	<i>as (temporal)</i>
		<i>as a result of</i>	<i>since</i>	<i>as if</i>	<i>as soon as</i>
n/a		<i>so that</i>	<i>with</i>	<i>in case</i>	
		<i>unless</i>		<i>given that</i>	
				<i>considering</i>	

Table 124 EFCAMDAT2 Hypotaxis and Enhancement High Density Low Frequency MBUs per CEFR Level.

EFCAMDAT2 Level 4 findings support Level 1 hypotaxis results (Figure 13) and Level 2 hypotaxis and enhancement results (Figure 16) in that HDLF MBUs follow the same trend which leads to peak frequency at C1 in English L2 texts. The DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed for hypotaxis and enhancement HDLF MBUs in the EFCAMDAT2.

CEDEL2/CAES findings contained a considerable amount of MBUs being categorized under zero frequency. A total of 15 MBUs composed the zero frequency category including one temporal (*a medida que*), four cause/reason (*en caso de que*, *como consecuencia de*, *en caso de* and *gracias a que*), three purpose (*con el objetivo de*, *a cambio de que* and *con la intención de que*) and seven positive, negative and concessive condition (*siempre que*, *siempre y cuando*, *con tal de que*, *a condición de que*, *sin*, *aun si* and *pese a que*).

Results from the CEDEL2/CAES mirrored those of the EFCAMDAT2 in the sense that the majority of zero-use MBUs, 14 out of 15 subordinators, came from the causal-conditional MBC (Table 125). *A medida de que* was the only temporal indicator with zero-use. Following the same logic as in the EFCAMDAT2, there is the possibility that this group of subordinators require a higher level of cognitive complexity in the acquisition process (Kortmann, 1999). Additionally, this may happen because of a lack of use and/or exposure (Ellis, 2008; Kemmer & Barlow, 2000).

Seventeen HDLF MBUs were found in the CEDEL2/CAES in the 0.1-0.9 frequency range. Subordinators were dispersed among a variety of MBCs. There were six MBUs in the temporal MBC (*tan pronto como*, *en el momento en que*, *al*, *cada vez de*, *desde que* and *una*

vez *que*), two in the manner MBC (*por* and *como si*), seven in the reason category (*puesto que*, *visto que*, *dado que*, *considerando que*, *con que*, *a causa de* and *debido a que*) and two in the causal-conditional category (*a menos que* and *a pesar de que*). Results resembled the Wenhui Xuan (2019) study as both temporal and reason subordinators were used at a higher frequency than manner subordinators. However, causal-conditional subordinators saw a much lower use in CEDEL2/CAES data than reported by Wenhui Xuan (2019).

Zero Frequency	0.1-0.9
<i>a medida que</i>	<i>tan pronto como</i> (A1)
<i>en caso de que</i>	<i>el momento en que</i> (C1)
<i>como consecuencia de</i>	<i>al</i> (C2)
<i>en caso de</i> (infinitive)	<i>cada vez de</i> (B1/C1)
<i>gracias a que</i>	<i>desde que</i> (A1)
<i>con el objetivo de</i>	<i>una vez que</i> (A2)
<i>a cambio de que</i>	<i>por</i> (B2)
<i>con la intención de que</i>	<i>como si</i> (C2)
<i>siempre que</i>	<i>puesto que</i> (C2)
<i>siempre y cuando</i>	<i>visto que</i> (C2)
<i>con tal de que</i>	<i>dado que</i> (C2)
<i>a condición de que</i>	<i>considerando que</i> (C1)
<i>sin</i>	<i>con que</i> (C1/C2)
<i>aun si</i>	<i>a causa de</i> (A1/C2)
<i>pese a que</i>	<i>debido a que</i> (C2)
	<i>a menos que</i> (A1)
	<i>a pesar de que</i> (C1)

Table 125 CEDEL2/CAES High Density Low Frequency MBUs.

Table 126 shows the breakdown of MBUs in HDLF category in the CEDEL2/CAES. There were 15 MBUs which showed null use. The 0.1-0.9 frequency range exhibited 17 MBUs which were classified in this category. Regarding the HDLF category, Spanish L2 learners had more zero-use MBUs as well as HDLF MBU use than L2 English writers.

# MBUs	15	17
Frequency	0	0.1-0.9
	Zero	HDLF

Table 126 CEDEL2/CAES2 Hypotaxis through Enhancement HDLF MBUs across Frequency Ranges.

MBUs for the HDLF frequency range in the CEDEL2/CAES were distributed in all six CEFR levels as seen in Table 127. Four MBUs were located in A1 (*tan pronto como*, *desde que*, *a causa de* and *a menos que*) one in A2 (*una vez que*), one in B1 (*cada vez de*), one in B2 (*por*), five in C1 (*el momento en que*, *cada vez de*, *considerando que*, *con que* and *a pesar de que*) and eight in C2 (*al*, *como si*, *puesto que*, *visto que*, *dado que*, *con que*, *a causa de* and *debido a que*). The fact that the largest grouping of HDLF MBUs happened at the C1 and C2 levels points to the evidence that L2 Spanish writers in this study use the largest quantity and most diverse array of adverbial subordinators at high proficiency levels.

A1	A2	B1	B2	C1	C2
<i>tan pronto como</i>	<i>una vez que</i>	<i>cada vez de</i>	<i>por</i>	<i>el momento en que</i>	<i>al</i>
<i>desde que</i>				<i>cada vez de</i>	<i>como si</i>
<i>a causa de</i>				<i>considerando que</i>	<i>puesto que</i>
<i>a menos que</i>				<i>con que</i>	<i>visto que</i>
				<i>a pesar de que</i>	<i>dado que</i>
					<i>con que</i>
					<i>a causa de</i>
					<i>debido a que</i>

Table 127 CEDEL2/CAES Hypotaxis and Enhancement High Density Low Frequency MBUs per CEFR Level.

This study concludes that the majority of HDLF MBUs peak at the advanced levels when learners have acquired more knowledge and have been exposed more often to target forms. The DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed for HDLF MBUs in the CEDEL2/CAES.

Level 1 taxis (Figure 13) and Level 2 expansion through hypotactic enhancement (Figure 16) showed peak frequency at C1 with a slight decrease at C2 in CEDEL2/CAES

data. While the HDLF grouping of MBUs might augment frequency at the C1 level, findings suggest that said category does not have much of an impact on the overall frequency in Level 1 hypotaxis and Level 2 hypotaxis and enhancement. This is backed up by the fact the largest grouping of MBUs, which should influence frequency, appears at C2. According to Level 1 taxis and Level 2 hypotactic enhancement results, a frequency decrease was seen at C2. Said decrease is counterintuitive to the notion in which a large grouping of low frequency MBUs have an impact on the C2 level.

An interesting observation with CEDEL2/CAES MBUs in the 0.1-0.9 frequency range is that the three levels with the highest density of are A1, C1 and C2. If we eliminate MBUs that have peak frequency at more than one level, we see the two levels with the highest density of subordination are A1 and C2. While findings suggest that subordination occurs at an advanced proficiency level, data from the A1 level shows that hypotaxis and enhancement also occurs in a variety of ways at the lower beginner level. A further matter which complicates this finding is that the A1 word count (23,020) in the CEDEL2/CAES is comparatively much lower than any of the other five categories, however there is more HDLF subordination.

A possible explanation for the increased amount of MBUs at the A1 level might be the way the CEDEL2 classifies proficiency levels through two measurements. The first one is an objective measurement in which learners are given a 43-point standardized placement exam. The second one is a subjective measurement in which learners self-rate their proficiency in Spanish according to speaking, listening, reading and writing (CEDEL2, 2023). Example (188) *tan pronto como* was extracted from the A1 level in the CEDEL2. The example shows the context in which *tan pronto como* appears in combination with the perfect and imperfect past tenses which have appeared in this study as an indicator of higher proficiency levels than A1. An in-depth explanation of which verb forms appear at different CEFR levels is beyond the reach of this study. Nevertheless, this study suggests that there exists a possibility that L1 English learners were not classified in the correct proficiency level due to the subjective standard of the second CEDEL2 measurement.

(188) *Recuerdo que **tan pronto como** Caitlin se lavaba la cabeza la lluvia paró.*
*I remember that **as soon as** Caitlin washed her head the rain stopped.*

The elephant in the room that must be acknowledged regarding the HDLF MBUs in both EFCAMDAT2 and CEDEL2/CAES data sets is why there is a large grouping of subordinators occurring at such a low frequency. The frequency of exposure might be an explanation (Ellis, 2008; Kemmer & Barlow, 2000), thus putting forth the notion that learners might not be taught these subordinators, or they might not appear in ESL curriculums.

Moving forward to focus on LDMF MBUs (Table 128) in the EFCAMDAT2, we can see that subordination goes from an increased number of indicators at the lower frequency ranges of 1.0-1.9 and 2.0-2.9 to a much smaller quantity at higher ranges. There are four MBUs at the 1.0-1.9 range (*since, once, because of* and *even though*), five at the 2.0-2.9 range (*while, before, until, by* and *as*), zero in the 3.0-3.9 range, two in the 4.0-4.9 range (*where* and *as*), one in the 5.0-5.9 range (*after*), zero in the 6.0-6.9 range, one in the 7.0-7.9 range (*to*), one in the 8.0-8.9 range (*although*) and zero in the 9.0-9.9 range.

1.0-1.9	2.0-2.9	3.0-3.9	4.0-4.9	5.0-5.9	6.0-6.9	7.0-7.9	8.0-8.9	9.0-9.9
<i>since</i> (temporal) (C2)	<i>while</i> (C1)		<i>where</i> (C1)	<i>after</i> (C2)		<i>to</i> (C2)	<i>although</i> (C1)	
<i>once</i> (C2)	<i>before</i> (B1)		<i>as</i> (manner)(C1)					
<i>because of</i> (B1)	<i>until</i> (C2)	n/a			n/a			n/a
<i>even though</i> (C1)	<i>by</i> (C1)							
	<i>as</i> (cause)(C1)							

Table 128 EFCAMDAT2 Hypotaxis and Enhancement Low Density Medium Frequency.

A large grouping of subordinators at the lower end of the frequency scale appears in LDMF. The quantity of MBUs gradually decreases as the frequency increases in what appears to be a scale effect that starts with a high number of HDLF MBUs at 0.1-0.9. The effect continues in MBUs which appear to diminish in number as frequency increases in the LDMF ranges of 1.0-1.9, 2.0-2.9 and 3.0-3.9. The important point to highlight is that L2 English writers use a large variety of MBUs at lower frequencies. As frequency begins to increase, the variety of subordinators learners use rapidly decreases (Table 129).

# MBUs	14	13	4	5	0	2	1	0	1	1	0
Freq.	0	0.1- 0.9	1.0- 1.9	2.0- 2.9	3.0- 3.9	4.0- 4.9	5.0- 5.9	6.0- 6.9	7.0- 7.9	8.0- 8.9	9.0- 9.9
	Zero	HDLF	LDMF								

Table 129 EFCAMDAT2 Hypotaxis through Enhancement MBU Density across Frequency Ranges.

There were a variety of EFCAMDAT2 MBUs scattered through six different MBCs in LDMF. The temporal MBC (*since, while, before, until and after*) contained the highest number of indicators with five in total. Concessive conditional (*even though and although*) and cause/reason (*because of and as*) MBCs both had two subordinators making them the categories with the second largest amount of MBUs. The final four MBCs included one MBU in spatial (*where*), manner means (*by*), manner comparison (*as*) and purpose (*to*).

Five EFCAMDAT2 temporal indicators were used by learners in LDMF. Except for *after* 5.0-5.9, the remaining four MBUs appeared between 1.0 to 2.9, thus making temporal subordination category the most frequently used MBC at a low bandwidth. MBUs spread across the remaining MBCs contained fewer subordinators but were used at higher frequencies including the spatial *where* (4.0-4.9), manner *as* (4.0-4.9), purpose *to* (7.0-7.9) and concessive *although* (8.0-8.9).

When the temporal indicator *after* (5.0-5.9) is added to the equation with spatial *where* (4.0-4.9), manner *as* (4.0-4.9), purpose *to* (7.0-7.9) and concessive *although* (8.0-8.9), it further sheds light on how Spanish L1 use hypotaxis and enhancement in English L2 writing. Out of the fourteen MBUs, only five subordinators are used at a frequency superior to 2.0-2.9. While explanations involving cognitive linguistic theory focus on generalities of L1 transfer to the L2, there appears to be a void when applying these fundamental aspects to fine-grained measure within the construct of subordination frequency. What the current study has been able to identify is that L1 to L2 transfer appears to take place in LDMF MBUs in frequency levels above 2.0-2.9 when L2 writers dynamically apply L1-based cognitive resources when needed in L2 learning or use (Cummins, 2008; Forbes, 2019; Siegel, 2003; Yan, 2010).

Table 130 shows how LDMF MBUs establish a clear pattern in which the majority of indicators had a peak frequency at advanced levels. The C1 level contained seven subordinators (*even though, while, by, as* <<cause>>, *where, as* <<manner>> and *although*) while the C2 level included five subordinators (*since, once, until, after* and *to*). There were a total of twelve subordinators appearing at advanced levels. In contrast, only two MBUs (*because of* and *before*) were observed at the B1 level with zero indicators at the A1, A2 and B2 levels. With only two MBUs at the B1 level and zero at the B2 level, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed for LDMF MBUs in the EFCAMDAT2.

A1	A2	B1	B2	C1	C2
		<i>because of</i>		<i>even though</i>	<i>since</i>
		<i>before</i>		<i>while</i>	<i>once</i>
				<i>by</i>	<i>until</i>
n/a	n/a		n/a	<i>as (manner)</i>	<i>after</i>
				<i>where</i>	<i>to</i>
				<i>as (cause)</i>	
				<i>although</i>	

Table 130 EFCAMDAT2 Hypotaxis and Enhancement Low Density Medium Frequency.

The tendency for LDMF MBUs suggests the majority of subordination is used at C1 and C2 with only *because of* and *before* appearing at B1. *Before* (Table 90) had a steady use case with users applying this subordinator consistently from A1 through C2, thus making it a consistent linguistic resource used by learners. *Because of* (Table 94) may be a clear case of task-based formulaic sequencing as a singular elevated frequency was recorded at the B1 level, thus indicating the possibility that learners remembered this sequence of words and incorporated them into a communicative task designed to extrapolate said sequence (Alexopoulou et al., 2015; Chen et al., 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002).

Results for LDMF MBUs corresponded with EFCAMDAT2 Level 1 (Figure 13) and Level 2 (Figure 16) findings in the fact that subordination peaked at C1. The explanation for the comparatively high density of subordinators at the C2 level can be explained by learners increased use of temporal subordination (Figure 24). While MBU frequency tended to

substantially decrease at the C2 level for hypotaxis and enhancement, temporal subordination decreased at C1 only to rebound at C2 hence explaining why this phenomenon took place.

Nine LDMF MBUs were extracted from the CEDEL2/CAES data set (Table 131). A total of four MBUs appeared in the 1.0-1.9 range (*antes de que*, *después de que*, *hasta que* and *ya que*) one in 2.0-2.9 (*para que*), one in 3.0-3.9 (*mientras*), two in 5.0-5.9 (*como* and *aunque*) and one in 6.0-6.9 (*donde*). Relatively few subordinators were encountered in the LDMF category with the highest number of MBUs appearing at 1.0-1.9 with four subordinators, 5.0-5.9 with two subordinators. LDMF categories 2.0-2.9, 3.0-3.9 and 6.0-6.9 all contained one MBU. No established pattern was seen other than a larger grouping of subordinators at the lowest range of the medium frequency spectrum.

1.0-1.9	2.0-2.9	3.0-3.9	4.0-4.9	5.0-5.9	6.0-6.9	7.0-7.9	8.0-8.9	9.0-9.9
<i>antes de que</i> (B2)	<i>para que</i> (C2)	<i>mientras</i> (C2)		<i>como</i> (manner) (C2) <i>aunque</i> (C2)	<i>donde</i> (B2)			
<i>después de que</i> (B2)			n/a			n/a	n/a	n/a
<i>hasta que</i> (A1)								
<i>ya que</i> (C2)								

Table 131 CEDEL2/CAES Hypotaxis and Enhancement Low Density Medium Frequency MBUs.

Table 132 shows a broad overview of the total number of MBUs at the HDLF and LDMF range in the CEDEL2/CAES. Subordination mostly occurs between 0.1-0.9 with seventeen MBUs appearing in this frequency classification. A sharp decrease happens as the LDMF begins with four indicators at 1.0-1.9, one at 2.0-2.9, one at 3.0-3.9, two at 5.0-5.9 and one at 6.0-6.9. As frequency increases, the density of MBUs per CEFR level plummets with no subordination taking place between 7.0-7.9 to 9.0-9.9.

# MBUs	15	17	4	1	1	0	2	1	0	0	0
Freq.	0	0.1- 0.9	1.0- 1.9	2.0- 2.9	3.0- 3.9	4.0- 4.9	5.0- 5.9	6.0- 6.9	7.0- 7.9	8.0- 8.9	9.0- 9.9
	HDLF		LDMF								

Table 132 CEDEL2/CAES MBU Density Across Frequency Ranges.

A trend appearing in both data sets is the cluster of MBUs at lower frequencies and a thinning out of subordination frequency at medium frequency ranges. Findings for the CEDEL2/CAES had fewer instances of LDMF subordination than the EFCAMDAT2 with only eight MBUs appearing in the category. When compared to the EFCAMDAT2 with fourteen instances, Spanish L2 writers in the CEDEL2/CAES used a smaller variety of MBUs. Subordination was additionally used at higher frequencies in the EFCAMDAT2 which had one MBU in the 7.0-7.9 and 8.0-8.9 frequency ranges. Results indicate that CEDEL2/CAES writers use fewer subordinators at a lower frequency than those from the EFCAMDAT2 in the LDMF range.

MBUs in the LDMF range were scattered across six different MBCs in the CEDEL2/CAES. Temporal subordinators (*antes de que*, *después de que*, *hasta que* and *mientras*) composed the category with the most indicators. The categories including spatial (*donde*), cause and reason (*ya que*), purpose (*para que*), manner (*como*) and concessive (*aunque*) all contained one MBU. The only correlation which can be vaguely drawn is with the Wenhui Xuan (2019) study and the Rasool and Mahmood (2023) study which showed a high percentage of temporal subordination. However, because of the relatively low density of MBUs in the medium frequency ranges other than temporal indicators, it is difficult to extrapolate a solid conclusion based on the evidence at hand.

Table 133 shows the pattern for LDMF MBUs across CEFR levels in the CEDEL2/CAES. The largest number of peak frequency MBUs appeared at the C2 level. Three subordinators appeared at the B2 level with no MBUs showing up at any other proficiency level. Being aware of the large grouping of MBUs at the C2 level, this study cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) is valid for LDMF MBUs in the CEDEL2/CAES.

A1	A2	B1	B2	C1	C2
<i>hasta que</i>			<i>antes de que</i>		<i>ya que</i>
			<i>después de que</i>		<i>para que</i>
	n/a	n/a	<i>donde</i>	n/a	<i>mientras</i>
					<i>como</i> (manner)
					<i>aunque</i>

Table 133 CEDEL2/CAES Hypotaxis and Enhancement Low Density Medium Frequency MBUs.

While it can be confirmed that hypotaxis frequency peaks at the C2 level in the LDMF category, we must take a step back and look at the case of *antes de que* (Table 90) and *donde* (Table 91). *Antes de que* and *donde* are consistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). In the case of *antes de que*, it reaches peak frequency in the 1.0-1.9 frequency category which leads to the presumption it does not have a substantial impact on the overall temporal MBC. In the case of *donde*, it appears as the MBU with the highest use frequency (6.0-6.9) in LDMF with the primary and secondary peak frequency being reached at the B1 and B2 levels and decreasing at advanced levels. What both *antes de que* and *donde* have in common is they are simple coordinators in English (Quirk et al., 1985). While *antes de que* and *donde* might be the indicators of a trend in which there is a pattern that simple subordination does apex at the intermediate levels in certain MBUs, more research will have to be done to confirm to what extent, which MBCs, at what frequencies and at which CEFR levels this takes place.

Después de que in Table 90 displays a much different pattern than *antes de que* and *donde*. From A1 (0.4) there is an increase to A2 (1.5) and a leveling off at B1 (1.5). Following B1, there is a slight increase to peak frequency at B2 (1.7) then a minor decrease at C1 (1.3). The trend for this temporal indicator was different than any other MBU in this category as it started at a relative stable frequency at the A2 and use steadily continued through C1.

Hasta que (Table 90) was the only MBU to have a peak frequency at A1 (1.1). Additionally, it was the only subordinator to not have a peak use frequency in intermediate or advanced proficiency levels. The fact that this MBU appears as an anomaly might suggest

that task-based formulaic sequencing may be occurring (Alexopoulou et al., 2015; Chen et al., 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002).

Comparing CEDEL2/CAES results for LDMF MBUs to overall results for Level 1 hypotaxis (Figure 13) and Level 2 enhancement (Figure 16), the first take away is a correspondence with a lower frequency from fewer MBUs at the B2 level surmounted by a higher frequency of use at the C2 level. In contrast, there is a slightly higher frequency rate in Level 1 at C1 (103.1) versus C2 (100.5). The same tendency repeats itself at Level 2 with C1 (97.6) appearing at a higher frequency than C2 (93.9).

CEDEL2/CAES MBUs which had an effect on the frequency curve in their respective Level 3 MBC were *donde* (6.0-6.9) in Table 91 which peaked at B2 (Figure 25) in the spatial MBC and the manner MBU *como* (5.0-5.9) in Table 93 in manner which showed an upward trend from C1 to C2 (Figure 26). Nevertheless, if looked at more in-depth, there is only one MBU in spatial and two in manner so it would be an obvious assumption that both subordinators would influence the overall frequency curve. On the other hand, the higher frequency shown by *aunque* (5.0-5.9) in Table 98 might add to the elevated frequency in the Level 3 causal-conditional MBC as there is an increase from C1 to C2 (Figure 27). However, it must be taken into account that causal-conditional MBUs form a varied and numerous range of subordinators. Therefore, whatever positive impact *aunque* might have on the MBC frequency would only add to the overall pattern.

EFCAMDAT2 LDHF MBUs are the final category for hypotaxis and enhancement (Table 134). There are a total of three subordinators used at a high frequency. *Because* (Table 94) appeared in the 10.0-19.9 range and is the only LDHF MBU categorized in the MBC for cause and reason. *When* (Table 89) shows up in the 20.0-29.9 and is the only LDHF subordinator in the temporal MBC. Finally, *if* (Table 96) occurs with a peak frequency in the 40.0-49.9 and is the only LDHF subordinator found in the causal-conditional positive condition MBC.

Wenhui Xuan (2019) in addition to Rasool and Mahmood (2023) found that hypotaxis through temporal and causal-conditional indicators displayed the highest frequencies in enhancement in their respective research. While both studies investigated different L1s, results from EFCAMDAT2 data showed that the causal-conditional reason indicator *because*,

the causal-conditional positive condition indicator *if* and the temporal indicator *when* were used at elevated frequencies within their MBC. Findings from the EFCAMDAT2, Wenhui Xuan (2019) as well as Rasool and Mahmood (2023) indicate a possible cross-cultural and cross-linguistic trend in which learners use temporal and causal-conditional MBUs at an elevated frequency.

10.0-19.9	20.0-29.9	30.0-39.9	40.0-49.9
<i>because</i> (B1)	<i>when</i> (A2)	n/a	<i>if</i> (C1)

Table 134 EFCAMDAT2 Hypotaxis and Enhancement Low Density High Frequency MBUs.

As there are a total of thirteen HDLF MBUs and fourteen LDMF MBUs (Table 129) in EFCAMDAT2 data, a stark contrast is seen with a total of three MBUs for the LDHF category. Both HDLF (0.1-0.9) and LDMF (1.0-9.9) MBUs were analyzed at a much lower frequency scale than LDHF MBUs. In contrast, LDHF MBUs encompassed a larger frequency range starting with 10.0-10.9 and finishing with 40.0-49.9. The point of highlighting the differences in frequency ranges is to underline that the number of LDHF MBUs are much fewer than any other category. The trend that appears is L2 English learners use more subordinators at a low frequency. Results indicate that as frequency increases, the number of subordinators used by learners drastically decreases showing that only a select group of MBUs are frequently employed in written texts in the LDHF category.

The cause of so few subordinators at a high frequency in EFCAMDAT2 data may be explained by developmental formulaic sequencing. Previous studies have clarified that using a large number of frequent formulaic sequences in the L2 is a key trait of fluent comprehension and production (Boers & Lindstromberg, 2012; Henriksen et al., 2013; Peters, 2014; Schmitt & Carter, 2004). Learners who produce a high density of accurate and appropriate L2 phrasal expressions are viewed to be relatively proficient (Boers et al., 2006; Dai & Ding, 2010; Schmitt, 2008; Stengers et al., 2010).

The MBU *if* achieved its peak frequency at C1 (Table 135) and has the highest frequency out of any MBU extracted from EFCAMDAT2 data in hypotaxis and enhancement. Therefore, it might be suggested that due to being at an advanced proficiency level and being at a high frequency range, *if* might be the product of developmental formulaic sequencing. As many formulaic sequences are comprised of high-frequency words (Martinez & Murphy, 2011), it means that learning these items consists not so much in learning new

words, but in strengthening bonds among words that are already known (Lindstromberg et al, 2016). An additional factor might be that learners tend to have a small inventory of formulaic sequences that they overuse (Wray, 2012, p. 235).

A1	A2	B1	B2	C1	C2
	<i>when</i>	<i>because</i>		<i>if</i>	

Table 135 EFCAMDAT2 Hypotaxis and Enhancement Low Density High Frequency MBUs per CEFR proficiency level.

If use of subordination at advanced proficiency levels and high frequencies indicates formulaic sequencing, then this explanation does not prove true for *because* (10.0-19.9) with a peak frequency at B1 and *when* (20.0-29.9) with a peak frequency at A2. This phenomenon might be due to a small inventory of formulaic sequences which are overused. However, having peak frequencies at relatively low CEFR levels is counterintuitive to the idea that learners produce a high density of accurate and appropriate L2 phrasal expressions which are a sign of proficiency (Boers et al., 2006; Dai & Ding, 2010; Schmitt, 2008; Stengers et al., 2010).

Very advanced learners can be expected to show knowledge of formulaic sequences which is like that of native speakers (Boers & Lindstromberg, 2012). There have been cross-sectional and longitudinal studies comparing learners at different proficiency levels which confirmed this genuine, but slow, development of formulaic competence when learning a foreign language (Appel & Wood, 2016; Huang, 2015; Li & Schmitt, 2009; Qi & Ding, 2011; Siyanova-Chanturia, 2015; Verspoor et al., 2012). While this might hold true for *if*, the same cannot be applied to *because* at the B1 level and *when* at the A2 level.

A tangible interpretation for the use of *because* and *when* by learners at a high frequency at A2 and B1 levels may be a result of cross-linguistic transfer of the L1 to the L2. Cross-linguistic transfer tends to contain a strategic purpose during processing. This is likely to be caused by the constraint of limited target language knowledge, which is associated with the difficulty experienced in completing a task. L2-related knowledge constitutes the foundation of L2 use, which is directly linked to L2 performance. However, if this knowledge is not adequate, an L1 counterpart or other relevant L1 resources will play a role in learners performing L2 tasks (Odlin, 2003).

When looking at the numbers of MBUs deployed across LDHF in the EFCAMDAT2 (Table 134) and across CEFR proficiency levels (Table 135), it is not exactly clear if higher frequency subordination peaks at intermediate levels and decreases at advanced levels. Table 89 shows how there is a sharp increase at A2 (21.9) with the temporal MBU *when*. The initial peak at A2 is thought to be due to task-based formulaic sequencing (Alexopoulou et al., 2015; Chen, 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002) yet there is a continued elevated frequency at B1(14.4) and B2 (14.7) with a gradual decline from C1 (9.1) to C2 (9.2). *If* (Table 96) shows a gradual rise in frequency from A1 (3.1) to a sharp peak at C1 (45.0) and a decline at C2 (6.7). In the case of *because* (Table 94), there is a clear rise in frequency from A1 (8.5) to A2 (18.4) with a double plateau effect at A2 (18.4) and B1 (18.8). There is a secondary plateau at B2 (14.4) and C1 (14.6) with a drop to zero-use at C2.

The above synopsis gives no evidence of a visible pattern. On the one hand, *if* appears to exemplify the MBU with the highest frequency in enhancement with a peak at the C1 level. On the other hand, there is no clear evidence that supports the idea in which subordination frequency peaks at the intermediate levels only to decrease at the advanced levels with *when* since the frequency decrease begins at B2 and continues through C2. In the case of *because*, there is a subordination peak from A2 to B1, yet it remains relatively stable at B2 and C1. Given the nature of the findings, this study cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) with LDHF MBUs for *if*, *because* or *when* in the EFCAMDAT2.

Although two of the three most frequently used MBUs occur at the A2 and B1 levels, Table 136 shows that L2 English writers use hypotaxis through enhancement mainly at advanced levels in terms of frequency as well as sheer number of MBUs. The overall finding from the EFCAMDAT2 data points to the trend in which subordination peak frequency happens at advanced levels, and not at intermediate levels since there is one MBU at A2, seven at B1, three at B2, twelve at C1 and eight at C2.

40.0-49.9					1
30.0-39.9					
20.0-29.9		1			
10.0-19.9			1		
9.0-9.9					
8.0-8.9				1	
7.0-7.9					1
6.0-6.9					
5.0-5.9					1
4.0-4.9				2	
3.0-3.9					
2.0-2.9			1	3	1
1.0-1.9			1	1	2
0.1-0.9			4	3	5
	A1	A2	B1	B2	C1
					C2

Table 136 EFCAMDAT2 Hypotaxis and Enhancement MBUs per Frequency and CEFR Level.

CEDEL2/CAES LDHF MBUs consisted of four separate subordinators (Table 137). *Cuando* and *si* were found to peak in the 10.0-19.9 frequency range. At a slightly higher frequency, *para* and *porque* appeared at the 20.0-29.9 range. *Cuando* was the only temporal subordinator, *si* was the only positive condition causal-conditional indicator, *para* showed up as the only causal-conditional purpose MBU and *porque* was the only causal-conditional cause and reason subordinator.

10.0-19.9	20.0-29.9	30.0-39.9	40.0-49.9
<i>cuando</i> (B1)	<i>para</i> (C1/C2)	n/a	n/a
<i>si</i> (B2)	<i>porque</i> (B1)		

Table 137 CEDEL2/CAES Hypotaxis and Enhancement Low Density High Frequency MBUs.

Findings from the CEDEL2/CAES are consistent with Wenhui Xuan (2019) since temporal, causal-conditional reason and causal-conditional positive condition MBUs occurred at a higher frequency than purpose, manner, result and concession. The Rasool and Mahmood (2023) study is consistent with findings from the current study in not only do they confirm that temporal, causal-conditional reason and causal-conditional positive condition are used at higher frequencies, but they found that purpose MBUs are used in L2 texts at a relatively high level. The correlation seen in the two previously mentioned studies and L2 Spanish writers show a possible and intriguing cross-linguistic trend in which more research will need to be carried out to determine why this is happening in three L1s which are fundamentally different.

The HDLF range is composed of 17 MBUs (Table 125), the LDMF range consists of nine subordinators (Table 131) and the LDHF range is made up of four indicators of

subordination (Table 137). There is an obvious trend in which there are over three times more subordinators at the lowest frequency level than the highest. Moreover, there are twice as many subordinators at the medium frequency level than the high frequency level. Results indicate that as frequency increases, L2 Spanish writers use a greatly reduced number of subordinators, yet they use them at a higher frequency.

As the same trend was seen in data extracted from the EFCAMDAT2 (Table 136), the reasons behind high frequency levels yet few MBUs in the CEDEL2/CAES may be due to an elevated number of developmental formulaic sequences in the L2 as a principal trait in comprehension and production (Boers & Lindstromberg, 2012; Henriksen et al., 2013; Peters, 2014; Schmitt & Carter, 2004). This assertion is based in learners who are viewed as proficient produce a high density of accurate and appropriate L2 phrasal expressions (Boers et al., 2006; Dai & Ding, 2010; Schmitt, 2008; Stengers et al., 2011).

Looking at this phenomenon from a different perspective, Ortega (2009) suggests that beginning L2 learners will opt for easier solutions in L2 written language production, using mostly simpler and high-frequency words. While this could be true in some contexts, this study finds the larger percentage of hypotactic enhancement to not occur at beginning levels. On the contrary, high frequency subordination appears to be more of an indicator in intermediate and advanced L2 learners in data from the EFCAMDAT2 and the CEDEL2/CAES. The only MBU to reach peak frequency at A2 was *when* (21.9) in the EFCAMDAT2. It should be noted that *when* at B1 (14.4) and B2 (14.7) was also observed at an elevated frequency when compared to A1, C1 and C2.

At brief glance, *si* reached its highest frequency at B2 (11.9). Nevertheless, if we look at Table 96, we can see there is not much difference between the primary peak frequency at B2 (11.9) with the secondary peak at C2 (11.8). In the case of the positive condition causal-conditional MBU *si*, it was found that L2 Spanish writers used this MBU at higher levels of proficiency thus not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

If there were a clear sign in this study indicating that subordination peaks at intermediate proficiency levels, it would be *porque* in the CEDEL2/CAES data set. Table 94

shows the peak frequency being reached at B1 (25.4). However, the secondary peak frequency comes into play at B2 (24.8). The obvious plateau for this often-used MBU confirms that hypotaxis is most often used, in this case, at the intermediate proficiency levels. Moreover, frequency of use gradually increases at the beginning levels and decreases at advanced levels, thus confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). While findings in this study generally point to hypotaxis as being an indicator of higher-level proficiency, there might be small number of frequently used simple subordinators which display this trend.

The results for *para* showed opposite trend to those displayed with *porque*. Although the general frequency was relatively high for every proficiency level, double peak frequencies for *para* occurred at C1 (28.3) and C2 (28.3) as seen in Table 95. Results indicate that L2 Spanish writers used the *para* to indicate purpose more often at advanced levels thus suggesting as they gain proficiency, their use of subordination increases. Evidence from data extracted from the CEDEL2/CAES for *para* cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Cuando showed an interesting trend since the highest frequency occurred at B1 (19.7), yet the overall trend continued through B2 (18.8) and C1 (19.1). Table 89 shows the pattern for subordination as it happens at intermediate and lower advanced levels. Furthermore, the temporal *cuando* has a generally elevated frequency range across all proficiency levels when compared to other MBUs across all MBCs in data from the CEDEL2/CAES. We cannot conclude this trend only happens at the intermediate level as it continues through C1, therefore the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) does not apply to *cuando*.

Table 138 gives a summary of how each LDHF MBU in the CEDEL2/CAES is classified under CEFR proficiency levels. Superficially, it appears as if the higher numbers of MBUs are located at the intermediate levels of B1 (*porque* and *cuando*) and B2 (*si*). However, if the fact that the peak frequency plateau for *cuando* starts at B1 (19.7) and continues through B2 (18.8) and C1 (19.1), the perspective on MBU frequency and CEFR

proficiency levels begins to change. Moreover, *if* reaches a primary peak B2 (11.9) with the secondary peak at C2 (11.8). If we consider that *para* has a double peak at C1 (28.3) and C2 (28.3), the way of looking at the subordination frequency versus proficiency level distribution changes even more.

A1	A2	B1	B2	C1	C2
n/a	n/a	<i>porque</i> <i>cuando</i>	<i>si</i>	<i>para</i>	<i>para</i>

Table 138 CEDEL2/CAES Hypotaxis and Enhancement Low Density High Frequency MBUs per CEFR Level.

Table 139 shows how CEDEL2/CAES MBUs span over frequency levels and CEFR proficiency levels. The HDLF category is composed of twenty MBUs. As we move onto LDMF MBUs, we can see the number is greatly reduced with a total of nine. LDHF MBUs decrease resulting in only five subordinators, yet it is important to note that *para* reaches its peak frequency at C1 and C2 thus constituting five occurrences. The majority of subordinators for hypotaxis and enhancement taken from the CEDEL2/CAES occurred in large numbers at a low frequency (0.1-0.9). As frequency of use increased, the amount of MBUs decreased leaving us with only five subordinators at the highest level.

45.0-49.9						
30.0-39.9						
20.0-29.9			1		1	1
10.0-19.9			1	1		
9.0-9.9						
8.0-8.9						
7.0-7.9						
6.0-6.9				1		
5.0-5.9						2
4.0-4.9						
3.0-3.9						1
2.0-2.9						1
1.0-1.9	1			2		1
0.1-0.9	4	1	1	1	5	8
	A1	A2	B1	B2	C1	C2

Table 139 CEDEL2/CAES MBUs per Frequency and CEFR Level.

A different way to look at the matter is through the focal lens of the CEFR proficiency levels. The amount of peak frequency MBUs vary at every level in the CEDEL2/CAES with there being five at A1, one at A2, three at B1, five at B2, six at C1 and fourteen at C2. The highest density of subordination is used at advanced levels. More specifically, the greatest number of MBU density occurred at C2. Twenty of the thirty-four MBUs in this study have a peak frequency between C1 and C2. This study suggests that use of MBUs in the category of hypotaxis and enhancement is a phenomenon which takes place at a high level of proficiency. In other words, adverbial subordination increases as proficiency increases but not overall frequency.

As we begin to bring the discussion for hypotaxis and enhancement to a close, it is worth cross-analyzing Table 136 and Table 139 to see how MBUs are distributed through frequency levels and CEFR proficiency levels. Considering there are MBUs that have a peak frequency at two proficiency levels, EFCAMDAT2 data shows that there are thirty-one frequency peaks with twenty occurring at advanced levels. Results from the CEDEL2/CAES show a total of thirty-four frequency peaks with twenty taking place at advanced levels. The overall conclusion for L2 Spanish and L2 English writers is that they use hypotaxis and enhancement at peak frequencies at advanced levels, thus suggesting proficiency is an indicator of acquisition of subordination.

Taking a point of view from the bottom up in Table 140, we can see there are many subordinators at the 0.1-0.9 HDLF category in both the EFCAMDAT2 and CEDEL2/CAES. As we move to LDMF subordinators, peak frequencies occur much less often throughout the nine ranges. However, when looking at Table 140 in contrast with Table 136 and Table 139, it is clear that English L2 writers use a higher medium frequency of subordination than L2 Spanish writers. High frequency MBUs compose the lowest density of MBUs out of the three categories for both data sets with three MBUs appearing in the EFCAMDAT2 and four in the CEDEL2/CAES.

	EFCAMDAT	CEDEL2/CAES
10.0-49.9	3	4
1.0-9.9	14	11
0.1-0.9	14	20

Table 140 Frequency Summary for EFCAMDAT2 & CEDEL2/CAES.

While there were three MBUs at the highest frequency in the EFCAMDAT2 and four from CEDEL2/CAES data, it is important to look at exactly which fine-grained measures were used. Table 134 and Table 137 show out of the four MBUs used at a high frequency for both L2 Spanish and English writers, *cuando/when*, *porque/because* and *si/if* all had parallel translations which demonstrated that L1 Spanish and English writers resort to the same linguistic resources at higher frequencies in L2 writing thus posing a strong case for L1 to L2 transfer in both languages.

4.6.2.3 Hypotaxis and Elaboration

The final category in hypotaxis is elaboration which is composed of five separate MBUs. Three MBUs (*who*, *when* and *where*) in EDFAMDAT2 data fall into the HDLF range of 0.1-0.9. *That* (1.0-1.9) and *which* (3.0-3.9) both occurred at their highest frequencies in the LDMF range (Table 141).

High Density Low Frequency	Low Density Medium Frequency		
0.1-0.9	1.0-1.9	2.0-2.9	3.0-3.9
<i>who</i> (B2)	<i>that</i> (C1)		<i>which</i> (C2)
<i>when</i> (B1)		n/a	
<i>where</i> (B1/B2)			

Table 141 EFCAMDAT2 Hypotaxis and Elaboration MBU Peak Frequency.

When we look at the representative sample for hypotaxis and elaboration, there are fewer measures than extension (7) and even a larger gap than in enhancement (45). However, the overall tendency was for a relatively larger grouping of MBUs in the HDLF range. The same trend appeared for extension (Table 116) and enhancement (Table 122) in which the majority of MBUs occurred at a high density yet a lower frequency in data extracted from the EFCAMDAT2.

Table 142 shows the layout of how MBUs are distributed throughout hypotaxis and elaboration. *Where* has frequency peaks at both B1 (0.1) and B2 (0.1). *When* peaks at B1 (0.6) while *who* apexes at B2 (0.7). There are two MBUs which peak at the advanced level. *That* has a peak frequency at C1 (1.8), and *which* has a peak frequency at C2 (3.3). It should be highlighted that all MBUs at the B1 and B2 level are classified under the HDLF range. In

contrast, *that* and *which* peaked at the advanced level and were classified in the LDMF range. What is apparent is that there are no instances of peak use in elaboration at the beginning levels thus suggesting that L2 English writers use non-defining relative clauses at intermediate and advanced levels.

A1	A2	B1	B2	C1	C2
n/a	n/a	<i>when</i> <i>where</i>	<i>who</i> <i>where</i>	<i>that</i>	<i>which</i>

Table 142 EFCAMDAT2 Hypotaxis and Elaboration MBU per CEFR Level.

The largest grouping of subordinators appeared at the intermediate level thus suggesting subordination might peak between B1 and B2. However, when we consider that MBUs at the intermediate level all occur at the frequency between 0.1-0.9, the storyline changes. Figure 17 shows how high density low frequency MBUs have little effect on the overall trend. Subordinators have more of a frequency impact in advanced levels thus demonstrating that hypotaxis and elaboration peaks at C1 and decreases at C2 in EFCAMDAT2 data. This indicates that hypotaxis in elaboration is not a trend that is consistent with the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Looking at Figure 17 for results of hypotaxis and elaboration, it can be observed that there is a gradual increase from A1 to C1 followed by a decrease at C2. Results mirror those from the only other known study using EFCAMDAT2 data in which learners from four separate L1 show the same increase in relative clause use from A1 to C1 and a decrease at C2 (Chen et al., 2021).

Results from hypotaxis and elaboration from the CEDEL2/CAES (Table 143) showed the same pattern as with the EFCAMDAT2 in the sense that the majority of individual subordinators occurred in the HDLF range of 0.1-0.9 (*quien*, *cuando* and *donde*). The LDMF range was composed of two MBUs *que* with the meaning *that* in a range of 1.0-1.9 and *que* with the meaning of *which* in 3.0-3.9.

High Density Low Frequency	Low Density Medium Frequency		
0.1-0.9	1.0-1.9	2.0-2.9	3.0-3.9
<i>quien</i> (C1)	<i>que</i> (that) (C2)		<i>que</i> (which) (B1)
<i>cuando</i> (C2)		n/a	
<i>donde</i> (B2/C1)			

Table 143 CEDEL2/CAES Hypotaxis and Elaboration MBU Peak Frequency Per 5k.

The relative majority of non-defining relative clauses occurred in the HDLF range in the CEDEL2/CAES. The same trend appeared in extension (Table 119) and enhancement (Table 125) thus solidifying a pattern in hypotaxis in which the majority of MBUs are in HDLF. Additionally, results from elaboration, extension and enhancement demonstrate that as frequency increases, the amount of MBUs decreases. This pattern was present in findings from the CEDEL2/CAES and the EFCAMDAT2.

Peak frequency in hypotaxis and elaboration occurs more at advanced levels (Table 144) than beginning levels. Taking this into account, the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) cannot be confirmed in the CEDEL2/CAES for hypotaxis through elaboration.

Cuando, *quien*, *donde* and *que* (that) all appear at advanced C1 and C2 proficiency levels. *Que* (which) and *donde* were the only two MBUs to occur at intermediate levels. When looking at MBU density, it would appear that elaboration happens at advanced levels. However, Figure 17 shows that Level 2 elaboration has a primary peak at B1 and a secondary peak at C2. The explanation behind this pattern is the primary peak of *que* (which) is at B1 (3.0) and the secondary peak is at C2 (2.2) which influences the overall frequency pattern. Furthermore, the same trend occurs with *que* (that) since the primary peak appears at C2 (1.2) and the secondary peak at B1(1.1).

A1	A2	B1	B2	C1	C2
		<i>que</i> (which)	<i>donde</i>	<i>quien</i>	<i>cuando</i>
n/a	n/a			<i>donde</i>	<i>que</i> (that)

Table 144 CEDEL2/CAES Hypotaxis and Elaboration MBU Per CEFR Level.

Looking at primary and secondary peak frequency ranges in Figure 17 for Level 2 and separate MBUs in Table 99, it can be seen that that two non-defining-relative clause indicators drive the trend towards peaks at B1 and C2. *Que* (which) peaks at B1 (3.0) and has a steep decrease at B2 (1.1) but shows a gradual upward trend at C1 (1.9) and C2 (2.2). The issue under investigation is why is there such a big drop at B2 with a rebound at advanced levels. *Que* (that) shows a primary peak at C2 (1.2) and a secondary peak at B1 (1.1) with lower frequencies at B2 and C1. Formulaic sequencing may be a factor in peak frequency with either marker of elaboration. More research into non-defining relative clauses may shed light on this phenomenon.

4.6.2.4 Parataxis and Extension

When looking at the frequency ranges for paratactic extension in the EFCAMDAT2 (Table 145), there is one coordinator which had zero frequency (*and yet*). There are two HDLF coordinators (*neither ... nor* and *except*) in the range of 0.1 to 0.9. Two LDMF MBUs appear at ranges 1.0-1.9 (*not only ... but also*) and 5.0-5.9 (*but* variative). LDHF coordinators contained three MBUs spanning three frequency ranges. The alternative *or* reached peak frequency between 10.0-19.9 and the adversative *but* came in between 20.0-29.9. The single highest frequency MBU in the EFCAMDAT2 data set was *and* which had a peak frequency between 90.0-99.9.

Zero Frequency	High Density Low Frequency	Low Density Medium Frequency		Low Density High Frequency		
0.0	0.1-0.9	1.0-1.9	5.0-5.9	10.0-19.9	20.0-29.9	90.0-99.9
<i>and yet</i>	<i>neither ... nor</i> (A1, A2, B1 & C1) <i>except</i> (A1 & A2)	<i>not only ... but also</i> (C2)	<i>but</i> (variation) (C2)	<i>or</i> (C2)	<i>but</i> (adversative) (A1)	<i>and</i> (B1)

Table 145 EFCAMDAT2 Parataxis and Extension MBU Peak Frequency per 5k.

The separate frequency ranges in Level 1 taxis (Table 79) and Level 2 expansion categories (Table 80) show that extension is the main driver of the overall category of parataxis in the EFCAMDAT2. While extension has three MBCs (Table 83), the MBSC of positive addition (Table 100) is clearly the backbone for the overall trend in extension and of parataxis in data from the EFCAMDAT2.

The frequency peak for positive addition in the EFCAMDAT2 appears at B2 and declines until C2. Findings from Zarco-Tejada et al. (2016) are consistent with the current study on two separate points. Firstly, the positive additive category of coordinators had the highest frequency in both studies. Secondly, although the previous study investigated use at the A2, B1 and B2 levels, findings showed a frequency increase from A2 to B1 and a decrease from B1 to B2 which coincided with the trend seen in the positive addition MBSC in the current study (Table 83).

EFCAMDAT2 frequency patterns for the three highest MBUs in extension were similar to the Zarco-Tejada et al. (2016) study. *And* (Table 100) and the adversative *but* (Table 102) saw an increase from A2 to B1 followed by a decrease at the B2 level. The variative conjunction *or* (Table 104) had a frequency peak at A2 in either study then a frequency decrease at B2. It would be interesting to see how the results would have been influenced if Zarco-Tejada et al. (2016) had incorporated A1, C1 and C2 in their study.

Unlike the Zarco-Tejada et al. (2016) study, the MBF in this study included CEFR levels A1 through C2. In the case of *and*, there is a continual decrease from B1 (95.7) to C2 (74.2). Despite *but* having a peak frequency at A1 (23.9), there is a decrease at A2 (17.4) followed by secondary increase at B1 (21.7). *But* experienced a continuous decrease at B2 (17.2), C1 (12.4). and C2 (9.2). *Or* demonstrates a decrease from A2 (7.9) to B2 (4.2), yet there is a significant increase from B2 (4.2) to C2 (11.7).

When looking at the role of MBUs in parataxis and extension in the EFCAMDAT2, various studies stand out as having parallel results even when the L1 is not the same. Findings from the Lahuerta Martínez (2018b) study at the A2 and B1 proficiency levels coincided with the results of the current study in which the additive *and* and the adversative *but* were the most frequently used coordinator in each or their respective categories. This may be a result of the overuse of the additive *and* and the adversative *but* since it has been cited as a common occurrence among various L1 learners writing in L2 English (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012; Yang et al., 2017) and might be the reason for such high frequencies.

Table 146 shows the spread of parataxis and extension MBUs across CEFR levels. The three main driver of high frequency are located at different proficiency levels. *But* achieves peak frequency at A1, *and* reaches peak frequency at B1, and finally *or* reaches peak frequency at C2. Although they appear in the 0.1-0.9 frequency range, *neither ... nor* is used at its peak frequency at A1, A2, B1 and C1. *Except* additionally reaches peak frequency at A1 and A2.

A1	A2	B1	B2	C1	C2
<i>neither ... nor</i>	<i>neither ... nor</i>	<i>neither ... nor</i>		<i>neither ... nor</i>	<i>not only ... but</i>
					<i>also</i>
<i>except</i>	<i>except</i>	<i>and</i>	n/a		<i>but</i> (variation)
<i>but</i>					<i>or</i>
(adversative)					

Table 146 EFCAMDAT2 Parataxis and Extension MBU per CEFR Level.

The grouping of MBUs at different CEFR levels (Table 146) gives no apparent pattern as to how they indicate a frequency pattern. Nonetheless, if we consider frequency trends in Table 145, the obvious peak frequency of parataxis and extension appears at B1 and is driven by *and*. This occurrence falls in line with other studies in which markers of extension, especially additive, are reported as having their highest frequency in the lower intermediate range (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012; Yang et al., 2017; Zarco-Tejada et al., 2016).

This study concludes that parataxis, especially extension, in the EFCAMDAT2 is not a phenomenon which has a peak frequency at beginner proficiency levels and decreases at intermediate levels (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012). Instead, the current study proposes that extension, primarily driven by the MBU *and*, reached highest frequency at B1 and is not a phenomenon indicative of beginner learners thus not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

The CEDEL2/CAES (Table 147) displayed a different frequency pattern than the EFCAMDAT2. MBUs were scattered among a variety of frequency ranges with there being

three coordinators in the HDLF range from 0.1-0.9 (*no solo ... sino también, y aun así* and *excepto*). Three MBUs are also categorized in the LDMF range including variative *pero* (1.0-1.9), *ni ... ni* (3.0-3.9) and *o* (7.0-7.9). There were two LDHF MBUs with the adversative *pero* reaching a peak between 20.0-29.9. The overall highest single MBU frequency in the entire study occurred with *y* which had a peak frequency between 110.0-119.0.

High Density Low Frequency	Low Density Medium Frequency			Low Density High Frequency	
0.1-0.9	1.0-1.9	3.0-3.9	7.0-7.9	20.0-29.9	110.0-119.9
<i>no solo ... sino también</i> (C2)	<i>pero</i> (variation) (B1)	<i>ni ... ni</i> (C1)	<i>o</i> (A1)	<i>pero</i> (adversative) (B1)	<i>y</i> (A1)
<i>y aun así</i> (B2, C1 & C2)					
<i>excepto</i> (A2, B2 & C1)					

Table 147 CEDEL2/CAES Parataxis and Extension MBU Peak Frequency Per 5k.

Comparing Level 1 parataxis (Table 79) and Level 2 (Table 80) expansion categories, we can see that extension has a significant influence on parataxis in the CEDEL2/CAES. This pattern mirrors the results in the EFCAMDAT2 and confirms the study conducted by Zarco-Tejada et al. (2016) in which MBUs in paratactic extension contained a much higher use frequency than enhancement and elaboration.

Zarco-Tejada et al. (2016) found when looking at proficiency levels A2, B1 and B2 that L2 writers used paratactic addition at a peak frequency at the B1 level. The current study observed that frequency in the CEDEL2/CAES MBSC of addition was driven by the single MBU *y*. Looking more in-depth into the matter, *y* reaches a frequency peak at A1 (116.6) with there being a secondary peak at B1 (116.4) as seen in Table 100. If data from the A1 level were removed, this study would mirror that of Zarco-Tejada et al. (2016). Nevertheless, results cannot confirm the previous studies findings and would beg the question as to why they excluded A1 from their study.

In line with results from the EFCAMDAT2 with *and* and the adversative *but*, the two MBUs with the highest frequency in the CEDEL2/CAES were the equivalent *y* and the

adversative *pero*. Previous research shows that the English equivalents to the previously mentioned Spanish MBUs are commonly overused in L2 writing (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012; Yang et al., 2017). L1 to L2 transfer may be an explanation as to why use frequency is so high with *y* and *pero* in the CEDEL2/CAES. The pattern shown by the adversative and variative *pero* matched findings from Zarco-Tejada et al. (2016) in which their peak frequency was at B1 followed by a continual decrease.

Table 148 shows how CEDEL2/CAES MBUs for parataxis and extension are distributed throughout CEFR levels. There appears to be a number of coordinators throughout proficiency levels. HDLF MBUs such as *excepto* and *y aun así* appear at several proficiency levels while *no solo ... sino también* appears in one. However, they do not have much of an effect on the overall frequency trend. LDMF MBUs are distributed in three proficiency levels with the variative *pero* in B1, *o* in A1 and *ni ... ni* in C1.

A1	A2	B1	B2	C1	C2
<i>y</i>	<i>excepto</i>	<i>pero</i> (adversative)	<i>y aun así</i>	<i>y aun así</i>	<i>no solo ...</i> <i>sino también</i>
<i>o</i>		<i>pero</i> (variation)	<i>excepto</i>	<i>excepto</i>	<i>ni ... ni</i> <i>y aun así</i>

Table 148 CEDEL2/CAES Parataxis and Extension MBU Peak Frequency per CEFR Level.

Of the two LDHF MBUs, *y* reaches peak frequency at A1 and the adversative *pero* at B1. Reiterating what has previously been mentioned, *y* has a primary frequency peak at A1 with 116.6, yet a secondary peak occurs at B1 116.4 thus appearing to solidify the importance of coordination at A1 and B1. If we cross-reference Table 100 and Table 102, we can see there is not much of a deviation in frequencies between either *pero* or *y*. In the case of *y*, peak frequency is achieved at 116.6 (A1) and the lowest frequency is a 106.4 (A2), yet the frequency does not much vary in B1 (116.4), B2 (114.6), C1 (115.7) and C2 (112.9). The adversative *pero* follows the same trend with the peak frequency at B1 (26.3) and not much variation through A1 (20.9), A2 (20.0), B2 (24.6), C1 (25.0) and C2 (22.7).

The DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) in which

coordination increases during beginner stages and wanes as proficiency advances (Bardovi-Harlig, 1992; Homburg, 1984; Ishikawa, 1995; Sharma, 1980; Vyatkina, 2012) appears to not be applicable to Spanish L2 writers using parataxis and extension in the CEDEL2/CAES.

4.6.2.5 Parataxis and Enhancement

Enhancement is the second expansion category in the general grouping of parataxis. MBUs are grouped into zero-use, HDLF and LDMF categories (Table 149). Comparing results from EFCAMDAT2 data between extension and enhancement, although enhancement has more MBUs, the use frequency is much lower than extension. Looking at the frequency patterns for the eight MBUs for parataxis and enhancement (Figure 19) in relation to the overall category of parataxis (Figure 14), enhancement appears to have a very little impact on parataxis as a whole.

Zero Frequency	High Density Low Frequency	Low Density Medium Frequency	
0.0	0.1-0.9	1.0-1.9	4.0-4.9
<i>and meanwhile</i>	<i>and before that</i> (A2/B2)	<i>and then</i> (A2)	<i>and there</i> (A1)
<i>and afterwards</i>	<i>but before that</i> (A2/B1/C1)		
<i>and in that case</i>	<i>and in that way</i> (A1)		
<i>and yet</i>	<i>thus</i> (B2)		
	<i>and so</i> (A2/B1/B2/C1)		
	<i>but nevertheless</i> (B2/C1)		

Table 149 EFCAMDAT2 Parataxis and Enhancement MBU Peak Frequency per 5k.

Four MBUs (*and meanwhile*, *and afterwards*, *and in that case* and *and yet*) had zero frequency showing they were foreign to L2 Spanish writers. The largest grouping of MBUs occurred between 0.1-0.9 in HDLF (*and before that*, *but before that*, *and in that way*, *thus*, *and so* and *but nevertheless*). The final two MBUs in this category appeared in LDMF range between 1.0-1.9 (*and then*) and 4.0-4.9 (*and there*).

When we analyze paratactic temporal, means, comparison, cause/reason, positive condition and concessive condition MBCs, findings from the present study deviate quite

noticeably from previous studies. The Wenhui Xuan (2019) study as well as Rasool and Mahmood (2023) study found causal-conditional MBUs had highest use frequency. However, the current study saw very low use in the causal-conditional MBC but saw higher frequency in temporal coordinators. This may be due to difference in L1s, or the larger data set used in the current study.

The temporal *and there* is the single MBU at the A1 level (Table 150) that is paramount to the lower beginner level having the highest frequency (Figure 19) in paratactic enhancement. The distribution of MBUs between A2 to C1 account for an overall stable yet descending frequency trend. With there being no MBUs at the C2 level and the prime frequency indicator at the A1 level, it appears the eight MBUs which make up this category show that parataxis and enhancement is a tool used by beginner L2 English writers and decreases as their level increases thus confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

A1	A2	B1	B2	C1	C2
<i>and in that way</i>	<i>and before that</i>	<i>but before that</i>	<i>and before that</i>	<i>but before that</i>	
<i>and there</i>	<i>but before that</i>	<i>and so</i>	<i>thus</i>	<i>and so</i>	
	<i>and so</i>		<i>and so</i>	<i>but</i>	n/a
				<i>nevertheless</i>	
	<i>and then</i>		<i>but</i>		
			<i>nevertheless</i>		

Table 150 EFCAMDAT2 Parataxis and Enhancement MBU Peak Frequency per CEFR Level.

A reason for the relatively high peak frequency at A1 might be task-based formulaic sequencing (Boers & Lindstromberg, 2012; Henriksen et al., 2013; Peters, 2014; Schmitt & Carter, 2004) since there is a possibility that an assignment required learners at the A1 level to use the temporal *and there*. Looking at the peak frequency at A1 from a different perspective, if we removed *and there* from the equation, the overall frequency for parataxis and extension would be very low with the highest peak frequency being between 1.0-1.9. Further research is needed with a larger data set and additional indicators to add to the knowledge base around parataxis and enhancement.

Results from the CEDEL2/CAES (Table 151) shows a similar pattern as seen in the EFCAMDAT2. The zero frequency range contained four separate MBUs (*y mientras tanto*, *y de esa manera*, *y aun así* and *pero aun así*). The largest grouping of MBUs occurred in the HDLF range with seven MBUs (*y luego*, *y después*, *y antes*, *pero antes*, *y allí*, *y por lo tanto* and *y en este caso*). The LDMF contained one MBU (*así*) at 1.0-1.9.

Zero Frequency 0.0	High Density Low Frequency 0.1-0.9	Low Density Medium Frequency 1.0-1.9
<i>y mientras tanto</i>	<i>y luego</i> (C1/C2)	<i>así</i> (C2)
<i>y de esa manera</i>	<i>y después</i> (C1/C2)	
<i>y aun así</i>	<i>y antes</i> (A2/B1/B2/C1/C2)	
<i>pero aun así</i>	<i>pero antes</i> (B1/C1)	
	<i>y allí</i> (A1/C2)	
	<i>y por lo tanto</i> (B2/C1/C2)	
	<i>y en este caso</i> (B2)	

Table 151 CEDEL2/CAES Parataxis and Enhancement MBU Peak Frequency per 5k.

Looking at the overall frequency pattern of parataxis and enhancement in the CEDEL2/CEDEL (Figure 19), we can see that MBU use starts at 0.9 at the A1 level and displays an ascending tendency until C2 with an ending frequency at 3.6 (Table 80). This occurrence is mainly due to the distribution of peak frequency MBUs being fewer at lower proficiency levels and increasing in number as they reach the advanced levels (Table 152).

A1	A2	B1	B2	C1	C2
<i>y allí</i>	<i>y antes</i>	<i>y antes</i>	<i>y antes</i>	<i>y luego</i>	<i>y luego</i>
		<i>pero antes</i>	<i>y por lo tanto</i>	<i>y después</i>	<i>y después</i>
			<i>y en este caso</i>	<i>y antes</i>	<i>y antes</i>
				<i>pero antes</i>	<i>y allí</i>
				<i>y por lo tanto</i>	<i>y por lo tanto</i>
					<i>así</i>

Table 152 CEDEL2/CAES Parataxis and Enhancement MBU Peak Frequency per CEFR Level.

Various studies report different findings in terms of the frequency of coordination across CEFR proficiency levels with the majority being contradictory. Bardovi-Harlig (1992) saw a decrease in coordination as proficiency increased. There have been various studies which have found no significant difference in the frequency of coordination across proficiency levels (Ai & Lu, 2013; Lu, 2011) or when looking more in-depth between intermediate and advanced L2 learners (Neary-Sundquist, 2016). It must be considered that

previous studies investigated a broader range of coordination. Nevertheless, findings from the CEDEL2/CAES with parataxis and enhancement contradict findings from previous studies in the sense that this type of coordination increased with proficiency (Figure 19).

The current study found coordination through parataxis and enhancement in the CEDEL2/CAES to be a distinguishing factor among proficiency levels. Contrary to previous studies, frequency use tended to increase as the L2 learners' proficiency level increased thus distinguishing beginner, intermediate and advanced levels. This study cannot confirm that coordination through parataxis and enhancement decreases as proficiency increases in L2 Spanish texts (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012; Yang et al., 2017) and it cannot confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

When looking at MBCs, CEDEL2/CAES data cannot confirm findings from the Wenhui Xuan (2019) study in which zero frequency was found in temporal enhancement. In contrast, the current study found parataxis through temporal coordinators to be the most often used MBSC. CEDEL2/CAES results are also inconsistent with the Rasool and Mahmood (2023) study which found that causal-conditional parataxis was categorically used at the highest MBC while temporal indicators were used at the second highest frequency by L2 learners.

Table 149 and Table 152 compare frequency rates between the EFCAMDAT2 and CEDEL2/CAES. L2 English writers used parataxis and enhancement with a higher frequency than L2 Spanish writers due to the elevated use of *and there* between the range of 4.0-4.2. On the other hand, when we look at how MBUs are distributed throughout CEFR levels, an opposite trend appears. *And there* at the A1 level establishes a pattern in which L2 English writers use a higher frequency of MBUs at lower levels. The opposite pattern emerged in the results from the CEDEL2/CAES in which L2 Spanish writers used more indicators of parataxis and enhancement as their level progressed. This is mainly due to a larger distribution of MBUs over proficiency levels and multiple frequency peaks.

4.6.2.6 Parataxis and Elaboration

The final category in parataxis involves elaboration. There are a total of thirteen MBUs distributed across the MBSC of exposition, exemplification and clarification. Table 153 shows the largest amount of EFCAMDAT2 MBUs were found in the HDLF category (*in other words, that is to say, I mean, in particular, what and indeed*). The remainder of MBUs were distributed across four LDMF ranges. There were two MBUs in the range of 2.0-2.9 (*for example and in fact*), one in 3.0-3.9 (*actually*), three in 4.0-4.9 (*such as, when and at least*) and one in 5.0-5.9 (*how*).

High Density		Low Density Medium Frequency			
Low Frequency					
0.1-0.9	1.0-1.9	2.0-2.9	3.0-3.9	4.0-4.9	5.0-5.9
<i>in other words</i>		<i>for example</i>	<i>actually</i>	<i>such as</i>	<i>how</i>
(C1)		(B1)	(C2)	(C2)	(C2)
<i>that is to say</i>		<i>in fact</i>		<i>when</i>	
(C1)		(C2)		(B2)	
<i>I mean</i>				<i>at least</i>	
(C2)				(C2)	
<i>in particular</i>	n/a				
(C1)					
<i>what</i>					
(B2)					
<i>indeed</i>					
(B2/C1)					

Table 153 EFCAMDAT2 Parataxis and Elaboration MBU Peak Frequency Per 5k.

The frequency ranges found in parataxis and elaboration (Figure 20) were much lower than those of extension (Figure 18) yet they were higher than enhancement (Figure 19). Table 154 shows how the overall peak frequency curve of MBUs per CEFR level in EFCAMDAT2 data begins at B1 and increases to B2 followed by a plateau at C1 and a sharp increase at C2. However, to correctly decipher the data, we must look at individual MBUs.

A1	A2	B1	B2	C1	C2
		<i>for example</i>	<i>what</i>	<i>in other words</i>	<i>I mean</i>
			<i>indeed</i>	<i>that is to say</i>	<i>in fact</i>
n/a	n/a		<i>when</i>	<i>in particular</i>	<i>actually</i>
				<i>indeed</i>	<i>such as</i>
					<i>at least</i>
					<i>how</i>

Table 154 EFCAMDAT2 Parataxis and Elaboration MBU Peak Frequency per CEFR Level.

Besides *that is to say* with a peak frequency at C1 (0.2), the remaining HDLF MBUs had occurrences throughout various proficiency levels in exposition (Table 113), exemplification (Table 114) and clarification (Table 115) thus adding to the overall augmentation of frequency in parataxis and elaboration. MBUs in the LDMF range had occurrences throughout the six proficiency levels with the tendency for a rising frequency. As proficiency increased, EFCAMDAT2 writers used more paratactic elaboration.

Results from EFCAMDAT2 data confirm findings from the Zarco-Tejada et al. (2016) study looking into parataxis and elaboration at the A2, B1 and B2. Either study found elaborative MBUs occurred at a relatively low frequency. However, frequency increased as learners progressed through proficiency levels and larger number of MBU occurrences appeared with more varied coordinators at B1 when compared to A2.

Table 154 shows paratactic elaboration MBUs across CEFR proficiency levels. The trend appears that peak frequency increases as proficiency rises. If we cross-reference peak frequency with overall frequency in Level 2 elaboration (Figure 20), except for a decrease from B2 to C1, there is an obvious correlation with an increase in frequency as proficiency level rises from A1 to C2. Evidence from data taken from the EFCAMDAT2 suggests that parataxis and elaboration is not a phenomenon which decreases after beginner levels (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012; Yang et al., 2017) thus not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

Findings for elaboration from the CEDEL2/CAES (Table 155) stand out when compared to the EFCAMDAT2 in the sense there is a larger grouping of MBUs in the HDLF category (*en otras palabras, quiero decir, tal como, en particular, cuándo, qué, de hecho, de verdad* and *en efecto*). LDMF MBUs were fewer and appeared at lower frequencies than those of the EFCAMDAT2 with two at the 1.0-1.9 (*es decir* and *por lo menos*) and two at the 2.0-2.9 range (*por ejemplo* and *cómo*).

High Density Low Frequency	Low Density Medium Frequency	
0.1-0.9	1.0-1.9	2.0-2.9
<i>en otras palabras</i> (B2/C1)	<i>es decir</i> (B1)	<i>por ejemplo</i> (C1)
<i>quiero decir</i> (A1/C2)	<i>por lo menos</i> (C2)	<i>cómo</i> (C2)
<i>tal como</i> (A1/B1)		
<i>en particular</i> (B2/C2)		
<i>cuándo</i> (B1/C1)		
<i>qué</i> (C2)		
<i>de hecho</i> (C2)		
<i>de verdad</i> (C1)		
<i>en efecto</i> (C2)		

Table 155 CEDEL2/CAES Parataxis and Elaboration Peak Frequency Per 5K.

Es decir has a single peak frequency at B1 (1.7), yet there is zero-use frequency at A1 and A2 and very low frequencies between 0.1 to 0.9 at B2, C1 and C2. *Por lo menos* has a peak at C2 (1.0) then occurrences between 0.1-0.9 at A2, B2 and C1. In terms of the HDLF category, *en efecto* was the only MBU which materialized in one proficiency level at 0.1 at C2. Of the remaining eight MBUs with frequency between 0.1-0.9, *en otras palabras* appeared in B2 and C1. *En particular* and *de hecho* appeared at B2, C1 and C2. *Tal como* appeared at A1, A2, B1 and C1. *Cuándo* at A2, B1, B2 and C1. *De verdad* at B1, B2, C1 and C2. Finally, *quiero decir* as well as *qué* materialized in all six CEFR levels. The point of detailing the multitude of various MBU occurrences is to firstly show that parataxis and elaboration is a linguistic resource used at many proficiency levels and secondly demonstrate that entire body of MBUs add to increase in categorical frequency.

CEDEL2/CAES findings suggest that learners gradually use more parataxis and elaboration as they progress through proficiency levels A2, B1 and B2 thus confirming

Zarco-Tejada et al. (2016). However, the caveat comes in the sense that Zarco-Tejada et al. (2016) found that learners used more varied conjunctions at a higher frequency at the B2 level. The current study observed a slight increase in frequency from B1 (5.0) to B2 (5.2) in Table 80, yet cannot confirmed that L2 Spanish writers used a larger variety of conjunctions at B2 as there are a plethora of MBUs at many different frequencies distributed throughout CEFR proficiency levels. One reason for the deviance in results is that Zarco-Tejada et al. (2016) utilized a representative sample of 10,000 words per proficiency level while the current study incorporated a significantly larger representative sample.

Table 156 shows the overall spread of how MBUs are categorized by their peak frequency throughout CEFR levels. The distribution of MBUs at peak frequency appears higher at C1 and C2. Bearing in mind the wide range of MBUs spread out over proficiency levels, we must consider how this factor affects the frequency curve. Figure 20 shows how frequency continuously rises from A1 to C2. This trend provides evidence that the use frequency of parataxis and elaboration increases as proficiency increases, therefore not confirming the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998).

A1	A2	B1	B2	C1	C2
<i>quiero decir</i>		<i>tal como</i>	<i>en otras palabras</i>	<i>en otras palabras</i>	<i>quiero decir</i>
<i>tal como</i>		<i>cuándo</i>	<i>en particular</i>	<i>cuándo</i>	<i>en particular</i>
	n/a	<i>es decir</i>		<i>de verdad</i>	<i>qué</i>
				<i>por ejemplo</i>	<i>de hecho</i>
					<i>en efecto</i>
					<i>por lo menos</i>
					<i>cómo</i>

Table 156 CEDEL2/CAES Parataxis and Elaboration Peak Frequency per CEFR Level.

Data sets from the EFCAMDAT2 and the CEDEL2/CAES both show that as use frequency increases, proficiency also increases. Data extracted from the EFCAMDAT2 showed L2 English writers used parataxis and elaboration at a higher frequency range than data from the CEDEL2/CAES. In terms of how learners used different MBUs, both data sets provided evidence that learners used a multitude of conjunctions with frequency increasing as learners became more proficient. An area of future research might include a bigger data set

with more MBUs yet have use frequency classified by the genre. This would show how and when conjunctions are used thus providing information on how teachers might insert parataxis and elaboration into second language classrooms to further L2 learning.

4.6.3 Conclusion

This section brings to a close the final quantitative analysis in the current study incorporating fine-grained MBUs. Firstly, we set out to test research question two to see if DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) can be confirmed through MBUs. Secondly, to address research question three, we investigate MBU patterns across CEFR proficiency levels to achieve an understanding of how individual units are used. Lastly, we delve into research question four by seeing which MBUs are being used and if frequency patterns can distinguish L2 acquisition as well as if L1 to L2 transfer may be taking place.

Findings from the current study for research question two indicate the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) is not consistent with the learner data from the EFCAMDAT2 and CEDEL2/CAES in hypotaxis through extension, enhancement and elaboration. However, certain MBUs did subscribe to the notion that subordination increases at the intermediate levels only to decrease at advanced levels. Three cases appeared in CEDEL2/CAES data in which *antes de que*, *donde* and *porque* peaked at intermediate levels. Although *porque* peaked at intermediate levels, frequency continued to be elevated at A1-A2 and C1-C2.

MBUs in the categories of parataxis through extension and elaboration produced results in the EFCAMDAT2 and the CEDEL2/CAES which were unable to confirm the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). Data from the EFCAMDAT2 could confirm the beforementioned hypothesis in paratactic enhancement, mainly due to the high use of *and there* at A1 which was thought to be the result of task-based formulaic sequencing (Alexopoulou et al., 2015; Chen et al., 2021; Alhassan & Wood,

2015; Lewis, 1997; Wray, 2002). Findings from the CEDEL2/CAES showed a different pattern with frequency increasing as learners achieved higher proficiency levels.

Research question three investigated MBU use across CEFR proficiency levels. Hypotaxis through extension, enhancement and elaboration in the EFCAMDAT2 and the CEDEL2/CAES saw the majority of MBUs reaching peak frequency at C1 or C2. However, due to the low representative sample at the C2 level in the EFCAMDAT2, which was also a factor in findings (Table 12) from the Chen et al. (2021) study, this study suggests the significant frequency decrease at the upper advanced level is due to a small representative sample. Results indicate that hypotactic subordination is employed in L2 Spanish and English writing more often as proficiency increases through advanced levels.

What stood out in hypotaxis and extension in the EFCAMDAT2 data is that five MBUs (*whereas, besides, apart from, instead of* and *without*) out of seven showed peak frequency at advanced levels. In contrast, the frequency driver of hypotaxis and extension in the CEDEL2/CAES consisted of one MBU (*si no*) which was the major factor in the overall frequency peak at C2 as well as the frequency booster in the extension MBC.

Hypotaxis and enhancement showed the majority of MBUs having a peak frequency at advanced levels in both the EFCAMDAT2 and the CEDEL2/CAES. Although individual MBUs displayed different frequency patterns, one of the main conclusions of this study is that L2 Spanish and English writers use a higher frequency of hypotaxis through enhancement with increasingly varied subordinators as they progress through CEFR levels. This finding indicates proficiency is a marker of language acquisition through adverbial subordination.

Several subordinators in both data sets should be highlighted since they were the main drivers of hypotaxis and enhancement in their respective categories. Firstly, same time temporal *when* and the equivalent *cuando* should be underscored as learners used either MBU as a constant linguistic resource from A2 to B2 in the EFCAMDAT2 and A2 to C2 in the CEDEL2/CAES. It must be noted that *when* had a peak frequency at A2 which is suggested to be the consequence of task-based formulaic sequencing (Alexopoulou et al., 2015; Chen et al., 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002) while elevated frequencies at B1 and B2 are thought to be due to L1 to L2 transfer.

A second phenomenon to call attention to was that of the causal-conditional cause through reason MBUs *because* and *porque*. Both MBUs were consistently used by L2 learners at relatively high frequencies within the parameters of both data sets between A2 to C1. In the case of *because* and *porque*, there are two possible explanations which may exist. Firstly, L1 to L2 transfer might be taking place in which learners are employing L1-based cognitive resources when needed in L2 learning (Cummins, 2008; Forbes, 2019; Siegel, 2003; Yan, 2010). Secondly, developmental formulaic sequencing may be occurring in which a high density of accurate and appropriate MBUs are being used as they are viewed as markers of proficiency (Boers et al., 2006; Dai & Ding, 2010; Schmitt, 2008; Stengers et al., 2011).

The causal-conditional through positive condition *if* in the EFCAMDAT2 and *si* in the CEDEL2/CAES were used at relatively high frequencies which matched the trend with hypotaxis and enhancement reaching peak frequency at advanced levels. In the case of *if*, there was an obvious peak at C1 (45.0) which was the highest frequency reached by any subordinator in the EFCAMDAT2. *Si* showed a different tendency with peak frequency at B2 (11.9) and C2 (11.8). The causal-conditional purpose indicator *para* in the CEDEL2/CAES was the only other singular MBU to be used at a high enough frequency at C1 (28.3) and C2 (28.3) to have an impact on hypotaxis and enhancement.

The two most common MBUs used by L2 learners in hypotaxis through elaboration were *which* and *that* in the EFCAMDAT2 which peaked at advanced levels. The analogous MBUs *que* (which) and *que* (that) were used at a high frequency in the CEDEL2/CAES. L2 Spanish writers employed *que* (which) reaching maximum frequency at B1 (3.0) and *que* (that) demonstrating what appears to be a relative double peak at B1 (1.1) and C2 (1.2). The remaining subordinators in either data set occurred at a low frequency and did little to affect the frequency curve in the Level 2 elaboration category.

Parataxis and extension showed the highest frequencies of categorical use and singular MBU use in the entire study. A commonality observed in both data sets was increased frequencies in three parallel MBUs which included *and/y*, *but/pero* and *or/o*. Coordination through *and* peaked at B1 which is consistent with previous studies looking into positive addition (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012;

Yang et al., 2017; Zarco-Tejada et al., 2016). Coordination remained relatively high with *and* from B1 (95.7) to C1 (85.8) and was the single MBU with the highest frequency in the EFCAMDAT2 data set. Although *and* is found to be commonly overused in English L2 writing (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012; Yang et al., 2017), this study believes that it's consistent use is a product of L1 to L2 transfer.

The positive additive *y* in the CEDEL2/CAES saw the highest frequency at A1 (116.6) of any MBU in this study. Besides having a decrease at A2 (106.4), the difference between peak frequency at A1, then B1 through C2, was only a variation of 3.7 words per 5k which is consistent with previous studies that observed no significant differences between L2 proficiency levels (Lu, 2011; Ai & Lu, 2013) nor significant differences in the rate of coordination between intermediate and advanced L2 learners (Neary-Sundquis, 2016). This study concludes that *y* is a consistent linguistic resource used by L2 Spanish writers, yet it is not a factor in distinguishing proficiency between levels. *And* is the equivalent to *y* and has been found to be commonly overused in English L2 writing (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012; Yang et al., 2017). The current study suggests overuse maybe a factor in such high frequency use of *y*. Nevertheless, consistent use in A1, B1, B2, C1 and C2 indicates L1 to L2 transfer may play a role in augmenting frequency.

The adversative *but* in the EFCAMDAT2 as well as the equivalent *pero* in the CEDEL2/CAES saw distinctive patterns. *But* reached peak frequency at A1 (23.9) and a second peak frequency at B1 (21.7). Frequency with *but* remained consistent from A1 (23.9) to B2 (17.2) with a variation of 6.7 words. The highest frequency for *pero* was observed at B1 (26.3), yet the difference in peak frequency between A1 (20.9) and C2 (22.7) was that of 6.3 words. *But* and *pero* were observed as a consistent linguistic resource used by L2 English and Spanish writers with probable L1 transfer to the L2 taking place.

Frequency for parataxis and enhancement in the EFCAMDAT2 was driven by *and in that way* and *and there* which both saw frequency peaks at A1 with a trend of descending frequency until C2. The peak frequency of *and there* at A1 was interpreted as the probable result of task-based formulaic sequencing (Alexopoulou et al., 2015; Chen et al., 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002). The remaining MBUs exhibited little to no usage. Data from the CEDEL2/CAES saw low-use frequency with no standout MBUs and

the overall trend for a slight increase from the lowest proficiency levels to the highest levels. Parataxis through enhancement was the category least used by learners in either data set.

Parataxis and elaboration in the EFCAMDAT2 showed six of the thirteen MBUs (*I mean, such as, how, in fact, actually and at least*) all demonstrating peak frequency at C2. The MBUs which demonstrated the highest frequencies out of the grouping were *such as, for example, when, in fact, actually and at least*. The two main MBUs frequency drivers in the CEDEL2/CAES were *por ejemplo* and *cómo* which both achieved their highest use rate at C2. The majority of the MBUs had a low frequency and saw minor increases from beginner to advanced levels but augmented overall categorical frequency.

Research question four delved into the subject of researching if frequency trends could be differentiated by the MBF to detect overall patterns of L1 to L2 transfer. There were two trends which should be underlined in hypotaxis and extension. First, EFCAMDAT2 writers used more adversative MBUs. In contrast, the second trend is CEDEL2/CAES writers used alternation at a much higher frequency. In general, data showed L2 Spanish learners used hypotaxis and extension at a higher rate which may suggest the English L1 has a stronger pull towards using this type of subordination.

Hypotaxis and enhancement had a large percentage of zero-use MBUs. EFCAMDAT2 contained a total of fourteen while the CEDEL2/CAES had fifteen MBUs which were not used by learners. The majority of zero-use MBUs occurred in the causal-conditional MBC in both data sets.

The HDLF range in hypotaxis and enhancement showed EFCAMDAT2 learners used a total of 13 MBUs which was a lower quantity when compared to the 17 MBUs used by learners in the CEDEL2/CAES. The temporal and cause reason MBC had the highest number of HDLF MBUs in either data set. Looking at the overall numbers, over half of the MBUs from both data sets were categorized in either zero-use or HDLF categories. As for the zero-use category, these subordinators may require a higher degree of cognitive complexity than learners are able to attain in the L2 (Kortmann, 1999). However, this is not the case for the HDLF as the majority of MBUs occurred at advanced levels which leads this study to believe an alternative phenomenon is taking place.

L2 English writers used almost twice the amount of LDMF MBUs than L2 Spanish writers in hypotaxis and enhancement. An important point to underline is that data from the CEDEL2/CAES showed that learners tended to use fewer subordinators at a lower frequency in the LDMF category. This appears to be the result of a larger quantity of HDLF MBUs used by L2 Spanish writers. In both data sets, a vast array of temporal MBUs were used by L2 learners which showed they may be an important resource for learners in this frequency range.

LDHF MBUs consisted of three MBUs in the EFCAMDAT2 and four in the CEDEL2/CAES for hypotaxis and enhancement. Parallels were seen in the cause reason with *because/porque*, temporal *when/cuando* and the positive condition *if/si*. Bearing this in mind, L2 learners in English and Spanish employ the same subordinators at high frequencies. The only exception to the above three MBUs was *para* for purpose in the CEDEL2/CAES.

A reason for high frequency levels yet few MBUs in both data sets might be due to frequent formulaic sequences in the L2 as a trait in comprehension and production (Boers & Lindstromberg, 2012; Henriksen et al., 2013; Peters, 2014; Schmitt & Carter, 2004). Learners are viewed as proficient when they produce a high concentration of accurate and appropriate L2 phrasal expressions (Boers et al., 2006; Dai & Ding, 2010; Schmitt, 2008; Stengers et al., 2011). While this explanation might hold true, it would also be possible to conclude that *because/porque*, *when /cuando* and *if/si* are cross-linguistic phenomena which indicate probable L1 to L2 transfer bi-directionally across two different L2s.

The frequency sequence seen in the EFCAMDAT2 and CEDEL2/CAES presented a pattern which corresponded in both data sets in which there were elevated numbers of zero-use and HDLF MBUs. Additionally, there were fewer MBUs at the LDMF range and then even fewer in the LDHF range. An important take away is that as frequency increases, L2 learners used fewer subordinators at a higher frequency with a set of preferred MBUs which are largely equivalent. On the other hand, as proficiency increases it was observed that a larger variety of MBUs were used in both the EFCAMDAT2 as well as the CEDEL2 at lower frequencies. An area of future research would be to investigate why this is happening and see if it can be replicated in English and Spanish, as well as with other L2s.

Hypotaxis and elaboration contained fewer MBUs than extension and significantly fewer than enhancement. The majority of non-defining relative clauses occurred in the HDLF range. The two most common MBUs were comparable in either data set with *que/that* and *que/which* indicating a possible cross-linguistic trend of L1 to L2 transfer.

Parataxis and extension saw the highest frequencies in the entire study. The two MBUs with the highest frequency in parataxis and extension in the CEDEL2/CAES and the EFCAMDAT2 were the additive coordinator *and/y* and the adversative *but/pero*. Previous research shows that *and* and *but*, which are the equivalents to *y* and *pero*, are commonly overused in English L2 writing (Bolton et al., 2003; Chen, 2006; Wenhui Xuan, 2019; Xu & Liu, 2012; Yang et al., 2017). Considering the high frequencies of *and/y* and *but/pero*, results from the current study suggest that both previously mentioned coordinators might be overused by L2 writers in the present study as well.

The results for L1 transfer to L2 for parataxis and enhancement would benefit from a larger representative sample and more measures. L2 English writers used parataxis and enhancement at a higher frequency than L2 Spanish writers due to the elevated use of *and there*. Nevertheless, English L2 writers used a higher frequency of MBUs at lower levels which decreased as proficiency increased. The opposite pattern emerged in the results from the CEDEL2/CAES in which L2 Spanish writers used more indicators of parataxis and enhancement as their level progressed. The overall conclusion suggests that if paratactic extension MBUs *and/y* and *but/pero* suffer from overuse, the logical suggestion would be for L2 English and Spanish teachers to put more emphasis on learning parataxis and enhancement MBUs.

The EFCAMDAT2 and the CEDEL2/CAES showed that as frequency increases in MBUs for parataxis and elaboration, proficiency also increased. English L2 writers used elaborative MBUs at a higher frequency range than L2 Spanish learners. Both data sets provided evidence that learners used a multitude of conjunctions at a range of proficiency levels. The general conclusion is that as proficiency increases, so does use of paratactic elaboration. Future research could examine a larger representative sample with more MBUs. Moreover, including the variable of writing genre might provide further insight.

4.6.3.1 The Developmental Sequence of Meaning-Based Interlanguage Complexification Hypothesis for Adverbial Subordination

Few studies have investigated tracking fine-grained measures, such as MBUs, over a series of ranges to measure which types of adverbial subordination are used, how frequently they are used and at what proficiency level they are used. Results from the current study cannot be explained by the DSSICH (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998). Findings can partially explain acquisition through frequency and exposure, L1 to L2 transfer and formulaic sequencing. To further clarify results from L2 Spanish and English texts, this study proposes the developmental sequence of meaning-based interlanguage complexification hypothesis (DSMBICH) for adverbial subordination. Findings from hypotaxis and enhancement lead to the proposal of several key elements which form the foundation of this alternative hypothesis.

Central to cognitive linguistics is that much of a speakers' language represents the accumulation of exposure through usage events which manifests the importance of frequency in shaping and learning a language (Kemmer & Barlow, 2000). Evidence suggests that language processing and acquisition relies on a previous statistical knowledge of the language the learner has encountered, and that frequency is a key determinant of language learning (Ellis, 2008). If we take this approach to explain the lack of frequency for numerous MBUs with zero occurrences, then it would be logical to assert that L2 English and Spanish learners had little or no exposure to these subordinators both in token and types of frequency (Bybee, 2008) with the consequence being an accumulation of zero-use MBUs.

A large number of subordinators in this study were classified at advanced proficiency levels in the HDLF range due to the fact their peak frequency occurred between 0.1-0.9. It is possible that subordinators in this frequency range require a high degree of cognitive complexity which will typically be the output of pragmatic enrichment (Kortmann, 1999). The application of this perspective would mean that HDLF MBUs necessitate a higher cognitive function than a certain percentage of L2 English and Spanish writers are able to fulfill at lower proficiency levels. However, the matter which remains to be explained is why a large grouping of HDLF MBUs were used at intermediate and advanced proficiency levels.

This study proposes that HDLF MBUs are largely a product of frequency and exposure (Ellis, 2008; Kemmer & Barlow, 2000) to the target MBU. The evidence supporting this proposal lies in the vast number of MBUs in the HDLF range which are used at peak frequency in the intermediate to advanced proficiency levels in the EFCAMDAT2 and advanced levels in the CEDEL2/CAES. Assuming that learners have been exposed to the language for a longer duration of time, they have had ample opportunity to gain an understanding of how and when to use MBUs in the HDLF category. The result of this process is exemplified by a low use frequency of a given MBU. Nevertheless, frequency gradually increases through the learning process culminating in a large grouping of subordinators at advanced levels. The implication of this proposal is that learners using LDHF MBUs have operationalized this resource into their linguistic stockpile thus indicating L2 acquisition.

The category of LDMF MBUs is proposed to be the line of demarcation as to when exposure and frequency diminish being a factor in subordination use and when formulaic sequencing and L1 to L2 transfer begin to play a role. EFCAMDAT2 data shows many subordinators reaching peak frequency at the C1 and C2 levels. Data from the CEDEL2/CAES indicates the larger percentage of subordination takes place at C2 with B2 also having an elevated percentage of MBUs.

Findings demonstrate that highly-proficient students' use of medium frequency and high frequency MBUs is not a sign of acquisition or learning, but rather is an indicator of L1 influence on the L2. Evidence from this study suggests the more frequently an MBU is used, the stronger the cognitive connection pulls from known correlations which are often used in the L1 and then applied to the L2. The broad definition of cross-linguistic transfer entails a language learner's use of linguistic knowledge in one of their languages to leverage learning in another language (Yang et al., 2017).

The challenge of clarifying when LDMF MBUs are a product of L1 to L2 transfer takes place when there is a stable and consistent spread of use over sequential CEFR proficiency levels. On a surface level it appears that many medium frequency subordinators peak at advanced levels. However, when looking more profoundly into the matter, we

uncover the trend in which certain MBUs show habitual use by learners in a successive order of proficiency levels.

The DSMBICH introduces the concept in which certain LDMF subordinators exhibit cross-linguistic L1 to L2 transfer. The reason for this behavior is that learners are using L1-based cognitive resources on their own initiative when needed in L2 learning (Cummins, 2008; Forbes, 2019; Siegel, 2003; Yan, 2010). Data from both the EFCAMDAT2 as well as the CEDEL2/CAES suggests that the majority of MBUs which share a stable frequency rate across CEFR proficiency levels in the LDMF category are equivalent in meaning. To illustrate a few examples, *while/mientras*, *after/después de*, *before/antes de*, *where/donde*, *although/aunque* and *como/as* (manner) all show stable frequencies across at least three or more proficiency levels.

In the case of LDMF MBUs, the DSMBICH introduces the notion in which there are a certain number of MBUs that appear without having elongated frequency patterns in a series of proficiency levels before and after peak frequency. Examples from the EFCAMDAT2, such as the temporal *since* and *once*, peak at C2. In this case, very advanced learners can be expected to show knowledge of developmental formulaic sequences which is like that of native speakers (Boers & Lindstromberg, 2012). The DSMBICH suggests that certain subordinators with increasing frequency leading up to a peak at advanced levels is a result of a mix of frequency and exposure which suggests native-like proficiency through developmental formulaic sequencing which is not task-based.

In contrast to the example with *since* and *once*, the LDMF temporal *hasta que* reaches peak frequency at A1 with relatively low occurrence rates from A2 through C2. This study advocates that MBUs which have a singular frequency peak at lower proficiency levels with demonstrably low frequency or zero use across remaining levels tend to be an example of task-based formulaic sequencing in which a learner remembers a particular sequence of words and incorporates them into a communicative task designed to extrapolate a predetermined sequence (Alexopoulou et al., 2015; Chen et al., 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002).

Low density high frequency subordination is a category with rather few MBUs which share a meaning-based commonality. *Because/porque*, *when/cuando* and *if/si* all appear at

relatively high frequencies across consecutive proficiency levels in the EFCAMDAT2 and CEDEL2/CAES. MBUs in this category, although few in quantity yet often used, are what the present study found to be the most commonly used resource in the linguistic stockpile of L2 English and Spanish writers. DSMBICH asserts that LDHF MBUs embellish L1 to L2 transfer as they are implemented at an elevated frequency in writing over a myriad of levels ranging from beginner to advanced.

Task-based and developmental formulaic sequencing can materialize within a high frequency series across consecutive proficiency levels in the shape of a subordinator reaching an elevated peak frequency which is substantially higher than the rest. A case and point of task-based formulaic sequencing involved *when* as it appeared at elevated frequencies from A2 to B2. The spike in use at A2 (21.9) was notable in the fact that frequency was categorically low at A1 (2.9). The intermediate levels involving B1 (14.4) and B2 (14.7) demonstrated a stabilization in frequency which suggest L1 to L2 transfer. The point of this explanation is to show how various factors are involved in the stabilization of high frequency through a series of CEFR levels which can be partially due to formulaic sequencing and L1 and L2 transfer.

L1 writing knowledge, which is linked with L2 writers' ability, is predominant to their conscious actions during L2 writing (Bhowmik, 2016; Rinnert & Kobayashi, 2016; Rinnert et al., 2015). L1 transfer in L2 development is either conscious or intuitive (DePalma & Ringer, 2011). Learners may exploit their L1-based cognitive resources on their own initiative when needed in L2 learning (Cummins, 2008; Forbes, 2019; Siegel, 2003; Yan, 2010). Transfer might be classified as an automatic or procedural process that adheres to habitual modes of thinking in an L1 (Cohen, 2014; Jarvis & Pavlenko, 2008). Whatever the case, L2 writers' agency makes for the dynamic application of L1 prior knowledge (DePalma & Ringer, 2011) which is thought to be the case with LDHF subordinators in both data sets with *when/cuando*, *porque/because* and *if/si* in both data sets as well as *para* in the CEDEL2/CAES.

In summary, this study found a lack of research into fine-grained meaning-based measures in adverbial subordination. Using EFCAMDAT2 and CEDEL2/CAES in conjunction with the MBF has allowed the current study to track learner trends throughout all six CEFR proficiency levels. Observed trends were incompatible with the DSSICH thus

requiring this study to search for an alternative explanation. The DSMBICH serves to expound on why hypotaxis and enhancement reached peak frequency at advanced levels. Furthermore, it provides an evidence and data backed explanation as to why a large grouping of subordinators are in the zero-use category which is thought to be due to lack of exposure and frequency. HDLF MBUs are thought to indicate L2 acquisition because of the large grouping of low frequency subordinators at advanced proficiency levels. Lastly, the LDHF category may indicate L1 to L2 transfer occurs in a small number of MBUs that appear at high frequencies across a series of CEFR proficiency levels with the caveat that task-based formulaic sequencing might take place at beginner to low intermediate levels while developmental formulaic sequencing could possibly happen at high intermediate to advanced levels.

5. Conclusions

The aim of this thesis and the focal point of research question one was to explore an alternative method of investigating syntactic complexity by discarding traditional measures and applying Systemic Functional Grammar to create a meaning-based framework using fine-grained measures to evaluate L2 English and Spanish texts throughout the CEFR proficiency levels of A1 to C2. The *Método de los Relojes* (2018) Reloj 2 along with *An Introduction to Functional Grammar* (2014) served as a theoretical foundation to bridge Spanish and English with the qualitative purpose of creating a fine-grained meaning-based framework.

The MBF was applied to a large data set which incorporated three separate learner corpora to provide a sizeable representative sample of L2 English and Spanish texts. Data was extracted from the second version of the *Education First-Cambridge Open Language Database* for L1 Spanish writers in L2 English. A combination of the *Corpus escrito del español L2* and the *Corpus de aprendices de español* provided a comparative representative sample for L1 English writers of L2 Spanish. The data analysis explored frequency patterns through the structural functional categories of taxis, expansion, meaning-based categories involving logico-semantic relations and individual subordinators and coordinators which were termed meaning-based units. The fine-grained meaning-based framework was applied to the four previously mentioned layers to provide an alternative perspective to traditional syntactic complexity measures.

A central aspect of this study which was projected through research question two was to examine the merits of the developmental sequence of syntactic interlanguage complexification hypothesis (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) which ascertains that coordination is prevalent at the beginner stages while subordination becomes dominant at intermediate levels to be replaced by phrasal complexity through grammatical metaphor at an advanced juncture. Findings from the EFCAMDAT2 and the CEDEL2/CAES were inconsistent with the aforementioned hypothesis in hypotaxis and parataxis as well as with expansion through extension, enhancement and elaboration. The majority of MBCs did not exhibit behavior indicative of the said theory. The only exceptions occurred in the CEDEL2/CAES in hypotaxis through enhancement with spatial indicators and in the

EFCAMDAT2 in parataxis through enhancement in the temporal and spatial categories. A limitation of this study and an area for future research is to analyze the effect of grammatical metaphor throughout CEFR proficiency levels paying special attention to advanced levels.

The goal of research question three was to gauge to what extent the meaning-based framework identified cross-linguistic meaning-based complexity patterns in relation to CEFR proficiency levels. There were three meaning-based categories which stood out as the workhorses of the entire study. Temporal and causal-conditional subordination in hypotaxis through enhancement demonstrated increased levels of subordination across CEFR proficiency levels in both the EFCAMDAT2 and the CEDEL2/CAES. Parataxis and extension through addition was the meaning-based category with the highest frequency rates in the entire study for both data sets. Frequency was driven by the singular coordinator *and* as well as the equivalent *y* which displayed consistent use from beginner to advanced CEFR levels demonstrating either MBU was a constant linguistic resource for Spanish and English L2 learners. Future research might want to focus on including additional measures to meaning-based categories, especially in paratactic enhancement, to investigate if different patterns appear.

Research question four investigated to what extent the MBF could pinpoint frequency trends which might indicate L2 acquisition and L1 to L2 transfer. The analysis of meaning-based units, individual subordinators and coordinators, necessitated the creation of four categories based on frequency to quantify results. The zero, low, medium and high use frequency categories are based on the parameters of the representative sample used by this study. Findings may be different for other studies trying to replicate results depending on the number of fine-grained measures and data involved. The expansive category of zero-use MBUs in hypotaxis and enhancement is thought to be due to lack of frequency and exposure to the target unit. There were a vast number of MBUs used at a low frequency which appeared at advanced levels. It is suggested that this phenomenon is a result of accumulated frequency and exposure of learners to the second language which culminates in visible use patterns at advanced proficiency levels indicating L2 acquisition.

The final two frequency categories for adverbial subordination occurred at medium and high use frequencies. The low number of medium frequency MBUs is thought to be a result of lack of exposure to target subordinators in the L2. As frequency increases in the low

density medium frequency category, the notion of L1 to L2 transfer comes into play as it is thought to be a driving factor in the upper ranges of subordination use. There were only three high frequency MBUs extrapolated from EFCAMDAT2 data (*when, because and if*) and four from the CEDEL2/CAES (*cuando, para, porque and si*). Stable yet high frequencies across a series of proficiency levels suggests this is an indication of L1 to L2 transfer. However, certain instances in the form of tasked-based and developmental formulaic sequencing appeared thus leading this study to believe there are several factors which influence MBUs as frequency increases.

The frequency trends observed in this study clearly did not confirm the developmental sequence of syntactic interlanguage complexification hypothesis (Byrnes et al., 2010; Colombi, 2002; Halliday & Mathiessen, 2006; Neary-Sundquist, 2016; Norris & Ortega, 2009; Wolf-Quitero et al., 1998) yet a solid explanation as to why the meaning-based framework using fine-grained measures produced alternative findings did not seem to be apparent. This study holds the viewpoint that the aforementioned hypothesis takes a broad perspective on the notion of subordination within syntactic complexity. The answers to research questions two, three and four serve as the basis which establishes the foundation for an alternative perspective with the singular focal point on adverbial subordination in which current study introduces the developmental sequence of meaning-based interlanguage complexification hypothesis.

The genesis for the developmental sequence of meaning-based interlanguage complexification hypothesis which proposes, in line with the results from this study, that L2 English and Spanish learners employ zero-use MBUs due to a lack of exposure to target units. If they have been exposed to subordinators in this group, the cognitive demand of using zero-use MBUs is beyond their abilities at that moment in their learning trajectory. Low frequency adverbial subordination is the accumulation of exposure and acquisition throughout L2 learning which culminates in a variety of MBUs being used at advanced levels. As frequency increases to the highest rates, L1 to L2 transfer starts to take place. Evidence from both data sets shows higher frequency MBUs tend to have a consistent elevated frequency throughout a sequence of proficiency levels.

It must be noted that formulaic sequencing may have an influence on higher frequency MBUs. We propose that formulaic sequencing with singular MBUs in isolated

lower proficiency levels may be tasked-based (Alexopoulou et al., 2015; Chen, 2021; Alhassan & Wood, 2015; Lewis, 1997; Wray, 2002). Developmental formulaic sequencing (Boers et al., 2006; Boers & Lindstromberg, 2012; Dai & Ding, 2010; Schmitt, 2008; Stengers et al., 2011), being the native-like production of a target unit due to exposure, may appear to be a factor in isolated frequency peaks at advanced levels.

There are many aspects to this study which would benefit from future research. Namely, the developmental sequence of meaning-based interlanguage complexification hypothesis requires further investigation in L2 English and Spanish to determine the replicability of results. Applying the previously mentioned hypothesis to a different range of L2s would also prove useful in comparing cross-linguistic trends. Separating the analysis by writing genre would also provide insight as to how coordination and subordination can be analyzed through the meaning-based framework. A line of research which would be interesting to investigate would be to apply *Método de los Relojes* R2 hours to research how L1 English learners use L2 Spanish subordination categories. The original intention was to include R2 hours as well as embedded clauses in this thesis. However, the overall extension and the complexity of creating and applying the meaning-based framework limited this study.

What could be seen as a limitation of the creation and the application of the meaning-based framework is that it requires an exhaustive effort to sort through and classify data. Not to mention the discretion of the researcher(s) in defining the meaning of MBUs as well as the comparability of MBUs in a cross-linguistic study. While it is believed that individual meaning-based units were correctly extrapolated and classified under correct categories, a slight margin of error exists. To ensure increased accuracy, a natural language processing program could be created within the parameters of the meaning-based framework. As of the completion of this thesis, there is no known application to remove and analyze the scope of meaning-based measures employed in this thesis.

References

- Achugar, M., & Colombi, M.C., (2009). Systemic Functional Linguistics explorations into the longitudinal study of advanced capacities. In H. Byrnes & L. Ortega (Eds.), *The longitudinal study of advanced L2 capacities* (1st ed., pp. 36-57). Routledge.
<https://doi-org.unican.idm.oclc.org/10.4324/9780203871652>
- Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in nns and ns university students writing. In A. Díaz-Negrillo, N. Ballier & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249-264). John Benjamins. <https://doi.org/10.1075/scl.59.15ai>
- Alderson, J. (2007). The CEFR and the Need for More Research. *The Modern Language Journal*, 91(4), 659-663. <http://www.jstor.org/stable/4626093>
- Alexopoulou, T., Geertzen, J., Korhonen, A., & Meurers, D. (2015). Exploring big educational learner corpora for SLA research. *International Journal of Learner Corpus Research*, 1(1), 96-129. <https://doi.org/10.1075/ijlcr.1.1.04ale>
- Alhassan, L., & Wood, D. (2015). The effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills: A quantitative research study. *Journal of English for Academic Purposes*, 17, 51-62.
<http://dx.doi.org/10.1016/j.jeap.2015.02.001>
- Appel, R., & Wood, D. (2016). Recurrent word combinations in EAP test-taker writing: Differences between high-and low-proficiency levels. *Language Assessment Quarterly*, 13(1), 55-71.
<https://doi-org.unican.idm.oclc.org/10.1080/15434303.2015.1126718>
- Aray, C., Coloma, C.J., Quezada, C., & Benavente, P. (2023). Development of Clause Complexity in Children with Specific Language Impairment/Language Disorder: A Longitudinal Study. *Children*, 10(7), 1152-1166.
<https://doi.org/10.3390/children10071152>

- Arús, J., Lavid, J., & Zamorano-Mansilla, J.F. (2010). *Systemic Functional Grammar of Spanish: A Contrastive Study with English*. Continuum International Publishing Group.
- Asención-Delaney, Y., & Collentine, J. (2011). A Multidimensional Analysis of a Written L2 Spanish Corpus. *Applied Linguistics*, 32(3), 299-322.
<https://doi-org.unican.idm.oclc.org/10.1093/applin/amq053>
- Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26(2), 390-395.
<https://doi-org.unican.idm.oclc.org/10.2307/3587016>
- Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition* 11(1), 17-34. <https://doi.org/10.1017/S0272263100007816>
- Basterrechea, M., & Weinert, R. (2017). Examining the Concept of Subordination in Spoken L1 and L2 English: The Case of If-Clauses. *TESOL Quarterly*, 51(4), 897-919.
<https://doi-org.unican.idm.oclc.org/10.1002/tesq.361>
- Baten, K., & Håkansson, G. (2015). The development of subordinate clauses in German and Swedish as L2s: A theoretical and methodological comparison. *Studies in Second Language Acquisition*, 37(3), 517-547. <https://doi.org/10.1017/S0272263114000552>
- Bax, S., Nakatsuhara, F., & Waller, D. (2019). Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels. *System*, 83, 79-95.
<https://doi-org.unican.idm.oclc.org/10.1016/j.system.2019.02.010>
- Bhowmik, S. K. (2016). Agency, identity and ideology in L2 writing: Insights from the EAP classroom. *Writing & Pedagogy*, 8(2), 275-308.
<https://doi.org/10.1558/wap.26864>
- Biber, D. (2006). *University Language: A Corpus- Based Study of Spoken and Written Registers*. John Benjamins. <https://doi.org/10.1075/scl.23>

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Grammar Of Spoken and Written English* (6 ed.). Pearson Education Limited, Essex.
<https://doi.org/10.1075/z.232>
- Biber, D., Conrad, S. M., Reppen, R., Byrd, P., Helt, M., Clark, V., Cortes, V., Csomay, E., & Urzua, A. (2004). *Representing language use in the university: Analysis of the TOEFL 2000 spoken and written academic language corpus*. Princeton: Technical report Educational Testing Service.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5-35. <https://doi-org.unican.idm.oclc.org/10.5054/tq.2011.244483>
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10(3), 245-261.
<https://doi-org.unican.idm.oclc.org/10.1191/1362168806lr195oa>
- Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies formulaic sequences in a second language. *Annual Review of Applied Linguistics*, 32, 83-110.
<https://doi.org/10.1017/S0267190512000050>
- Bolton, K., Nelson, G., & Hung, J. (2003). A corpus-based study of connectors in student writing: Research from the International Corpus of English in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics*, 7(2). 165-182.
<https://doi.org/10.1075/ijcl.7.2.02bol>
- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21-46). John Benjamins.
<https://doi.org/10.1075/llt.32.02bul>

- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42-65.
<https://doi-org.unican.idm.oclc.org/10.1016/j.jslw.2014.09.005>
- Bulté, B., & Housen, A. (2018). Syntactic complexity in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics*, 28(1), 147-164.
<https://doi-org.unican.idm.oclc.org/10.1111/ijal.12196>
- Bybee, J. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Acquisition* (1st ed., pp 226-246). Routledge.
- Byrnes, H. (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education*, 20(1), 50-66.
<https://doi-org.unican.idm.oclc.org/10.1016/j.linged.2009.01.005>
- Byrnes, H., Maxim, H. H., & Norris, J. M. (2010). Realizing Advanced Foreign Language Writing Development in Collegiate Education: Curricular Design, Pedagogy, Assessment. *The Modern Language Journal*, 94, i-235.
<http://www.jstor.org/stable/40985261>
- CAES (2023). Corpus de aprendices de español como lengua extranjera.
 Retrieved October 14, 2023, from <https://galvan.usc.es/caes>
- Casanave, C. P. (1994). Language development in students' journals. *Journal of Second Language Writing*, 3(3), 179-201.
[https://doi-org.unican.idm.oclc.org/10.1016/1060-3743\(94\)90016-7](https://doi-org.unican.idm.oclc.org/10.1016/1060-3743(94)90016-7)
- CEDEL2 (2023). Corpus Escrito del Español L2. Retrieved August 29, 2023 from
<http://cedel2.learnercorpora.com/>
- Chen, C. (2006). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, 11(1). 113-130.
<https://doi.org/10.1075/ijcl.11.1.05che>

- Chen, X., Alexopoulou, T., & Tsimpli, I. (2021). Automatic extraction of subordinate clauses and its application in second language acquisition research, *Behavior Research Methods*, 53, 803-817.
<https://doi.org/10.3758/s13428-020-01456-7>
- Cohen, A. (2014). *Strategies in Learning and Using a Second Language*. Routledge.
<https://doi-org.unican.idm.oclc.org/10.4324/9781315833200>
- Collentine, J., & Collentine, K. (2020). Organic Models for Measuring Spanish Learners' Linguistic Complexity. *Current Theoretical and Applied Perspectives on Hispanic and Lusophone Linguistics*, 27, 39-62. <https://doi.org/10.1075/ihll.27>
- Collins Cobuild (1990). *Collins Cobuild English Grammar*. Collins Publishers.
- Colombi, C. (2002). Academic language development in Latino students' writing in Spanish. In M. Schleppergrell, & M. C. Colombi (Eds.), *Developing advanced literacy in first and second languages: Meaning with power* (pp. 67-86). Mahwah, NJ: Lawrence Erlbaum Associates.
- Council of Europe. Council for Cultural Co-operation. Education Committee, Modern Language Division. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
<https://rm.coe.int/1680459f97>
- Council of Europe. Council for Cultural Co-operation. Education Committee, Modern Language Division. (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment Companion Volume with New Descriptors*. Cambridge University Press.
<https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>

- Cummins, J. (2008). Teaching for transfer: Challenging the two solitudes assumption in bilingual education. In J. Cummins, & N. H. Hornberger (Eds.), *Encyclopedia of language and education Vol. 5. Bilingual education* (2nd ed., pp. 65-75). Springer. https://doi.org/10.1007/978-0-387-30424-3_116
- Dai, Z., & Ding, Y. (2010). Effectiveness of text memorization in EFL learning of Chinese students. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 71-87). Continuum.
- De Clercq, B., & Housen, A. (2017). A Cross-Linguistic Perspective on Syntactic Complexity in L2 Development: Syntactic Elaboration and Diversity. *The Modern Language Journal*, 101(2), 315-334. <https://doi-org.unican.idm.oclc.org/10.1111/modl.12396>
- DePalma, M. J., & Ringer, J. M. (2011). Toward a theory of adaptive transfer: Expanding disciplinary discussions of “transfer” in second-language writing and composition studies. *Journal of Second Language Writing*, 20(2), 134-147. <https://doi-org.unican.idm.oclc.org/10.1016/j.jslw.2011.02.003>
- Deygers, B., Carlsen, C., Saville, N., & Van Gorp, K. (2018a). The Use of the CEFR in Higher Education: A Brief Introduction to This Special Issue, *Language Assessment Quarterly*, 15(1), 1-2. <https://doi.org/10.1080/15434303.2017.1421957>
- Deygers, B., Zeidler, B., Vilcu, D., & Hamnes Carlsen, C. (2018b). One Framework to Unite Them All? Use of the CEFR in European University Entrance Policies. *Language Assessment Quarterly*, 15(1), 3-15. <http://dx.doi.org/10.1080/15434303.2016.1261350>
- Eastwood, J. (1994). *Oxford Guide to English Grammar*. Oxford University Press.
- Ellis, N. C. (2008). Usage-based and form-focused SLA: The implicit and explicit learning of constructions. In A. Tyler, Y. Kim, & M. Takada (Eds.), *Language in the context of use: Discourse and cognitive approaches to language* (pp. 93-121). Mouton de Gruyter. <https://doi.org/10.1515/9783110199123.1.93>

- Ellis, R., & Barkhuizen, R. (2005). *Analyzing learner language*. Oxford, England: Oxford University Press.
- Ehret, K., & Szmrecsanyi, B. (2019). Compressing learner language: An information-theoretic measure of complexity in SLA production data. *Second Language Research*, 9(1), 23-45. <https://doi.org/10.1177/0267658316669559>
- Ferris, D. R. (1994). Rhetorical Strategies in Student Persuasive Writing: Differences between Native and Non-Native English Speakers. *Research in the Teaching of English*, 28(1), 45-65. <http://www.jstor.org/stable/40171324>
- Figueras, N. (2007). The CEFR, a Lever for the Improvement of Language Professionals in Europe. *The Modern Language Journal*, 91(4), 673-675. <https://www.jstor.org/stable/4626097>
- Figueras, N. (2012). The impact of CEFR. *ELT Journal*, 66(4), 477-485. <https://doi.org/10.1093/elt/ccs037>
- Forbes, K. (2019). The role of individual differences in the development and transfer of writing strategies between foreign and first language classrooms. *Research Papers in Education*, 34(4), 445-464. <https://doi.org/10.1080/02671522.2018.1452963>
- Gast, V. (2012). Contrastive analysis: Theories and methods. In B. Kortmann, & J. Kabatek (Eds.), *Dictionaries of Linguistics and Communication Science: Linguistic theory and methodology*. Berlin: Mouton de Gruyter.
- Giagkou, M., Kantzou, V., Stamouli, S., & Tzevelekou, M. (2015). Discriminating CEFR levels in Greek L2: a corpus-based study of young learners' written narratives. *Bergen Language and Linguistic Studies*, 6. <https://doi.org/10.15845/bells.v6i0.813>

- Granger S. (2017). Learner Corpora in Foreign Language Education. In N. Van Deusen-Scholl and N.H. Hornberger (Eds.), *Encyclopedia of Language and Education. Volume 4. Second and Foreign Language Education* (pp. 427-440). Springer.
https://doi.org/10.1007/978-3-319-02237-6_33
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers*, 36, 193-202. <https://doi.org/10.3758/BF03195564>
- Green, A. (2012). *Language functions revisited: Theoretical and empirical bases for language construct definition across the ability range* (Vol. 2). Cambridge University Press.
- Gyllstad, H., Granfeldt J., Bernardini, P., & Källkvist, M. (2014). Linguistic correlates to communicative proficiency levels of the CEFR: The case of syntactic complexity in written L2 English, L3 French and L4 Italian. *EUROSLA Yearbook*, 14(1), 1-30.
<https://doi.org/10.1075/eurosla.14.01gyl>
- Halliday, M. A. K. (2004). *An Introduction to Functional Grammar* (3rd ed.). Hodder Arnold.
<https://doi.org/10.4324/9780203783771>
- Halliday, M.A. K., & Matthiessen, M. (2006). *Construing Experience through Meaning: A Language-Based Approach to Cognition*. London: Continuum. [Study Edition of 1999 book]
- Halliday, M. A. K. (2014). *An Introduction to Functional Grammar* (4th ed.). Routledge.
<https://doi.org/10.4324/9780203431269>
- Harrington, M. (1987). Processing transfer: Language-specific processing strategies as source of interlanguage variation. *Applied Psycholinguistics*, 8 (4), 351-377.
<https://doi.org/10.1017/S0142716400000370>

- Hawkins, J. A., & Filipović, L. (2012). *Criterial features in L2 English: Specifying the reference levels of the Common European Framework* (Vol. 1). Cambridge University Press.
- Henriksen, B. (2013). Research on L2 learners' collocational competence and development: A progress report. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use* (pp. 29-56). Eurosla Monographs.
- Hernández, A. E., Bates, E., & Avila, L.X. (1994). Sentence interpretation in Spanish-English bilinguals: What does it mean to be in-between?. *Applied Psycholinguistics*, 15(4), 417-466. <https://doi.org/10.1017/S014271640000686X>
- Hernández, A. E., Sierra, I., & Bates, E. (2000). Sentence interpretation in bilingual and monolingual Spanish speakers: Grammatical processing in a monolingual mode. *Spanish Applied Linguistics*, 4(2), 179-213.
- Homburg, T. J. (1984). Holistic evaluation of ESL compositions: Can it be validated objectively? *TESOL Quarterly*, 18(1), 87-107. <https://doi.org/10.2307/3586337>
- Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency, Language Learning & Language Teaching* (pp. 1-20). John Benjamins. <https://doi.org/10.1075/llt.32>
- Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, 53, 13-23. <https://doi.org/10.1016/j.system.2015.06.011>
- Hulstijn, J. (2007). The Shaky Ground beneath the CEFR: Quantitative and Qualitative Dimensions of Language Proficiency. *The Modern Language Journal*, 91(4), 663-667. <https://www.jstor.org/stable/4626094>

- Hulstijn, J.H., Alderson, J.C., & Schoonen, R. (2010). Developmental stages in second-language acquisition and levels of second-language proficiency: Are there links between them. In I. Bartning, M. Martin, & I. Vedder (Eds), *Communicative Proficiency and Linguistic Development: Intersections between SLA and Language Testing Research* (pp. 11-20). Eurosla Monographs.
- Ishikawa, S. (1995). Objective measurement of low-proficiency EFL narrative writing. *Journal of Second Language Writing*, 4(1), 51-69.
[https://doi.org/10.1016/1060-3743\(95\)90023-3](https://doi.org/10.1016/1060-3743(95)90023-3)
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. Routledge. <https://doi.org/10.4324/9780203935927>
- Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow, & S. Kemmer (Eds.), *Usage-based models of language* (pp. vii–xxviii). CSLI Publications.
- Kilborn, K. (1989). Sentence processing in a second language: The timing of transfer. *Language and Speech*, 32(1), 1-23. <https://doi.org/10.1177/002383098903200101>
- Kim, S. (2004). A study of development in syntactic complexity by Chinese learners of English and its implications on the CEF scales [unpublished MA dissertation]. Lancaster University.
- Khushik, A.B., & Huhta, A. (2020). Investigating Syntactic Complexity on EFL Learners' Writing across Common European Framework of Reference Levels A1, A2, and B1. *Applied Linguistics*, 41(4). 506-532. <https://doi.org/10.1093/applin/amy064>
- Kortmann, B. (1999). Iconicity, Typology and Cognition. In M. Nänny and O. Fisher (Eds.), *Form miming meaning* (pp. 375-392). John Benjamins.
<https://doi.org/10.1075/ill.1.27kor>
- Kuiken, F. (2023). Linguistic complexity in second language acquisition. *Linguistics Vanguard*, 9(s1), 83-93. <https://doi.org/10.1515/lingvan-2021-0112>

- Kuiken, F., Vedder, I., Housen, A., & De Clercq B. (2019). Variation in syntactic complexity: Introduction. *International Journal of Applied Linguistics*, 29(2), 161-170. <https://doi.org/10.1111/ijal.12255>
- Kuiken, F., & Vedder, I. (2019). Syntactic Complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. *International Journal of Applied Linguistics*, 29(2), 192-210. <https://doi.org/10.1111/ijal.12256>
- Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage- based indices of syntactic sophistication* [PhD dissertation, Georgia State University].
- Kyle, K., Crossley, S., & Berger, C. (2018). The tool for the automatic analysis of lexical sophistication (TAALES): Version 2.0. *Behavior Research Methods*, 50, 1030-1046. <https://doi.org/10.3758/s13428-017-0924-4>
- Kyle, K., & Crossley, S. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal*, 102(2), 333-349. <https://doi.org/10.1111/modl.12468>
- Lahuerta Martínez, A.C. (2018a). Analysis of syntactic complexity in secondary education EFL writers at different proficiency levels. *Assessing Writing*, 35, 1-11. <https://doi.org/10.1016/j.asw.2017.11.002>
- Lahuerta Martínez, A. C. (2018b). Conjunctions in the English Essay Writing of Spanish EFL Learners. In *Conference Proceedings. Innovation in Language Learning 2018*. <https://pdfs.semanticscholar.org/d884/69f385b920b622c3c836b840e67dac7e72c6.pdf>
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590-619. <https://doi.org/10.1093/applin/aml029>

- Lei, L., Wen, J., & Yang, X. (2023). A large-scale longitudinal study of syntactic complexity development in EFL writing: A mixed-effects model approach. *Journal of Second Language Writing*, 59. <https://doi.org/10.1016/j.jslw.2022.100962>
- Leung, C., & Lewkowicz, J. (2013). Language communication and communicative competence: A view from contemporary classrooms. *Language and Education*, 27(5), 398-414. <https://doi.org/10.1080/09500782.2012.707658>
- Lewis, M. (1997). Pedagogical implications of the lexical approach. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (255-270). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524643>
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18(2), 85-102. <https://doi.org/10.1016/j.jslw.2009.02.001>
- Li, X., & Yu, H. (2021). Parataxis or hypotaxis? Choices of taxis in Chinese English translation. *Lingua*, 251. <https://doi.org/10.1016/j.lingua.2020.103026>
- Lindstromberg, S. (2010). *English Prepositions Explained* (2nd ed.). John Benjamins Publishing Company. <https://doi.org/10.1075/z.157>
- Lindstromberg, S., Eyckmans, J., & Connabeer, R. (2016). A modified dictogloss for helping learners remember L2 academic English formulaic sequences for use in later writing. *English for Specific Purposes*, 41, 12-21. <https://doi.org/10.1016/j.esp.2015.08.002>
- Litosseliti, L. (2010). *Research Methods in Linguistics*. Continuum International Publishing Company.
- Little, D. (2007). The Common European Framework of Reference for Languages: Perspectives on the Making of Supranational Language Education Policy. *The Modern Language Journal*, 91(4), 645-655. <http://www.jstor.org/stable/4626091>

- Lorenzo, F., Granados, A., & Ávila, I. (2019). The development of cognitive language proficiency in multilingual education: Evidence of a longitudinal study on the language of history. *Journal of English for Academic Purposes*, 41. <https://doi.org/10.1016/j.jeap.2019.06.010>
- Lorenzo, F., & Rodríguez, L. (2014). Onset and expansion of L2 cognitive academic language proficiency in bilingual settings: CALP in CLIL. *System*, 47, 64-72. <https://doi.org/10.1016/j.system.2014.09.016>
- Lozano, C. (2022). CEDEL2: Design, compilation and web interface of an online corpus for L2 Spanish acquisition research. *Second Language Research*, 38(4), 965-982. <https://doi-org.unican.idm.oclc.org/10.1177/026765832110505>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45(1), 36-62. <https://doi.org/10.5054/tq.2011.240859>
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x
- Lu, X. & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing* 29, 16-27. <https://doi-org.unican.idm.oclc.org/10.1016/j.jslw.2015.06.003>
- Martinez, R., & Murphy, V. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly*, 45(2), 267-290. <https://doi.org/10.5054/tq.2011.247708>

- Matthiessen, C. (1989). [Review of book *Introduction to Functional Grammar* by M. A. K. Halliday]. *Language*, 65(4), 862-87.
- Maxim, H. (2021). A longitudinal case study of curriculum-based L2 writing development. *System*, 96, 2-17. <https://doi.org/10.1016/j.system.2020.102397>
- Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3-15. <https://doi.org/10.1016/j.jslw.2015.06.004>
- Müller, N., & Penner, Z. (1996). Early subordination: the acquisition of free morphology in French, German, and Swiss German. *Linguistics*, 34(1), 133-165. <https://doi.org/10.1515/ling.1996.34.1.133>
- Neary-Sundquist, C. A. (2016). Syntactic complexity at multiple proficiency levels of L2 German speech. *International Journal of Applied Linguistics*, 27(1), 242-262. <https://doi.org/10.1111/ijal-12128>.
- Neff, J., Dafouz, E., Díez, M., Prieto, R., & Chaudron, C. (2004). Contrastive discourse analysis: Argumentative text in English and Spanish. In: C.L. Moder & A. Martinovic-Zic (Eds.), *Discourse across Languages and Cultures* (pp. 267-283). John Benjamins. <https://doi.org/10.1075/slcs.68>
- Norrby, C., & Håkansson, G. (2007). The interaction of complexity and grammatical processability: The case of Swedish as a foreign language. *International Review of Applied Linguistics in Language Teaching*, 45(1), 45-68. <https://doi.org/10.1515/IRAL.2007.002>
- Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30 (4), 555-578. <https://doi.org/10.1093/applin/amp044>
- North, B. (2007). The CEFR Illustrative Descriptor Scales. *The Modern Language Journal*, 91(4), pp. 656-659. <https://www.jstor.org/stable/4626092>

- North, B., & Piccardo, E. (2016). Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR). *Language Teaching*, 49(3), 455-459. http://journals.cambridge.org/abstract_S0261444816000100
- Odlin, T. (2003). Cross-Linguistic Influences. In C. Doughty, & M. Long (Eds.), *The Handbook of Second Language Acquisition* (1st ed., pp. 436-486). Blackwell Publishing. <https://doi.org/10.1002/9780470756492.ch9>
- Ortega, L. (2000). *Understanding syntactic complexity: The measurement of change in the syntax of instructed L2 Spanish learners*. [PhD dissertation, University of Hawaii].
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492-518. <https://doi.org/10.1093/applin/24.4.492>
- Ortega, L. (2009). Studying writing across efl contexts: Looking back and moving forward. In R.M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research*, (pp. 232-255). Multilingual Matters.
- Ortega, L. (2012). Interlanguage complexity: A construct in search of theoretical renewal. In B. Szmrecsanyi, & B. Kortmann (Eds.), *Linguistic complexity: Second language acquisition, indigenization, contact* (pp. 127-155). De Gruyter. <https://doi.org/10.1515/9783110229226>
- Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing* 29, 2-94. <https://doi.org/10.1016/j.jslw.2015.06.008>
- Pallotti, G. (2009). CAF: Defining, Refining and Differentiating Constructs. *Applied Linguistics*, 30(4), 590-601. <https://doi.org/10.1093/applin/amp045>
- Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research* 31(1), 117-134. <https://doi.org/10.1177/0267658314536435>

- Pérez Saiz, M. (2018). *Método de los relojes. Gramática descriptiva del español* (4a ed.). Publican, Universidad de Cantabria.
- Peters, E. (2014). The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations. *Language Teaching Research*, 18(1), 75-94. <https://doi.org/10.1177/1362168813505384>
- Phakiti, A., De Costa, P., Plonsky, L., & Starfeild, S. (2018). *The Palgrave Handbook of Applied Linguistic Research Methodology*. Palgrave Macmillan.
[doi:10.1057/978-1-137-59900-1](https://doi.org/10.1057/978-1-137-59900-1)
- Polio, C., & Shea, C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10-27.
<https://doi-org.unican.idm.oclc.org/10.1016/j.jslw.2014.09.003>
- Qi, Y., & Ding, Y. (2011). Use of formulaic sequences in monologues of Chinese EFL learners. *System*, 39(2), 164-174. <https://doi.org/10.1016/j.system.2011.02.003>
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. Longman.
- Rasool, A., & Mahmood, A. (2023). Comparative Study of Logical Meta-Function in Pakistani High and Low-Graded Argumentative Essays. *Journal of Positive School Psychology*, 7(2), 1540-1557.
https://www.researchgate.net/publication/372570612_Comparative_Study_of_Logical_Meta-functions_in_Pakistani_High_and_Low-Graded_Argumentative_Essays
- Real Academia Española, & Asociación de Academias de la Lengua Española. (2010). *Nueva gramática de la lengua española*, 2. Madrid: Espasa.
- Restrepo-Remos, F. (2021). Writing to Increase Complexity. *Hispania*, 104(4), 671-690.
<https://www.jstor.org/stable/27127542>

- Reyes, I., & Hernández, A. (2006). Sentence interpretation strategies in emergent bilingual children and adults. *Bilingualism: Language and Cognition*, 9(1), 51-69.
<https://doi.org/10.1017/S1366728905002373>
- Rinnert, C., & Kobayashi, H. (2016). Multicompetence and multilingual writing. In R. M. Manchón, & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 365-386). Mouton de Gruyter.
<https://doi.org/10.1515/9781614511335>
- Rinnert, C., Kobayashi, H., & Katayama, A. (2015). Argumentation text construction by Japanese as a foreign language writer: A dynamic view of transfer. *Modern Language Journal*, 99(2), 213-245. <https://doi.org/10.1111/modl.12210>
- Ryshina-Pankova, M. (2015). A meaning-based approach to the study of complexity in L2 writing: The case of grammatical metaphor. *Journal of Second Language Writing*, 29, 51-63. <https://doi.org/10.1016/j.jslw.2015.06.005>
- Saldanha, G., & O'Brien, S. (2014). *Research Methodologies in Translation Studies*. Routledge. <https://doi.org/10.4324/9781315760100>
- Sasaki, Y. (1991). English and Japanese interlanguage comprehension strategies: An analysis based on the competition model. *Applied Psycholinguistics*, 12(1), 47-73.
<https://doi.org/10.1017/S0142716400009371>
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363.
<https://doi.org/10.1177/136216880808992>
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 1-22). John Benjamins.
<https://doi.org/10.1075/llt.9>

- Sharma, A. (1980). Syntactic maturity: Assessing writing proficiency in a second language. In R. Silverstein (Ed.), *Occasional papers in linguistics* (pp. 318–325). Southern Illinois University.
- Siegel, J. (2003). Substrate influence in creoles and the role of transfer in second language acquisition. *Studies in Second Language Acquisition*, 25(2), 185-209.
<https://doi.org/10.1017/S0272263103000093>
- Sinclair, J. (2005). How to make a corpus. In M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice* (pp. 79-83). Oxford.
<https://users.ox.ac.uk/~martinw/dlc/index.htm>
- Siyanova-Chanturia, A. (2015). Collocation in beginner learner writing: A longitudinal study. *System*, 53, 148-160. <https://doi.org/10.1016/j.system.2015.07.00>
- Sotillo, S. M. (2000). Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning and Technology* 4, 82-119.
- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2010). Does ‘chunking’ foster chunk-uptake? In S. De Knop, F. Boers, & A. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 99-117). Mouton de Gruyter.
<https://doi.org/10.1515/9783110245837>
- Stefanowitsch, A. (2020). *Corpus Linguistics: A guide to the methodology*. Language Science Press.
[file:///Users/timothylawrence/Documents/PhD/Final Draft/10.5281/zenodo.3735822](file:///Users/timothylawrence/Documents/PhD/Final%20Draft/10.5281/zenodo.3735822)
- Sulistyaningrum, S., & Rasyid, Y. (2015). The Logico-semantic Relation of Students’ Presentation in acceleration program of SMA Labschool Jakarta. *International Journal of Language Education and Culture Review*, 1(1), 41-54.
<https://doi.org/10.21009/IJLECR.011.05>

- Tarone, E., & Swierzbin, B. (2009). *Exploring learner language*. Oxford, England: Oxford University Press.
- Tan, Y., & Römer, U. (2022). Using phrase-frames to trace the language developments of L1 Chinese learners of English. *System*, 108.
<https://doi.org/10.1016/j.system.2022.102844>
- Veel, R. (1997). Learning how to mean-scientifically speaking: Apprenticeship into scientific discourse in the secondary school. In F. Christie, & J. R. Martin (Eds.), *Genre and institutions: Social processes in the workplace and school* (pp. 161–195). Continuum.
- Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, 21(3), 239-263.
<https://doi.org/10.1016/j.jslw.2012.03.007>
- Velázquez-Mendoza, O. (2015). The role of grammatical metaphor in development of advanced literacy in Spanish as a first, second and heritage language. *Functional Linguistics*, 2(7). <https://doi.org/10.1186/s40554-015-0020-y>
- Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *Modern Language Journal*, 96(4), 576-598. <https://doi.org/10.1111/j.1540-4781.2012.01401.x>
- Wenhui Xuan, W. (2019). Understanding the logical meaning-making in Chinese high school students' EFL writing: A systemic perspective. *The Asian EFL Journal*, 23(6.1), 7-33.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). Second language development in writing: Measures of fluency, accuracy, & complexity. Technical report, Second Language Teaching & Curriculum Center, University of Hawaii at Manoa.
- Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge University Press.

- Wulfeck, B., Juarez, L., Bates, E., & Kilborn, K. (1986). Sentence interpretation strategies in healthy and aphasic bilingual adults. In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neurological perspectives* (pp. 199-220). Lawrence Erlbaum Associates.
- Xu, Y., & Li, Y. (2012). The Use of Adverbial Conjuncts of Chinese EFL Learners and Native Speakers-Corpus-based Study. *Theory and Practice in Language Studies*, 2(11), 2316-2321.
<https://www.academypublication.com/issues/past/tpls/vol02/11/12.pdf>
- Yan, H. (2010). The role of L1 transfer on L2 and pedagogical implications. *Canadian Social Science*, 6(3), 97-103.
- Yang, M., Cooc, N., & Sheng, L. (2017). An investigation of cross-linguistic transfer between Chinese and English: a meta-analysis. *Asian-Pacific Journal of Second and foreign Language Education*, 2. <https://doi.org/10.1186/s40862-017-0036-9>
- Yaqub, H., & Shakir, A. (2019). A Corpus-Based Study of Hypotactic and Paratactic Thematic Relations in English and Urdu Clause Complexes. *International Journal of English Linguistics*, 9(5), 430-448. <https://doi.org/10.5539/ijel.v9n5p430>
- Zarco-Tejada, A., Noya Gallardo, C., Merino Ferradá, M., & López Calderón, I. (2016). Analysing corpus-based criterial conjunctions for automatic proficiency classification. *Journals of English Studies*, 14, 215-237. <http://doi.org/10.18172/jes.3090>
- Zhang, X., & Lu, X. (2022). Revisiting the predictive power of traditional vs. fine-grained syntactic complexity indices for L2 writing quality: The case of two genres. *Assessing Writing*, 51. <https://doi.org/10.1016/j.asw.2021.100597>