

COGNICIÓN Y SELECCIÓN DE SUJETOS: ANÁLISIS COMPARATIVO ENTRE
NARRATIVAS ESCRITAS Y ORALES EN EL DISCURSO DE ESPAÑOL COMO L2.

by

INGRID ABISAMBRA MICCHELI

(Under the Direction of Margaret Lubbers Quesada)

RESUMEN

La *Givenness Hierarchy* (Gundel, Hedberg y Zacharski, 1993) propone que la cognición afecta la selección y uso de las formas referenciales en el discurso. La jerarquía identifica seis estados cognitivos (ECs) basados en las presuposiciones de los hablantes con respecto a los niveles de atención y memoria de los oyentes. Dichos estados pueden predecir las formas que usarán los hablantes para referirse a diferentes entidades. Un mayor nivel de atención y memoria predice el uso de expresiones mínimas y lo opuesto corresponde con el uso de formas más elaboradas. Blackwell y Lubbers Quesada (2012) revisaron la jerarquía para el español y probaron su poder predictivo en las narrativas orales de hablantes nativos y hablantes de español como segunda lengua (L2).

El presente estudio examina si la jerarquía revisada es igualmente válida para predecir el uso de los sujetos en español en los textos escritos, donde los escritores pueden hacer menos presuposiciones acerca de los niveles de atención y memoria de los lectores con respecto al referente. El objetivo secundario es el de establecer si el terreno común (TC) afecta la selección de las formas referenciales en las narrativas escritas.

Veinte hablantes avanzados de español L2 y 20 hablantes nativos vieron un segmento de la película muda de Charlie Chaplin, *A Woman*, y se les pidió que narraran por escrito lo que habían visto. Se diseñaron tablas de frecuencia y distribución para las formas de sujeto de tercera persona según EC y se llevaron a cabo modelos de regresión logístico donde se pudieran encontrar diferencias significativas.

Los datos corroboran la tendencia de los escritores de adherirse a la jerarquía, en particular en los estados a los extremos. Sin embargo, se observa un mayor número de frases nominales definidas en las narrativas escritas, lo cual sugiere que los poderes predictivos son más fuertes en las narrativas orales. Además, los hablantes L2 seleccionan formas de sujeto más en línea con el habla nativa en las narrativas escritas. No hay evidencia de un efecto cuantitativo del TC en la selección de las formas, pero se observa su codificación a través de otras estrategias lingüísticas.

Palabras claves: *adquisición de segunda lengua, cognición y memoria, expresiones referenciales, terreno común, español, discurso escrito*

COGNICIÓN Y SELECCIÓN DE SUJETOS: ANÁLISIS COMPARATIVO ENTRE
NARRATIVAS ESCRITAS Y ORALES EN EL DISCURSO DE ESPAÑOL COMO L2.

by

INGRID ABISAMBRA MICCHELI

B.A., University Externado de Colombia, Colombia, 1996

M.A., Georgia State University, 2016

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2022

© 2022
Ingrid Abisambra Miccheli
All Rights Reserved

COGNICIÓN Y SELECCIÓN DE SUJETOS: ANÁLISIS COMPARATIVO ENTRE
NARRATIVAS ESCRITAS Y ORALES EN EL DISCURSO DE ESPAÑOL COMO L2.

by

INGRID ABISAMBRA MICCHELI

Major Professor:	Margaret Lubbers Quesada
Committee:	Sarah. E. Blackwell
	Timothy M. Gupton

Electronic version approved:

Ron Walcott
*Vice Provost for Graduate Education
and Dean of the Graduate School*
The University of Georgia
August 2022

DEDICATION

To Oliver, because you fill my life with opportunities.

ACKNOWLEDGEMENTS

I cannot express enough gratitude to my committee chair, Dr. Margaret Lubbers Quesada, for her unwavering support and encouragement. She continually and convincingly guided me throughout my time at UGA, teaching me how to carry out my research and sharing her knowledge and experience with me. Without her persistent help this dissertation would not have been possible. Thank you for believing in me and for the kindness you have shown me.

I would like to thank my committee members, Dr. Sarah E. Blackwell and Dr. Timothy Gupton, who first introduced me to linguistics and fostered in me an enthusiasm with lasting effects. Thank you for teaching me and for your continued support.

It takes a village to achieve your most challenging goals, my completion of this project could not have been accomplished without the support of my classmates and friends, Darío, Camila, Jihee, Javier, Shannon, Jessica, David. You made the process fun, even during the toughest of times.

Thank you to my wonderful parents, Franca and Jaime, my beautiful stepdaughter, Isabelle, and my ever-supporting brother, Cristian, who don't always understand what I do, but nevertheless smile at me with pride.

Finally, and most importantly, a huge thank you to my husband and best friend, Oliver. You are my number one fan and you helped me create a space full of possibilities and opportunities. I have no doubt that I would not have taken the first step on this long road without you by my side, supporting me along the way with your infinite love for me.

TABLA DE CONTENIDOS

	Página
AGRADECIMIENTOS	v
LISTA DE TABLAS	ix
LISTA DE FIGURAS	xvii
CAPÍTULO	
1 INTRODUCCIÓN	1
1.1 Objetivos	1
1.2 Enfoque teórico	3
1.3 El lenguaje oral vs el lenguaje escrito	6
1.4 Escritura y memoria: ¿mayor costo cognitivo?	10
1.5 Descripción general de la disertación	15
2 EXPRESIONES REFERENCIALES DE SUJETO	19
2.1 Introducción	19
2.2 Propiedades sintácticas	21
2.3 Propiedades pragmáticas	30
2.4 Propiedades cognitivas	38
2.5 El terreno común (TC)	54
3 EL ESTUDIO DE LAS EXPRESIONES REFERENCIALES DE SUJETO EN LA ADQUISICIÓN DE SEGUNDA LENGUA.....	66
3.1 Introducción.....	66

3.2 La adquisición de sujetos en el generativismo.....	71
3.3 La adquisición de sujetos en el variacionismo.....	83
3.4 La adquisición de sujetos en el funcionalismo	91
3.5 La adquisición de sujetos en el cognitvismo	105
4 METODOLOGÍA.....	120
4.1 Introducción	120
4.2 Preguntas de investigación e hipótesis.....	121
4.3 Participantes	123
4.4 Instrumentos y recolección de datos	125
4.5 Codificación y metodología en el análisis de datos	128
5 RESULTADOS CUANTITATIVOS	137
5.1 Introducción.....	137
5.2 Resultados cuantitativos: estudio oral vs estudio escrito.....	138
5.3 Resultados cuantitativos: Terreno Común (TC) vs No Terreno Común (NTC).....	169
6 DISCUSIÓN Y ANÁLISIS CUALITATIVO	193
6.1 Introducción	193
6.2 Análisis y discusión INF.....	195
6. 3 Análisis y discusión ACTR	208
6.4 Análisis y discusión ACT	216
6.5 Análisis y discusión ACTNR	224
6.6 Análisis y discusión FAM	235
6.7 Análisis y discusión UNI	240

6.8 Análisis y discusión TYP	246
6.9 Conclusiones análisis cualitativo	256
7 CONCLUSIONES.....	258
7.1 Resumen	258
7.2 El discurso oral vs el discurso escrito	260
7.3 El papel del Terreno Común en la narrativa escrita	271
7.4 Observaciones generales con respecto a la relación entre la adquisición del español como L2 y la cognición	277
7.5 Limitaciones del estudio	281
7.6 Investigaciones futuras	282
BIBLIOGRAFÍA	284
APÉNDICES	
A Prueba adaptada DELE del Instituto Cervantes	296
B Expresiones generales — estado cognitivo ACTNR	297

LISTA DE TABLAS

	Página
Tabla 2.3.1: Distribución en español de los sujetos nulos y explícitos, según las nociones pragmáticas de tópico y foco	37
Tabla 3.2.1: Porcentaje de sujetos nulos en historias de tipo variable ligada y en historias referenciales en el estudio de la OPC	77
Tabla 3.5.1: Codificación y descripción de los estados cognitivos	110
Tabla 4.3.1: Hablantes nativos de español (L1)	124
Tabla 4.3.2: Hablantes no nativos de español (L2)	125
Tabla 4.5.1: Frecuencia y distribución de las formas de sujeto (todos los grupos) con GenNP.....	132
Tabla 4.5.2: Frecuencia y distribución de las formas de sujeto (todos los grupos) sin GenNP	133
Tabla 5.2.1: Comparación de instancias de expresiones referenciales para ECs entre datos orales (Blackwell y Lubbers Quesada) y datos escritos (presente estudio)	141
Tabla 5.2.2: Frecuencia y distribución de formas de sujeto por estado cognitivo (los dos grupos) en narrativas escritas (con GenNP)	142
Tabla 5.2.3: Frecuencia y distribución de formas de sujeto por estado cognitivo (los dos grupos) en narrativas escritas (sin GenNP)	144

Tabla 5.2.4: INF: Formas de sujeto entre L1 y L2 (instancias y porcentajes) en narrativas escritas	145
Tabla 5.2.5: INF: Regresión logística para Null, Overt, DefNP y PrpN entre L1 y L2 ..	146
Tabla 5.2.6: INF: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas ...	146
Tabla 5.2.7: INF: Regresión logística sobre Null, Overt, DefNP y PrpN para L1 entre narrativas orales y escritas	147
Tabla 5.2.8: INF: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre narrativas orales y escritas	147
Tabla 5.2.9: INF: Ajuste logístico sobre Null	148
Tabla 5.2.10: INF: Ajuste logístico sobre Overt	148
Tabla 5.2.11: INF: Ajuste logístico sobre DefNP	148
Tabla 5.2.12: INF: Ajuste logístico sobre PrpN	148
Tabla 5.2.13: ACTR: Formas de sujeto entre L1 y L2	149
Tabla 5.2.14: ACTR: Regresión logística para Overt y PrpN entre L1 y L2	150
Tabla 5.2.15: ACTR: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas	150
Tabla 5.2.16: ACTR: Regresión logística sobre Null, Overt, DefNP y PrpN para L1 entre narrativas orales y escritas	151
Tabla 5.2.17: ACTR: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre narrativas orales y escritas	151
Tabla 5.2.18: ACTR: Ajuste logístico sobre Null	151
Tabla 5.2.19: ACTR: Ajuste logístico sobre Overt	151
Tabla 5.2.20: ACTR: Ajuste logístico sobre DefNP	152

Tabla 5.2.21: ACTR: Ajuste logístico sobre PrpN	152
Tabla 5.2.22: ACT: Formas de sujeto entre L1 y L2	153
Tabla 5.2.23: ACT: Regresión logística sobre Null, Dem y PrpN entre L1 y L2	153
Tabla 5.2.24: ACT: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas	154
Tabla 5.2.25: ACT: Regresión logística sobre Null, Overt, Dem, DefNP y PrpN para L1 entre narrativas orales y escritas	154
Tabla 5.2.26: ACT: Regresión logística sobre Null, Overt, Dem, DefNP y PrpN para L2 entre narrativas orales y escritas	155
Tabla 5.2.27: ACT: Ajuste logístico sobre Overt	155
Tabla 5.2.28: ACT: Ajuste logístico sobre DefNP	155
Tabla 5.2.29: ACTNR: Formas de sujeto entre L1 y L2	156
Tabla 5.2.30: ACTNR: Regresión logística sobre Null, DefNP, PrpN y IndNP entre L1 y L2	156
Tabla 5.2.31: ACTNR: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas	157
Tabla 5.2.32: ACTNR: Regresión logística sobre Null, GenNP, DefNP y PrpN para L1 entre narrativas orales y escritas	158
Tabla 5.2.33: ACTNR: Regresión logística sobre Null, GenNP, DefNP y PrpN para L2 entre narrativas orales y escritas	158
Tabla 5.2.34: ACTNR: Ajuste logístico sobre GenNP	158
Tabla 5.2.35: ACTNR: Ajuste logístico sobre DefNP	158
Tabla 5.2.36: ACTNR: Ajuste logístico sobre PrpN	159

Tabla 5.2.37: FAM: Formas de sujeto entre L1 y L2	159
Tabla 5.2.38: FAM: Regresión logística sobre DefNP y PrpN entre L1 y L2	160
Tabla 5.2.39: FAM: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas	160
Tabla 5.2.40: FAM: Regresión logística sobre GenNP, DefNP y PrpN para L1 entre narrativas orales y escritas	161
Tabla 5.2.41: FAM: Regresión logística sobre GenNP, DefNP y PrpN para L2 entre narrativas orales y escritas	161
Tabla 5.2.42: FAM: Ajuste logístico sobre GenNP	161
Tabla 5.2.43: FAM: Ajuste logístico sobre DefNP	162
Tabla 5.2.44: FAM: Ajuste logístico sobre PrpN	162
Tabla 5.2.45: UNI: Formas de sujeto entre L1 y L2	163
Tabla 5.2.46: UNI: Regresión logística sobre PrpN entre L1 y L2	163
Tabla 5.2.47: UNI: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas	164
Tabla 5.2.48: UNI: Regresión logística sobre DefNP y PrpN para L1 entre narrativas orales y escritas	164
Tabla 5.2.49: UNI: Regresión logística sobre DefNP y PrpN para L2 entre narrativas orales y escritas	164
Tabla 5.2.50: UNI: Ajuste logístico sobre DefNP	165
Tabla 5.2.51: UNI: Ajuste logístico sobre PrpN	165
Tabla 5.2.52: TYP: Formas de sujeto entre L1 y L2	166
Tabla 5.2.53: TYP: Regresión logística sobre DefNP y IndNP entre L1 y L2	166
Tabla 5.2.54: TYP: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas	167

Tabla 5.2.55: TYP: Regresión logística sobre GenNP, DefNP y IndNP para L1 entre narrativas orales y escritas	167
Tabla 5.2.56: TYP: Regresión logística sobre GenNP, DefNP y IndNP para L2 entre narrativas orales y escritas	168
Tabla 5.2.57: TYP: Ajuste logístico sobre GenNP	168
Tabla 5.2.58: TYP: Ajuste logístico sobre DefNP	168
Tabla 5.2.59: TYP: Ajuste logístico sobre IndNP	168
Tabla 5.3.1: INF: Formas de sujeto entre TC y NTC	171
Tabla 5.3.2: INF: Regresión logística sobre Null, DefNP y PrpN entre TC y NTC	171
Tabla 5.3.3: INF: Formas de sujeto entre L1 y L2, TC y NTC	172
Tabla 5.3.4: INF: Regresión logística sobre Null, Overt, DefNP y PrpN para L1 entre TC y NTC	173
Tabla 5.3.5: INF: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre TC y NTC	173
Tabla 5.3.6: INF: Regresión logística sobre Null, Overt, DefNP y PrpN para TC entre L1 y L2	173
Tabla 5.3.7: INF: Regresión logística sobre Null, Overt, DefNP y PrpN para NTC entre L1 y L2	173
Tabla 5.3.8: INF: Ajuste logístico sobre Overt	174
Tabla 5.3.9: INF: Ajuste logístico sobre DefNP	174
Tabla 5.3.10: INF: Ajuste logístico sobre PrpN	174
Tabla 5.3.11: ACTR: Formas de sujeto entre TC y NTC	175

Tabla 5.3.12: ACTR: Regresión logística sobre Null, Overt, DefNP y PrpN entre TC y NTC	176
Tabla 5.3.13: ACTR: Formas de sujeto entre L1 y L2, TC y NTC	176
Tabla 5.3.14: ACTR: Regresión logística sobre Null, DefNP y PrpN para L1 entre TC y NTC	177
Tabla 5.3.15: ACTR: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre TC y NTC	177
Tabla 5.3.16: ACTR: Regresión logística sobre Null, DefNP y PrpN para TC entre L1 y L2	177
Tabla 5.3.17: ACTR: Regresión logística sobre Overt, DefNP y PrpN para NTC entre L1 y L2	178
Tabla 5.3.18: ACT: Formas de sujeto entre TC y NTC	178
Tabla 5.3.19: ACT: Regresión logística sobre Null, Overt, Dem, DefNP y PrpN entre TC y NTC	179
Tabla 5.3.20: ACT: Formas de sujeto entre L1 y L2, TC y NTC	179
Tabla 5.3.21: ACT: Regresión logística sobre Null, Overt, DefNP y PrpN para L1 entre TC y NTC	180
Tabla 5.3.22: ACT: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre TC y NTC	180
Tabla 5.3.23: ACT: Regresión logística sobre Overt, DefNP y PrpN para TC entre L1 y L2	180
Tabla 5.3.24: ACT: Regresión logística sobre Overt, Dem y PrpN para NTC entre L1 y L2	181

Tabla 5.3.25: ACTNR: Formas de sujeto entre TC y NTC	181
Tabla 5.3.26: ACTNR: Regresión logística sobre Null, DefNP, PrpN y IndNP entre TC y NTC	182
Tabla 5.3.27: ACTNR: Formas de sujeto entre L1 y L2, TC y NTC	182
Tabla 5.3.28: ACTNR: Regresión logística sobre DefNP, PrpN y IndNP para L1 entre TC y NTC	183
Tabla 5.3.29: ACTNR: Regresión logística sobre DefNP, PrpN y IndNP para L2 entre TC y NTC	183
Tabla 5.3.30: ACTNR: Regresión logística sobre DefNP, PrpN y IndNP para TC entre L1 y L2	183
Tabla 5.3.31: ACTNR: Regresión logística sobre DefNP, PrpN y IndNP para NTC entre L1 y L2	184
Tabla 5.3.32: FAM: Formas de sujeto entre TC y NTC	184
Tabla 5.3.33: FAM: Regresión logística sobre DefNP y PrpN entre TC y NTC	185
Tabla 5.3.34: FAM: Formas de sujeto entre L1 y L2, TC y NTC	185
Tabla 5.3.35: FAM: Regresión logística sobre DefNP y PrpN para L1 entre TC y NTC	186
Tabla 5.3.36: FAM: Regresión logística sobre DefNP y PrpN para L2 entre TC y NTC	186
Tabla 5.3.37: FAM: Regresión logística sobre DefNP y PrpN para TC entre L1 y L2 ..	186
Tabla 5.3.38: FAM: Regresión logística sobre DefNP y PrpN para NTC entre L1 y L2	186
Tabla 5.3.39: UNI: Formas de sujeto entre TC y NTC	187

Tabla 5.3.40: UNI: Regresión logística sobre DefNP y PrpN entre TC y NTC	1875
Tabla 5.3.41: UNI: Formas de sujeto entre L1 y L2, TC y NTC	188
Tabla 5.3.42: UNI: Regresión logística sobre DefNP y PrpN para L1 entre TC y NTC	189
Tabla 5.3.43: UNI: Regresión logística sobre DefNP y PrpN para L2 entre TC y NTC	189
Tabla 5.3.44: UNI: Regresión logística sobre DefNP y PrpN para TC entre L1 y L2 ...	189
Tabla 5.3.45: UNI: Regresión logística sobre DefNP y PrpN para NTC entre L1 y L2	189
Tabla 5.3.46: TYP: Formas de sujeto entre TC y NTC	190
Tabla 5.3.47: TYP: Regresión logística sobre DefNP y IndNP entre TC y NTC	190
Tabla 5.3.48: TYP: Formas de sujeto entre L1 y L2, TC y NTC	191
Tabla 5.3.49: TYP: Regresión logística sobre DefNP y IndNP para L1 entre TC y NTC	192
Tabla 5.3.50: TYP: Regresión logística sobre DefNP y IndNP para TC entre L1 y L2 .	192
Tabla 5.3.51: TYP: Regresión logística sobre DefNP y IndNP para NTC entre L1 y L2	192
Tabla 5.3.52: TYP: Ajuste logístico sobre IndNP	192
Tabla 6.7.1: Uso total de demostrativos en las narrativas escritas y orales, L1 vs L2, todos los ECs	243
Tabla 7.2.1: DefNP y PrpN vs Null y Overt en los ECs INF, ACT y ACTR	265
Tabla 7.2.2: Acercamiento al habla nativa por parte de los HNNs	267
Tabla Apéndice B: Expresiones generales utilizadas por HNs y HNNs: estado cognitivo ACTNR	297

LISTA DE FIGURAS

	Página
Figura 2.4.1: Escala de familiaridad	40
Figura 2.4.2: Escala de continuidad de tópico	43
Figura 2.4.3: “Initial Accessibility Marking”	44
Figura 2.4.6: “The Givenness Hierarchy”	47
Figura 2.5.1: Definición de "conocimiento compartido"	55
Figura 3.5.1: Givenness Hierarchy para el español	110

CAPÍTULO 1

1 INTRODUCCIÓN

1.1 Objetivos

El objetivo principal de esta investigación es el de realizar un análisis comparativo entre datos orales y escritos en narrativas espontáneas, donde podamos comprobar si existen diferencias cuantitativas y cualitativas en el uso y selección de expresiones referenciales de sujeto de tercera persona¹, según el nivel de atención y memoria que se presupone del lector. Al mismo tiempo, se busca identificar posibles diferencias en el uso y selección de dichas expresiones entre hablantes nativos (HNs) de español y hablantes no-nativos (HNNs) de español como L2, en las dos modalidades de producción: la modalidad de producción escrita y la modalidad oral.

Como objetivo secundario, en esta investigación se busca determinar hasta qué punto es posible manipular el terreno común (TC) en las narrativas escritas como variable independiente, y establecer su posible influencia en la selección de expresiones referenciales de sujeto por parte de los escritores, tanto nativos como no-nativos de español. Entendemos por TC aquella parte compartida entre el hablante/escritor y el oyente/lector de sus conocimientos, creencias y suposiciones (Carlson, 1992:60), la cual se establece a través de diferentes procesos de copresencia, descritos en el Capítulo 2 del

¹ Los pronombres de primera y segunda persona se consideran déicticos o con un valor enfático.

presente trabajo, los cuales incluyen la pertenencia a una comunidad, la simultaneidad física y la copresencia lingüística o textual.

Se analiza todo el espectro de expresiones referenciales de sujeto disponible a los hablantes de español, lo cual incluye el sujeto nulo (\emptyset), los pronombres personales (*él, ella, ellos*, etc.), los demostrativos (*este, ese, aquel*, etc.), los nombres propios (*Juan*) y las frases definidas e indefinidas (*la mujer/una mujer*). Las expresiones referenciales son ideales para mostrar la importancia de la cognición en el lenguaje, puesto que son construcciones lingüísticas usadas con el objetivo comunicativo de llevar la atención del oyente hacia el mismo marco de atención del hablante con respecto a la “cosa” a la cual hacemos referencia; en otras palabras, aquello acerca de lo cual predicamos algo (Tomassello, 2005:241). Las lenguas tienen diferentes herramientas convencionales para cumplir con este propósito, por ejemplo, sujetos nulos, pronombres personales y frases nominales. Los procesos necesarios para identificar, seleccionar y utilizar una determinada expresión referencial de sujeto en el discurso han sido objeto de estudios extensivos por parte de lingüistas, psicolingüistas, psicólogos y filósofos, quienes han aportado diferentes contribuciones teóricas y empíricas desde sus diversas perspectivas teóricas. A partir de los estudios generativos de los años 50 comienza la plétora de investigaciones que buscan entender la distribución entre el sujeto nulo y explícito en las lenguas que permiten el sujeto nulo, entre las cuales se encuentra el español. Subsecuentes marcos teóricos como el variacionista, el funcionalista y el cognitivista, proponen factores cognitivos y pragmáticos, así como extralingüísticos, que afectan la selección y uso de las expresiones referenciales de sujeto. El común denominador entre todos estos estudios es que, en su gran mayoría, los datos utilizados para el análisis

proviene de datos orales, bien sean datos artificiales como en el caso de los estudios generativistas, en los cuales es el mismo lingüista el que fabrica los enunciados bajo estudio y solicita a los hablantes juzgar la gramaticalidad o aceptabilidad de dichos enunciados, o bien sean datos naturales obtenidos a través de investigaciones empíricas, así como aquellos provenientes de corpora.

A través de la presente investigación se busca llenar el vacío existente con respecto al análisis del discurso escrito como modalidad válida de producción. Se toman como base aquellos factores cognitivos y pragmáticos que parecen restringir las opciones de los escritores con respecto a la selección y uso de las expresiones referenciales de sujeto en el español. Esperamos encontrar diferencias significativas entre los datos de ambos tipos de modalidad, la oral y la escrita, y que estos se deben principalmente a factores cognitivos de atención y memoria. En segundo lugar, se prevé que los HNNs demostrarán un acercamiento al habla nativa debido a la influencia directa de la escritura, ya que esta modalidad permite un mayor grado de cuidado en la realización de las expresiones seleccionadas. Por último, esperamos encontrar diferencias entre las narrativas que presentan TC versus aquellas donde no se asume TC entre el escritor y el lector, pero estas diferencias provendrán principalmente de los HNs debido a la mayor facilidad que tienen estos hablantes de incorporar información pragmática con información sintáctica y semántica en el discurso.

1.2 Enfoque teórico

En esta investigación se parte de la base de que el discurso es un proceso cognitivo que involucra la interacción, la adaptación y la negociación entre los interlocutores, además de ser un fenómeno social (Virtanen, 2004:3). Virtanen dice que las

presuposiciones que hacen los interlocutores en el discurso tienen restricciones mentales y de memoria que se manifiestan en el texto (p. 4). Esta posición va en línea con Tomasello (2005), quien habla de la atención que se requiere en el proceso referencial. Tomasello explica que al seleccionar una expresión lingüística que simboliza un evento referencial, el hablante hace contacto de varias maneras con el marco de atención que comparte con el oyente en el momento del enunciado, el cual se crea durante la interacción entre los interlocutores. Añade que algunos de estos contactos están más enfocados en el oyente dependiendo del conocimiento mutuo que el hablante presupone, las expectativas y las perspectivas que se tienen acerca del enunciado. Por ejemplo, el hablante seleccionará un pronombre sobre una frase nominal para referirse a un objeto en el foco de atención inmediato entre los interlocutores (Tomasello, 2005:197). Tanto Virtanen (2004) como de Beaugrande (2004) hablan acerca de este tipo de análisis del discurso como lingüística cognitiva textual (*cognitive text linguistics*). De Beaugrande explica que la lingüística textual siempre ha tenido una orientación cognitiva porque el texto debe describirse tanto como un producto así como un proceso (p. 27).

Como apunta Tomasello (2005), el argumento principal de la lingüística cognitivo-funcionalista² es que el objetivo fundamental del lenguaje es la comunicación y por lo tanto la estructura del lenguaje es simbólica, donde un símbolo es una unidad compuesta tanto de forma como de significado. Los símbolos son convenciones sociales a través de

² Aunque existen algunas diferencias entre el cognitivismo y el funcionalismo como enfoques lingüísticos, principalmente debido a la identidad científica que se le ha atribuido a la lingüística cognitiva, en general son más las similitudes y esto hace que sea difícil trazar una clara separación, llevando a muchos investigadores a calificarse a sí mismos como *cognitivo-funcionalistas* (Nuyts, 2010). La posición que tomo en esta investigación es una de integración entre estos dos enfoques, puesto que los procesos para la adquisición de lengua provienen de habilidades mentales relacionadas con la detección de las intenciones de los hablantes (*intention-reading*) y de detección de patrones (*pattern-finding*) (Tomasello, 2005), habilidades reconocidas tanto por el cognitivismo como el funcionalismo.

los cuales un individuo pretende llamar la atención de otro individuo dirigiendo la atención compartida o el estado mental del otro individuo hacia algo en el mundo exterior (Tomasello, 2005). Según Tomasello, es esta dimensión mental lo que les da a los símbolos lingüísticos el poder comunicativo, exclusivamente humano, de hacer referencia a y predicar algo acerca de un objeto, un evento o una situación en el mundo real. El análisis del discurso que parte desde este enfoque incluye nociones pragmáticas de tópico (información compartida/conocida), foco (información nueva) y TC. Investigadores como Givón (1993), Ariel (1988) y Gundel, Hedberg y Zacharski (1993), entre otros, dicen que los hablantes deben evaluar la disponibilidad cognitiva del referente para el oyente en cuanto a su accesibilidad, topicalidad y nivel de “lo dado” (*givenness*). Por lo tanto, las expresiones referenciales que selecciona un hablante se ven restringidas no solamente por factores sintácticos y cognitivos, sino también por estos factores pragmático-discursivos. En el presente trabajo se tienen en cuenta estas nociones, ya que hacen parte del marco de atención compartido entre los interlocutores y serán descritas en detalle en los Capítulos 2 y 3.

La hipótesis de posición del antecedente (PAH)³ de Carminati (2002) es un modelo sintáctico que se tiene en cuenta en la presente investigación ya que complementa los enfoques cognitivo y pragmático. Esta hipótesis plantea que la prominencia sintáctica del sujeto impone una restricción en la interpretación del pronombre anafórico y predice que un sujeto nulo prefiere recuperar un antecedente que se encuentre en la posición de sujeto, mientras que un pronombre explícito prefiere un antecedente en una posición más

³ Uso de la abreviatura PAH por el inglés Position of Antecedent Hypothesis, también conocida como PAS por Position of Antecedent Strategy.

baja en la estructura sintáctica (p.ej. posición de objeto). La PAH será descrita en más detalle en el Capítulo 2.

Como veremos a lo largo de este trabajo, hay una interacción directa entre los factores sintácticos relacionados con la posición pre o posverbal del sujeto en una oración y las nociones pragmáticas de tópico y foco. Dicha interacción es de particular relevancia en el estudio de la adquisición de una segunda lengua ya que se ha demostrado que los factores pragmáticos son de más difícil adquisición que los factores sintácticos (Sorace y Filiaci, 2006, Montrul y Rodríguez Louro, 2006, Lozano 2008, *inter alia*).

No parece haber hasta el momento ningún estudio de adquisición de segunda lengua para el español que tenga como marco teórico este enfoque cognitivo-funcionalista que he brevemente descrito en esta sección, donde se comparen las modalidades de producción oral y escrita. Esta disertación pretende llenar ese vacío con el propósito de contribuir al estudio del análisis del discurso y la cognición.

1.3 El lenguaje oral vs el lenguaje escrito

Para aquellos que vemos la escritura como una expresión lingüística con un alto valor de estudio por derecho propio, la primera pregunta que surge es por qué se ha relegado a un segundo plano en las investigaciones lingüísticas. En palabras de Olson, “it has been assumed that writing is a graphic device for transcribing speech” (1996:83) y la explicación de esta situación tiene raíces históricas. No es el objetivo de esta disertación el de presentar un recuento histórico acerca de la evolución de la escritura, pero es relevante notar cómo ha evolucionado esta posición desde lingüistas como Ferdinand de Saussure, Leonard Bloomfield y Noam Chomsky hasta hoy en día, con proponentes de

la escritura como modalidad digna de estudio con lingüistas como David R. Olson, Elinor Ochs, Deborah Tannen y Wallace L. Chafe.

Olson (1996) y Chafe (1994) describen cómo en la antigua tradición griega, la escritura tenía importancia prioritaria y Chafe añade que el lenguaje oral se veía como una versión corrupta (1994:45). Sin embargo, el estatus de la escritura fue reducido a un segundo lugar a partir de la lingüística moderna hacia finales del siglo XIX y principios del siglo XX, en parte debido a las opiniones de influyentes lingüistas como Saussure y Bloomfield, a quienes Olson y Chafe citan de las siguientes maneras: “The sole reason for the existence of writing is to represent speech. The linguistic object is not defined by the combination of the written word and the spoken word: the spoken form alone constitutes the object” (Saussure, 1916/1983, citado en Olson, 1996:85); “Language and writing are two distinct systems of signs; the second exists for the sole purpose of representing the first” (Saussure, 1916/1959 citado en Chafe, 1994:45). Por su parte, Bloomfield redujo la escritura a una forma de transcripción sin valor propio: “Writing is not language, but merely a way of recording language” (Bloomfield 1933, citado en Chafe, 1994:46 y en Olson, 1996:85). Estas opiniones fueron luego adoptadas por Chomsky (1965, citado en Chafe, 1994:47), quien describe a la teoría lingüística como principalmente preocupada por el lenguaje ideal entre un hablante y un oyente, y es a partir de esta posición que los lingüistas fueron llevados a dedicar sus esfuerzos a estudiar la gramaticalidad (o falta de esta) del lenguaje oral a partir de pseudo-oraciones.

Como mencionado por Olson (1996:89), ningún sistema de escritura puede transcribir a la perfección el sistema oral. Esto se debe simplemente a que el alfabeto carece de fuerza ilocucionaria, de entonación y de otros elementos prosódicos que hacen

más compleja la interpretación de un texto escrito. Aun así, a partir de los años 60 y 70 varios autores han reconocido en la escritura un sistema que afecta diferentes procesos en la cognición humana y que por lo tanto impone sus propias restricciones, diferentes al lenguaje oral. A continuación, destacamos algunas de las principales contribuciones a este subcampo de la lingüística.

En primer lugar, resalta el trabajo de Chafe (1994), quien analiza las modalidades oral y escrita como una experiencia que afecta el flujo de la consciencia y describe varias de las diferencias entre los dos modos de producción desde este enfoque teórico y filosófico. Algunas de estas diferencias son el nivel de espontaneidad, la evanescencia del lenguaje oral vs la permanencia del lenguaje escrito, las diferencias temporales (la inmediatez del lenguaje oral vs el desplazamiento del lenguaje escrito), la riqueza prosódica del lenguaje oral vs la pobreza de la representación prosódica del lenguaje escrito, el nivel de naturalidad y la copresencia. Aunque para Chafe la modalidad oral es la base del lenguaje, también el lenguaje escrito tiene la posibilidad de llevar a la consciencia experiencias exclusivas relativas a su producción.

La posición de Chafe no es del todo diferente a la de Olson (1996, 2016), para quien la escritura no es meramente la transcripción del lenguaje oral y, así como Chafe, Olson propone que la modalidad escrita es un tipo de lenguaje que contribuye a la construcción de los modelos conceptuales del lenguaje; es decir, que tiene un efecto cognitivo (1996:99). Olson (2016) dice que la escritura provee un grado más alto de consciencia del lenguaje que el habla, ya que es solo a través de la escritura que podemos *ver* el lenguaje, hablar de las palabras y de las oraciones, su longitud, su forma y

podemos comparar los diferentes significados. En otras palabras, la escritura permite el metalenguaje al hacer el lenguaje explícito en nuestras consciencias.

En los estudios de Ochs (1979) se plantea una distinción funcional entre el discurso planeado y no-planeado y sugiere que las dos modalidades del lenguaje presentan características de los dos tipos de discurso; por ejemplo, es posible un lenguaje escrito no-planeado en el caso de cartas personales y diarios, así como un lenguaje oral planeado, en el caso de charlas y seminarios formales. Para el presente trabajo se analizan narrativas escritas espontáneas, las cuales corresponden con un lenguaje no-planeado, siguiendo la distinción de Ochs. Este tipo de lenguaje es descrito por Ochs de la siguiente manera: “The discourse must lack forethought and prior organization” (1979:78); en otras palabras, se requiere que el hablante o escritor no tenga tiempo de planear y revisar su narrativa.

Por último, no cabe duda de que Tannen (1980, 1982a, 1982b), en su extensiva investigación en el estudio de la escritura, ha contribuido altamente al desarrollo de nuestro entendimiento de esta modalidad. Tannen presenta las diferencias entre las dos formas de producción y propone lo que ella denomina “el continuum oral/letrado en el discurso” (*the oral/literate continuum in discourse*) (Tannen, 1982a). Este continuum refleja el grado de implicación versus contenido que está presente en los dos tipos de lenguaje y concluye que, en la escritura, los procesos se enfocan en el contenido del mensaje, ignorando las dinámicas interpersonales que ocurren en el discurso oral. En un trabajo conjunto, Chafe y Tannen (1987) revisaron una gran cantidad de estudios que han analizado el lenguaje escrito. En su mayoría, los trabajos comparan las diferencias estructurales entre el lenguaje oral y el lenguaje escrito, mientras que otros se enfocan en

el desarrollo de la escritura en los niños. Ninguno de estos estudios se basa en el español o en el impacto de la escritura en la adquisición de una segunda lengua.

1.4 Escritura y memoria: ¿mayor costo cognitivo?

Las investigaciones que se han llevado a cabo acerca del costo de procesamiento de la memoria en la escritura provienen del campo de la psicología con investigadores como McCutchen (1996), Kellogg (2001) y Bourdin y Fayol (1994, 2002), entre otros.

McCutchen (1996) dice que el problema de las limitaciones de la capacidad mental durante el proceso de escritura debe considerarse dentro del más amplio contexto de la teoría de la memoria, según la cual información acerca del entorno e información proveniente de la memoria a largo plazo se encuentra almacenada en la memoria operativa (a corto plazo) durante el procesamiento. Debido a los recursos limitados del sistema de memoria, existe un intercambio entre el almacenamiento en la memoria operativa y las funciones de procesamiento; en otras palabras, entre más recursos se necesiten durante el procesamiento, habrá menos recursos disponibles para el almacenamiento. McCutchen añade que, para componer un texto, el escritor debe coordinar entre los objetivos de planeación en la memoria operativa (p. ej. planear por contenido, el tono, el tipo de audiencia) y los objetivos del producto en sí (p. ej. la gramaticalidad), al mismo tiempo que los procesos de generación del lenguaje recuperan palabras para expresar el contenido y organizan estas palabras en un texto apropiado. McCutchen concluye que la comprensión y la composición escrita imponen demandas considerables en el procesamiento y almacenamiento y, si una de estas dos funciones está comprometida, la ejecución se verá afectada. Sin embargo, tanto McCutchen (1996) como Bourdin y Fayol (1994, 2002) dicen que son los escritores menos sofisticados

(definidos por estos autores como escritores más jóvenes o con poca experiencia escribiendo) los que demostrarán un efecto negativo en la expresión del lenguaje escrito; mientras que en los escritores adultos no se perciben diferencias significativas entre la producción del lenguaje oral y el lenguaje escrito.

A este respecto, Bourdin y Fayol (1994, 2002) han conducido varios estudios que tienen como objetivo determinar si existen diferencias en la carga cognitiva entre el lenguaje oral y el escrito. Estos investigadores parten de tres premisas: (1) la capacidad de la memoria operativa es limitada, (2) cada proceso en la escritura tiene una carga cognitiva y (3) cada incremento en la carga cognitiva de una determinada actividad de uno de los procesos lleva a una reducción en los recursos disponibles para los otros componentes (1994:591). Estos procesos o componentes se categorizan como planeación, traducción y revisión. En el proceso de planeación se generan y organizan las ideas, que luego se traducen a oraciones que componen el texto y el cual incluye, en la escritura, una fase de ejecución motora; por último, el proceso de revisión incluye la lectura del texto y la detección de errores o problemas (Kellogg, 2001:176). Todos estos procesos consumen recursos y estos recursos provienen de la misma fuente; sin embargo, el costo cognitivo es imperceptible en adultos (Bourdin y Fayol, 1994:595). Esta es la conclusión a la que llegaron Bourdin y Fayol (1994, 2002) después de analizar los datos obtenidos en sus estudios al respecto de la carga cognitiva en la escritura. Específicamente, el trabajo publicado en el 2002 buscaba confirmar los resultados obtenidos en la investigación del 1994, según los cuales los adultos han automatizado la escritura a un nivel suficiente como para que el texto no se vea afectado por el mayor costo de procesamiento. Sin embargo, cuando hay una carga mayor en alguno de los procesos, por ejemplo, al aplicar

límites en el tiempo de la composición o al manipular la memoria a largo plazo, los recursos que se necesitan en la memoria operativa se ven reducidos y esto afecta la calidad del texto producido. La calidad se midió con respecto al número de ideas elaboradas y a la coherencia de los textos (Bourdin y Fayol, 2002).

No hay muchos estudios lingüísticos que analicen el costo cognitivo de la escritura, y ninguno de ellos se enfoca al mismo tiempo en la lingüística cognitiva, en el español y en la comparación de textos escritos con textos orales. Sin embargo, brevemente describo a continuación los únicos tres estudios encontrados que de cierta forma son pertinentes para el presente estudio.

En primer lugar, tenemos el trabajo de Schoonen, van Gelderen, de Gloppe, Hulstijn, Simis, Snellings y Stevenson (2003) que analiza las diferencias entre composiciones escritas en L1 y L2. En esta investigación se exploró la importancia del conocimiento lingüístico, el conocimiento metacognitivo (conocimiento acerca de lo que constituye un buen texto, así como de las estrategias de escritura necesarias) y la accesibilidad al conocimiento lingüístico de la L1 (en este caso, el holandés) y de la L2 (el inglés). Los participantes fueron estudiantes de segundo grado de una escuela secundaria en Holanda (lo cual corresponde con los 14 años de edad), con un promedio de 3.5 años de instrucción de inglés como L2. Schoonen et ál. afirman que escribir en una L2 requiere una mayor demanda cognitiva debido a la más baja competencia lingüística; desafortunadamente, estos autores no mencionan diferencias específicas según el nivel de adquisición (principiantes, intermedios, avanzados), aunque sí dicen que, entre mayor instrucción en la L2, mayor es la fluidez escrita (p. 8). Es relevante notar en los resultados que el nivel de competencia escrita en la L2 está altamente correlacionado con

el nivel de competencia escrita en la L1; al mismo tiempo, los resultados demuestran que al escribir en la L2, las variables lingüísticas parecen tener un papel mayor que en la L1, lo cual indica que los estudiantes estaban más preocupados con resolver problemas lingüísticos en la L2 que en la L1; mientras que en la L1 su mayor preocupación era de acceso al conocimiento metacognitivo (p. 29).

El segundo trabajo a resaltar, es un estudio de adquisición de español como segunda lengua donde se comparan textos escritos con textos orales, y es el de García-Alcaraz y Bel (2011). Este estudio tiene como base la PAH Carminati (2002) y los trabajos de Alonso-Ovalle, Fernández-Solera, Frazier y Clifton (2002) (ver Capítulo 2 para una descripción detallada de los trabajos de Carminati, 2002 y de Alonso-Ovalle et ál, 2002). García-Alcaraz y Bel (2011) analizan los pronombres personales de tercera persona, nulo y explícito, en textos de producción semi-espontánea entre hablantes nativos de español (nativos de Córdoba, España), hablantes bilingües de español-árabe de origen inmigrante y aprendices de español como L2 con árabe como L1, ambos grupos residentes en Cataluña. Entre las preguntas de investigación formuladas por los autores, se buscaba analizar las posibles diferencias en la influencia de la modalidad escrita y oral en la selección y uso de los pronombres bajo investigación. Según estos autores, existe una mayor “presión procedimental” en la modalidad oral, aunque esta aseveración no es sustentada con ningún tipo de explicación, evidencia o citas a lo largo del estudio y va en contra de los resultados encontrados por Bourdín y Fayol (1994, 2002) mencionados más arriba. El resultado con mayor relevancia para mi investigación es el hecho de que se observa un aumento en el uso del pronombre explícito en las narrativas escritas por parte

de los hablantes bilingües. Sin embargo, los autores notan que este resultado no es significativo debido al bajo número de instancias que se produjeron.

El tercer estudio que analiza las dos modalidades del lenguaje, aunque únicamente entre HNs, es el de Bel, Perera y Salas (2010). En este estudio el enfoque es la resolución anafórica en el catalán, en contextos intra-oracionales e inter-oracionales en el discurso, también con base en la PAH de Carminati (2002). Uno de los objetivos del estudio es la comparación de las narrativas escritas y orales en textos semi-espontáneos, para determinar si el modo de producción tiene una influencia en la selección entre los sujetos nulos y los sujetos explícitos (pronombres y frases nominales). Participaron únicamente HNs entre los nueve y los 16 años de edad. Aunque los resultados indican que no se encontraron diferencias significativas, el papel de la modalidad se vuelve relevante cuando interactúa con la edad de los participantes, confirmando la hipótesis de Bourdin y Fayol (1994, 2002), según quienes las demandas del costo cognitivo son mayores en la modalidad escrita, pero únicamente en los niños más jóvenes. Las mayores diferencias se encontraron en el uso de las frases nominales, las cuales fueron la forma más difícil de integrar en el discurso en el proceso de resolución anafórica. Para todas las edades, en las narrativas escritas se ve un aumento de frases nominales y una disminución del sujeto nulo. Las autoras concluyen que, a nivel intra-oracional, los sujetos nulos tienden a hacer correferencia con antecedentes en posición de sujeto, mientras que los sujetos plenos (p.ej.: nombres y frases nominales) pueden tener ya sea un antecedente en posición de sujeto o en otra posición oracional, resultados en línea con la investigación de Alonso-Ovalle et ál. (2002) para el español. Los resultados son consistentes en las dos modalidades de producción, oral y escrita.

La relevancia de los anteriores estudios para la presente investigación recae en los resultados que se obtuvieron con respecto al costo cognitivo entre la producción oral y escrita y a las selecciones que hacen los hablantes de las expresiones referenciales de sujeto, según el modo de producción. En esta disertación se toman como base los resultados obtenidos por McCutchen (1996) y Bourdin y Fayol (1994, 2002), según los cuales los adultos no demuestran diferencias en la producción de las narrativas escritas con respecto a la producción oral, a pesar de la mayor carga de procesamiento cognitivo atribuida al lenguaje escrito. Sin embargo, en el presente estudio hay dos factores que pueden afectar la carga cognitiva y en consecuencia la calidad del texto: en primer lugar, está el hecho de que a los participantes se les dio un tiempo límite dentro del cual debían redactar su narrativa (ver Capítulo 4 para la descripción de la metodología utilizada) y, en segundo lugar, se espera una mayor carga cognitiva en los HNNs que en los HNs al redactar un texto en una L2, en línea con los resultados obtenidos por Schoonen et ál. (2003). También, veremos que los resultados obtenidos en este trabajo concuerdan en gran medida con aquellos de García-Alcaraz y Bel (2011) y de Bel et ál. (2010) en cuanto a un mayor uso de pronombres explícitos y de expresiones referenciales más elaboradas en las narrativas escritas en los HNs versus las narrativas orales.

1.5 Descripción general de la disertación

En este estudio se utiliza como marco de referencia la Givenness Hierarchy⁴ propuesta por Gundel, Hedberg y Zacharski (1993) donde la forma de la expresión referencial depende del estado cognitivo que se asume del referente intencionado por

⁴ La Givenness Hierarchy puede traducirse al español como “la jerarquía de lo dado”; sin embargo, en esta disertación se mantiene el título en inglés, puesto que así aparece en la mayoría de la literatura revisada para este trabajo.

parte del hablante. Los factores que intervienen son el nivel de atención y memoria (operativa y a largo plazo) con respecto a dicho referente, factores de naturaleza cognitivos. Según Gundel (2010:154), la mayoría de los criterios para determinar los estados cognitivos se ven respaldados por investigaciones psicolingüísticas. Por ejemplo, el criterio de que se encuentran en foco entidades introducidas sintácticamente en una posición prominente, en una oración inmediatamente anterior hace posible y apropiado el uso de formas mínimas (sujetos nulos y pronombres explícitos).

La jerarquía fue posteriormente adaptada al español por Blackwell y Lubbers Quesada (2012), en cuyo trabajo se buscó examinar si la jerarquía aplica de igual manera a los aprendientes de español como segunda lengua. Blackwell y Lubbers Quesada (de aquí en adelante ByLQ) compararon narrativas orales de estudiantes de español de diferentes niveles de competencia lingüística, con narrativas de HNs de español. En este trabajo se lleva a cabo una comparación entre los datos orales de ByLQ y los datos escritos recogidos para la presente investigación.

En la presente investigación participaron un total de 40 personas: 20 HNs de español y 20 HNNs de español de nivel muy avanzado. La recolección de datos se hizo a través de una encuesta en línea donde la tarea principal era pedir a los participantes que vieran los primeros minutos de la película muda de Charlie Chaplin “A Woman” (1915) y que escribieran lo que habían visto. Esta es la misma película que vieron los participantes del estudio de ByLQ. Se prepararon cuatro encuestas con algunas diferencias en las instrucciones para reflejar narrativas con TC y sin TC, entre escritor y lector. En el estudio oral de ByLQ no se aisló el TC como variable, por lo tanto, la comparación entre HNs y HNNs con respecto a este factor se llevó a cabo únicamente entre las

composiciones escritas. Para la presente investigación participaron un total de 40 personas divididos entre 20 HNs de español y 20 HNNs de español L2 de nivel muy avanzado, con inglés como L1.

El estudio busca contestar las siguientes dos preguntas de investigación, con sus respectivas sub-preguntas. En primer lugar, cómo se comparan las narrativas derivadas de dos medios de producción distintos: producción oral y producción escrita, y establecer posibles diferencias entre los HNs los HNNs de español. En segundo lugar, se busca determinar si es posible manipular el TC para aumentar o disminuir el nivel de atención y memoria de los participantes en las narraciones escritas y evaluar si hay diferencias en las narrativas escritas entre HNs y HNNs, dependiendo del factor TC.

Para poder hacer una comparación adecuada entre ambos estudios, en el presente trabajo se mantiene una metodología muy similar a la del estudio de ByLQ. Específicamente, los procesos de codificación de las expresiones referenciales fueron los mismos y se utilizaron equivalentes herramientas estadísticas (ver Capítulo 4 para una descripción de la metodología utilizada). Sin embargo, para poder comparar los datos del presente estudio con aquellos de ByLQ, en colaboración con un equipo de estadísticos se diseñó un modelo de regresión logística que pudiera evidenciar los posibles efectos de los factores Producción (oral/escrito) y Grupo (L1/L2) en las selecciones de las expresiones referenciales de sujeto en las narrativas.

Esta disertación está organizada como sigue: en el Capítulo 2 se revisa la literatura previa con respecto a las propiedades sintácticas, pragmáticas y cognitivas de expresiones referenciales de sujeto que se han identificado en las principales investigaciones hechas hasta el momento. En la última sección del mismo capítulo resaltaré el papel que juega el

TC durante el acto comunicativo. En el Capítulo 3 se describen algunos de los estudios más relevantes que reflejan los diferentes enfoques lingüísticos dentro de la adquisición de segunda lengua (ASL), con un énfasis en la adquisición de sujetos en el español como L2. Estos enfoques son el generativismo, el variacionismo, el funcionalismo y el cognitivismo. En el Capítulo 4 se presenta la metodología del presente trabajo, la cual incluye las preguntas de investigación y correspondientes hipótesis, la descripción de los participantes, los procesos para la recolección de datos, el proceso de codificación de estos, las pruebas y, por último los análisis. En el Capítulo 5 se presentan los resultados cuantitativos y en el Capítulo 6 el análisis cualitativo. Por último, el Capítulo 7 resume los resultados obtenidos en las conclusiones, así como los argumentos presentados y se proponen áreas de investigación para futuros trabajos.

CAPÍTULO 2

2 EXPRESIONES REFERENCIALES DE SUJETO

2.1 Introducción

Durante el proceso comunicativo, bien sea al hablar o al escribir, proporcionamos información acerca de individuos y de objetos. Las lenguas naturales poseen herramientas que nos permiten identificar dichas entidades en el mundo real, o en el discurso. Estas herramientas son expresiones lingüísticas que hacen posible la función referencial del lenguaje, a través de la cual un hablante puede designar al referente que denotan. Por lo tanto, el objetivo de una expresión referencial de sujeto es el de identificar una entidad en particular. Según Jorgensen (1998:99-100), el uso de una expresión referencial en un acto comunicativo implica que el oyente hará una representación mental que denota la entidad a la cual hace referencia el hablante, de tal manera que el oyente pueda reconocer el objeto intencionado por el hablante. Esto quiere decir que el referente no es el objeto o la persona en el mundo real de manera directa, sino una representación mental del mismo.

Las expresiones referenciales pueden ser sintagmas nominales encabezados por artículos definidos o indefinidos (*el libro/un libro*), pronombres personales (*él/ella*), pronombres nulos (*pro*), nombres propios (*Juan*), expresiones demostrativas (*este/ese/aquel*), y en algunos casos, sintagmas adverbiales (*por mi parte*) (Matos Amaral y Schwenter 2005:119). Su función puede ser tanto no anafórica como anafórica, en donde estas últimas hacen referencia a una entidad previamente mencionada o implicada en el discurso. Levinson (2000) define la anáfora como “aquel fenómeno en donde una

expresión lingüística (la anáfora), al no tener una referencia independiente, puede seleccionar una referencia o una interpretación a través de una conexión con otra expresión lingüística (usualmente el antecedente)” (2000:267, mi traducción). En otras palabras, el hablante utiliza la anáfora como una forma lingüística que el oyente deberá utilizar para poder identificar el antecedente (el referente de la expresión anafórica). En (1) el pronombre *él* es anafórico porque designa al mismo referente de la oración, *Juan*, mientras que *Juan* es una referencia no anafórica puesto que su interpretación no depende de una interpretación de correferencia con un antecedente en el discurso anterior (los subíndices representan correferencia):

(1) *Juan_i* sabe lo que *él_i* quiere.

La relación que se establece entre el referente y la expresión lingüística conlleva un procesamiento cognitivo en tanto que las expresiones referenciales, al restringir el proceso de asignación del referente, ayudan a orientar “en la tarea de búsqueda y asignación del referente” (Figueras, 2002:55) en la mente de los hablantes. Este proceso incluye la selección, uso e interpretación de una expresión dada, y su estudio ha incluido investigaciones extensivas tanto en la L1 como en la L2. Se han definido propiedades lingüísticas y extralingüísticas que influyen en las formas utilizadas para codificar e interpretar las formas anafóricas y no anafóricas y, aunque algunas de estas propiedades parecen ser universales, también se han reconocido algunas específicas a cada lengua (Lubbers Quesada, 2015:20).

En las siguientes secciones describiré las propiedades sintácticas, pragmáticas y cognitivas de las expresiones referenciales de sujeto que se han identificado en las principales investigaciones hechas hasta el momento, concentrándome en la L1, con un

énfasis en el español. En la última sección resaltaré el papel que juega el terreno común (TC) en la codificación de la expresión referencial de sujeto seleccionada durante un acto comunicativo.

2.2 Propiedades sintácticas

Los estudios sintácticos son principalmente estudios derivados de la teoría lingüística generativa, la cual fue desarrollada primero en los años 50 por Chomsky (Rothman y Pascual y Cabo, 2014:46-47). En esta teoría se plantea que los humanos deben tener una estructura innata mental que les permite adquirir una lengua a pesar de la pobreza de estímulos (input incompleto y ambiguo que reciben los niños). Este conocimiento innato ocupa un lugar modular e independiente del resto de las otras capacidades cognitivas del ser humano. Al problema de la adquisición relacionada con la pobreza de estímulos se le denomina el problema lógico (Sánchez y Toribio, 2003) y la gramática universal (GU) surge como solución a este problema. La GU consta de principios universales que son verdaderos para todas las lenguas, y las diferencias se reducen a un grupo pequeño de parámetros, entre los cuales se encuentra el parámetro del sujeto nulo (PSN).

El español se considera una lengua pro-drop; es decir, el sujeto referencial puede ser nulo o explícito. El inglés no comparte esta propiedad del español, ya que es una lengua de débil concordancia verbal, y por lo tanto no permite el sujeto nulo, salvo en cláusulas paralelas de coordinación (Lubbers Quesada, 2015:21-22). Los primeros estudios generativos intentaban identificar un conjunto de propiedades relacionadas con el PSN. Dichos estudios formales se centraban en determinar la distribución del sujeto

nulo vs el sujeto pronominal explícito y, en general, no consideraban otras expresiones lingüísticas de sujeto, como lo son las frases nominales léxicas o los nombres propios. Inicialmente, Rizzi (1982, citado en Rothman e Iverson, 2007:329) propuso una serie de propiedades asociadas con el PSN. Estas propiedades son las siguientes (el símbolo * indica que una oración es agramatical):

1. El sujeto referencial puede ser explícito o nulo:

Yo vivo en Miami / *Pro* vivo en Miami

2. El sujeto expletivo es obligatoriamente nulo:

Llueve mucho /*Ello llueve mucho

3. Posibilidad de inversión sujeto-verbo:

Ellos llegaron / Llegaron ellos

4. Efecto complementante-huella:

¿Quién crees que habla español? / *¿Quién crees ____ habla español?

(Rothman e Iverson, 2007:329)

Aunque la primera propiedad indica que el sujeto referencial puede ser nulo o explícito, en realidad, esta distribución no se encuentra en variación libre como se demostrará a lo largo de esta disertación. Por ejemplo, en español, un sujeto puede ser nulo si el referente es identificable gracias a la inflexión morfológica de persona, género y/o número, los cuales son factores sintácticos, pero la distribución también se ve restringida por condiciones semánticas y pragmático-discursivas que indican el referente (Lubbers Quesada, 2015:23-24), así como condiciones cognitivas de atención y memoria, que veremos más adelante (Secciones 2.2 y 2.3, respectivamente). Igualmente, la posibilidad de inversión verbo-sujeto ilustrada en la tercera propiedad de Rizzi también

se ve condicionada por factores semánticos y pragmático-discursivos, además de sintácticos (Zubizarreta, 1998, Isabelli, 2003, Lubbers Quesada, 2015, *inter alia*). Por ejemplo, en oraciones declarativas sin foco, se prefiere el orden SV con verbos inergativos como *llamar, reír, hablar* (2), y se prefiere el orden VS con verbos inacusativos como *llegar, crecer, florecer* (3):

(2) María llamó (SV)

(3) Llegó María (VS)

(Lubbers Quesada, 2015:25)

Los ejemplos en (2) y (3) corresponden a restricciones semánticas relacionadas con el papel temático del participante: en los verbos inergativos el sujeto desempeña el papel de agente (el participante realiza la acción), mientras que en los verbos inacusativos el agente es el paciente o tema (el participante sufre la acción).

Contrariamente al inglés, los sujetos expletivos de la segunda propiedad, citada arriba, son obligatoriamente nulos en el español porque no tienen un referente en el mundo real (Lubbers Quesada, 2015:246), lo cual quiere decir que son semánticamente vacíos.

Más adelante, a las propiedades originales de Rizzi se añadió la restricción de pronombre explícito de Montalbetti (1984) (OPC⁵). Las lenguas que tienen contrastes entre sujetos nulos y plenos (explícitos), como lo son el español o el italiano, demuestran una asimetría que fue resaltada por Montalbetti en su tesis doctoral y la cual está relacionada con la posibilidad de ligamiento. La OPC aplica específicamente a la interpretación de pronombres nulos y explícitos en oraciones donde la matriz contiene un

⁵ OPC por el inglés *Overt Pronoun Constraint*.

sujeto variable como un cuantificador (p. ej. *nadie*) o una palabra *qu-* (p. ej. *quién*) (Zyzik, 2016:41 y Lubbers Quesada, 2015:27). Según la OPC, se evita el pronombre pleno cuando este está ligado a la forma variable. Por lo tanto, en (4) el pronombre *él* no puede hacer referencia al cuantificador *nadie* por ser una referencia desarticulada en donde el pronombre explícito no permite el ligamiento al sujeto cuantificador; mientras que en (5), *pro* puede interpretarse tanto como correferencial con el sujeto en la matriz *nadie*, o con otro referente conocido en el discurso:

(4) Nadie_i cree que *él*^{*i/j} es inteligente.

(5) Nadie_i cree que *pro*_i es inteligente.

(Adaptado de Montalbetti, 1984:26)

La OPC representa restricciones interpretativas sintácticas que han tenido repercusiones en los estudios de adquisición de lengua, lo cual se debe a que esta propiedad epitomiza la pobreza de estímulos desde la perspectiva de la gramática generativa, puesto que es una propiedad que no se aprende a base de input (Zyzik, 2016:41). En el Capítulo 3, se abordará una discusión más a fondo acerca de la OPC con respecto a la adquisición de lengua.

Otra propiedad sintáctica de los sujetos se detalla en la hipótesis de la posición del antecedente (PAH) propuesta por Carminati (2002), la cual plantea que la prominencia sintáctica del sujeto impone una restricción en la interpretación del pronombre anafórico. Esta hipótesis asume que hay una relación de prominencia entre un sujeto posverbal o preverbal, donde el sujeto nulo se relaciona más seguido con un antecedente más prominente (en posición preverbal) que el sujeto explícito. En otras palabras, la estructura sintáctica determina la prominencia del antecedente. La hipótesis predice que un sujeto

pro prefiere un antecedente en la posición de Spec IP (es decir, en posición preverbal), por lo tanto, en (6a) *pro* es correferente con *Mario*, mientras que el pronombre *él* en (6b) no está ligado al elemento en la posición de Spec IP, sino a *Juan*, elemento que se encuentra en una posición jerárquica inferior (objeto directo de la primera oración):

- (6) a. Cuando Mario_i llamó a Juan_j, *pro*_i había apenas terminado de comer
- b. Cuando Mario_i llamó a Juan_j, *él*_j había apenas terminado de comer

(Adaptado de Carminati, 2002:58)

Aunque Carminati basa sus resultados en el italiano, estudios ulteriores han demostrado que la PAH actúa también en el español y que, además, es más relevante que la OPC como restricción sintáctica interpretativa. En un estudio al respecto, llevado a cabo por Alonso-Ovalle, Fernández-Solera, Frazier y Clifton (2002), se realizaron una serie de experimentos con hablantes nativos (HNs) de español, estudiantes de la Universidad Complutense de Madrid, en donde se demuestran inconsistencias con la OPC, a favor de la PAH. Los experimentos 1 y 2 buscaban evidencia acerca de si la PAH era igualmente aplicable al español. Los resultados del experimento 1 muestran que, en oraciones tipo (6) más arriba, el 73.2% selecciona el sujeto de la primera oración cuando el pronombre es nulo, selección que se reduce al 50.2% cuando el pronombre es explícito (p.154). Por lo tanto, al igual que el italiano, *pro* en español favorece la selección de un sujeto en posición de Spec IP y el pronombre explícito desfavorece la selección del mismo sujeto. El experimento 2 demuestra que esta división de labores entre los pronombres nulo y explícito es independiente de cualquier ambigüedad que pueda presentarse en el proceso de resolución del antecedente (p.153); es decir, se excluye el

uso de un pronombre explícito con el objetivo de desambiguar entre dos posibles antecedentes. El experimento 3 pone a prueba la OPC con oraciones como las siguientes:

- (7) a. Ningún estudiante cree que *pro* pasó el examen.
b. Ningún estudiante cree que *él* pasó el examen.

(Alonso-Ovalle, Fernández-Solera, Frazier y Clifton, 2002:158)

Según la PAH, *pro* va a preferir el sujeto *ningún estudiante* en la oración (7a) como antecedente, sin importar su naturaleza variable, excluyendo así la necesidad de buscar un antecedente extraoracional. Sin embargo, esta predicción va directamente en contra de la OPC, según la cual se rechaza una interpretación ligada al sujeto variable. Los resultados muestran que en el 86.1% de los casos, *pro* recibe una interpretación ligada al sujeto variable, lo cual se reduce al 63.3% con un sujeto explícito como *él* en (7b). Los autores arguyen que, si la OPC fuera un principio gramatical, no habría respuestas que prefirieran una interpretación con *pro* ligada a los sujetos variables; sin embargo, la evidencia va en contra de este principio. Por este motivo, Alonso-Ovalle et ál. concluyen que la OPC debe subsumir a la PAH (p.160). En el presente trabajo también encontramos evidencia del papel que juega la PAH en el análisis de las expresiones referenciales. Como veremos en el análisis de los resultados (Capítulo 6), los datos demuestran que la correferencia entre el sujeto nulo o explícito con el antecedente está relacionada con los niveles de accesibilidad derivados de la prominencia sintáctica del referente en la posición de sujeto en la oración anterior.

La relación entre un elemento anafórico y un antecedente también se ve restringida por la relación sintáctica de mando-c, la cual es una relación de hermandad y dominancia entre constituyentes que fue propuesta por Reinhart (1976) en su tesis doctoral. Esta

relación intenta explicar la aceptabilidad de ciertas correferencias, mientras que otras se rechazan. Blackwell (2003:5-6) presenta ejemplos en español, en los cuales la regla estructural de mando-c nos permite ver cómo funciona esta restricción. En (8), *Alicia* es correferencial con el pronombre posesivo *su*, aunque el pronombre preceda al referente (catáfora), lo cual se debe al hecho de que el nodo *su* manda-c al nodo *Alicia*, según Blackwell (2003:6). En (9) no se presenta esta relación sintáctica y, por lo tanto, el pronombre nulo o explícito no pueden hacer referencia a *Alicia*, aunque este referente anteceda los pronombres:

(8) Con *su_i* madre *Alicia_i* se lleva muy bien.

(9) Con la madre de *Alicia_i* *{ella, \emptyset }_{*i}* se lleva muy bien.

(Blackwell, 2003:6)

Sin embargo, Blackwell (2003:6) añade que hay algunos ejemplos en los cuales parece que se violan las restricciones de mando-c:

(10) *He_i* didn't give her a diamond ring because, although he is madly in love with her, *Walter_i's* just not ready to tie the knot.

(McCray, 1980 citado en Blackwell, 2003:6)

Blackwell explica que el problema radica no en una ineficiencia de esta restricción sintáctica, sino al hecho de que hay factores pragmáticos que la invalidan. De hecho, más adelante Reinhart (1986:124) habla de la *intención* del hablante como factor que influye en la correferencia de un pronombre con su antecedente para explicar lo que ella denomina una coindexación opcional, como en (11), donde el pronombre *his* puede hacer referencia tanto al antecedente *Christopher* como a otro referente en el discurso:

(11) *Christopher_i* likes *his_{i/j}* bear

(Adaptado de Reinhart, 1986:123)

Aunque los estudios sintácticos han contribuido con una plétora de investigaciones en cuanto a la identificación de las propiedades sintácticas de las expresiones referenciales de sujeto, esta perspectiva se limita a la descripción estructural y paramétrica del pronombre nulo y explícito en oraciones aisladas, sin tener en cuenta toda la gama de expresiones referenciales con las que cuenta una lengua, como lo son las expresiones léxicas, los nombres propios o los demostrativos. Además, hemos visto cómo en varias ocasiones estos estudios no logran explicar de manera completa las condiciones que restringen la alternancia entre las diferentes expresiones, al ignorar factores semánticos, pragmáticos y cognitivos. Es por este motivo que más recientemente, el generativismo ha dado cabida a las interfaces sintáctico-pragmáticas al reconocer la importancia de incorporar factores semánticos y discursivos en su análisis. Autores como Zyzik reconocen este avance al decir: "...with the current emphasis on interfaces in generative theory (see Montrul, 2011; Sorace, 2011; White, 2011) both approaches are now looking at similar facets of the problem" (2017:34). Aquí Zyzik hace referencia al generativismo y al funcionalismo como dos enfoques que, al incluir las interfaces, estudian los mismos problemas desde dos perspectivas diferentes. Retomaré los estudios de las interfaces en el Capítulo 3, ya que estas investigaciones han tenido lugar primordialmente en el estudio de la adquisición de segundas lenguas.

Hemos visto en esta sección cuáles son las principales propiedades sintácticas que se han estudiado desde la perspectiva generativista. Estas incluyen las propiedades asociadas con el PSN, la OPC, la PAH y el mando-c. Las descripciones presentadas aquí de los sujetos en español no son exhaustivas, ya que el enfoque de este estudio no trata de la sintaxis generativa. Sin embargo, como veremos en el Capítulo 3, los planteamientos

que se han hecho de las propiedades sintácticas han permitido el avance de las investigaciones en el campo de la adquisición de lengua, incluyendo el estudio de las expresiones referenciales de sujeto, relevantes para esta investigación.

De las críticas más reconocidas y debatidas hacia la gramática generativa es el hecho de que no reconoce la función comunicativa del lenguaje, lo cual conlleva a análisis abstractos y a explicaciones modulares que ignoran factores por fuera de las estructuras sintácticas. Aunque se han hecho esfuerzos con las investigaciones de las interfaces, estos estudios siguen sin ser suficientes para explicar cómo se adquiere una lengua de una forma natural sin tener que recurrir a reglas y parámetros que algunos autores, como Tomasello (2005), consideran innecesarias. Sin embargo, el generativismo ha llevado a la luz estructuras que antes eran invisibles y con esto ahora tenemos la capacidad de explicar algunas de las razones que motivan ciertas selecciones lingüísticas por parte de los hablantes. La PAH de Carminati (2002) es un ejemplo concreto de cómo ciertas propiedades sintácticas permiten visualizar el comportamiento lingüístico de los hablantes con respecto a la posición de los sujetos. Como veremos en la siguiente sección, esta hipótesis es relevante en mi investigación ya que se ha demostrado que las nociones pragmáticas de tópico y foco tienen una influencia sobre la división del trabajo entre pronombres nulos y explícitos, codificada por la PAH.

A continuación, voy a describir algunas de las propiedades discursivo-pragmáticas que determinan la selección, uso e interpretación de las expresiones referenciales de sujeto, tanto anafóricas como no anafóricas.

2.3 Propiedades pragmáticas

Las propiedades pragmáticas son aquellas que siguen conceptos discursivos, entre los cuales se encuentran las nociones de tópico y foco. Estas no son propiedades sintácticas ya que se refieren a la estructura informativa de un enunciado, aunque, como veremos, los análisis pragmáticos no desestiman los factores sintácticos, en particular la PAH, que vimos en la sección anterior, debido a que tienen en cuenta la prominencia sintáctica de los sujetos según su posición en una oración. Dicha prominencia afecta las condiciones de tópico y foco, condiciones que en el presente trabajo hacen parte integral del análisis, en cuanto a que son factores que influyen el estado cognitivo de una expresión.

No hay una definición estándar para tópico y foco en la literatura; de hecho, Martín Rojo y Meeuwis (1993:92) hacen notar las diferentes terminologías utilizadas: tópico/comentario, información nueva/información dada, tema/remata. Chafe (1976) toma la definición de Hornby (1971 citado en Chafe 1976:47) y dice que tópico es aquella parte de la oración que constituye aquello acerca de lo cual está hablando el hablante, al resto se lo denomina comentario, que es aquello que provee información nueva acerca del tópico. Martín Rojo y Meeuwis añaden que la noción de tópico/comentario está separada de la noción información dada/nueva, ya que, según estos autores, un tópico puede ser nuevo o dado, acercándose más a la definición de Chafe en cuanto a que el tópico debe definirse en términos de *pragmatic aboutness* (1993:92); es decir, aquello de que trata la oración. Zubizarreta (1998:1) define foco en términos de lo presupuesto: foco es la parte de la oración que no es presupuesta; o sea, que no corresponde con presuposiciones compartidas entre el hablante y el oyente en el momento del enunciado en el discurso.

Según Gundel (1985:84), tópico es de lo que se está hablando, por lo tanto, es la información conocida, mientras que comentario es lo nuevo que se dice acerca del tópico, lo que se asevera. Lubbers Quesada (2015:36) relaciona la distribución de los sujetos nulos y explícitos con su función pragmática, donde el sujeto nulo se relaciona con el tópico/continuidad de tópico y tiene como función discursiva presentar información vieja o compartida; mientras que el sujeto explícito se relaciona con el foco/cambio de tópico y tiene como función introducir nueva información en el discurso o marcar un contraste. Siguiendo esta línea, para este trabajo adoptaré la terminología y definiciones de tópico y foco presentadas por Lubbers Quesada (2015).

Como anteriormente mencionado, la interpretación entre tópico y foco no solo se ve influenciada por factores pragmáticos, sino también por factores sintácticos, como lo es la posición pre o posverbal del sujeto. Varios autores (Zubizarreta, 1998, Pérez-Leroux y Glass, 1999, Alonso-Ovalle, Fernández-Solera, Frazier y Clifton, 2002) han demostrado que la posición de tópico en español generalmente se encuentra en posición preverbal. Por ejemplo, Alonso-Ovalle et ál. (2002) condujeron un estudio con el objetivo de demostrar que la PAH predice correctamente el comportamiento anafórico en español y articulan la evidencia de tópico/foco en la posición sintáctica. En el experimento número 4 del estudio, a ochenta participantes nativos de español se les presentó doce series de preguntas/respuestas, como en (12a y b):

- (12) a. ¿Quién vino? A. Vino Juan / B. Juan vino
b. ¿Qué sucedió con Juan? A. Vino Juan / B. Juan vino

(Alonso-Ovalle et ál., 2002:161)

Si la posición preverbal codifica el tópico de la oración, se espera que a la pregunta en (12b) la mayoría de las respuestas correspondan con B (*Juan vino*) (p.162). Los resultados muestran que el 92.83% de los hablantes nativos tienen una preferencia por un referente en la posición posverbal (*Vino Juan*) cuando la pregunta es de foco restringido, es decir, cuando la pregunta pide información nueva que contesta la pregunta (*¿Quién vino?*), mientras que en preguntas de foco amplio (*¿Qué sucedió con Juan?*), solamente el 48.5% prefieren el sujeto posverbal (*Vino Juan*). Según estos autores, esto es evidencia de que hay una clara asociación entre la posición sintáctica, la forma del sujeto, y la propiedad pragmática de foco/tópico en una oración; esto es, las nociones de tópico y foco tienen una influencia sobre la división del trabajo entre pronombres nulos y explícitos, codificada por la PAH.

Como mencionado anteriormente, el tipo de estrategia sintáctica o de *parsing* que ofrece la PAH es un análisis relevante en mi investigación puesto que los resultados demuestran que la prominencia del antecedente está determinada por su posición sintáctica en la oración: un sujeto explícito preverbal se considera más prominente y por lo tanto es más accesible que sujetos en otras posiciones de menor prominencia, como las posiciones de objeto. Sin embargo, Gundel (1985:86) asevera que el referente de un tópico sintáctico siempre es un tópico pragmático, pero que lo contrario no es verdadero; en otras palabras, una expresión que haga referencia a un tópico pragmático (es decir, que tenga el propósito de añadir información para el oyente acerca del referente), puede encontrarse en otras posiciones que no estén reservadas para los tópicos sintácticos (la posición de sujeto gramatical en una oración). Por ejemplo, en español, también son tópicos los complementos indirectos, como lo demuestra Blackwell (2001) en su estudio

con 105 participantes nativos de español. En este estudio, construcciones como en (13) fueron seguidas por una pregunta y cuatro posibles respuestas:

(13) Al llegar a casa Juan_i besó a su_i mujer_j y entonces *pro_{ij}* se puso a preparar la cena.

¿Quién preparó la cena?

- a. Juan
- b. la criada
- c. la esposa de Juan
- d. no se sabe

(Blackwell, 2001:935)

Juan es el tópico sintáctico de la oración y la respuesta (a.) *Juan*, fue seleccionada como correferente con *pro* el 83.81% de las veces, en línea con la prominencia sintáctica del referente en la posición de sujeto. Sin embargo, la respuesta (c.) *la esposa de Juan* (el referente indicado por el complemento indirecto, *a su mujer*), fue seleccionada como correferente con *pro* un 10.48% de las veces (Blackwell, 2001:935). Este último caso puede considerarse el tópico pragmático de la oración, ya que la selección se ve influenciada por procesos anafóricos interpretativos, los cuales a su vez pueden verse afectados por interpretaciones culturales relacionadas con nuestro conocimiento del mundo, ya que son las mujeres las que tradicionalmente preparan la cena en la mayoría de los casos de los participantes del estudio.

En contraste con el ejemplo anterior, en (14) la respuesta (a.), *la esposa de Juan*, fue seleccionada un 94.29% como correferente con *la mujer*:

(14) Al llegar a casa Juan_i besó a su_i esposaj y entonces *la mujer_j* se puso a preparar la cena.

¿Quién preparó la cena?

- a. la esposa de Juan
- b. la criada
- b. la madre de Juan
- c. no se sabe

(Blackwell, 2001:935)

Esto es importante porque, según Blackwell (2001:935), se demuestra que hay una tendencia a seleccionar e interpretar los sujetos nulos como correferentes con el tópico sintáctico y pragmático de la oración, mientras que un referente explícito tiende a ser correferente con una entidad en una posición de menor prominencia en una oración, como lo es la posición de objeto. Estos resultados están en línea con la propuesta de Huang (1991), según la cual las correferencias siguen una jerarquía preferencial con base en la prominencia del antecedente, de la siguiente manera: tópico > sujeto > objeto > otros (p.330).

Si concluimos que la información nueva se introduce en el discurso como foco, mientras que la información conocida se introduce como tópico, también debemos tener en cuenta la continuidad y el cambio de tópico en el uso de las expresiones referenciales de sujeto, ya que esta noción afecta la selección de la expresión referencial misma. Por ejemplo, Lubbers Quesada (2015:33) dice que información nueva es generalmente introducida en el discurso con una frase nominal, o con nombres propios, como en (15) con el uso de *Charles*:

- (15) Pues, *Charles_i* llegó, *pro_i* vio a la muchacha que estaba sola; *pro_i* se sienta a platicar con ella pero, bueno, pues casi, casi *pro_i* tropezó con ella.

(Lubbers Quesada y Blake, 2003, citados en Lubbers Quesada, 2015:33)

Además, vemos que esta entidad (*Charles*) continúa siendo el tópico del discurso, lo cual se representa fonéticamente con un elemento más mínimo (*pro*), al igual que en el caso de continuidad de tópico, o mismo referente, como sucede en “casi *pro* tropezó con ella”. Sin embargo, en los casos de cambio de tópico (cambio de referente), el hablante

nativo tiende a utilizar de nuevo un marcador menos mínimo, como se ve en (16) con el uso del pronombre explícito *él*:

- (16) Entonces *pro_i* jala al señor de los ojos vendados y cuando *él_j* se da cuenta de que *pro_i* no es la muchacha_k,

(Lubbers Quesada y Blake, 2003, citados en Lubbers Quesada, 2015:33)

Varios otros autores analizan la selección de las diferentes formas lingüísticas según el foco, la continuidad de tópico y el cambio de tópico (Pérez-Leroux y Glass, 1999, Montrul y Rodríguez Louro, 2006, Lozano, 2009 *inter alia*). Según Pérez-Leroux y Glass (1999), las lenguas [+pro drop] hacen un uso frecuente del pronombre nulo ya que este codifica continuidad de tópico en el discurso. En contraste, el pronombre pleno es interpretado como foco, codificando información nueva o contrastiva. Por lo tanto, un pronombre nulo no es posible si el sujeto es nuevo en el discurso y da el siguiente ejemplo en (17), donde el pronombre explícito *ella* es información nueva y por lo tanto el foco:

- (17) ¿Quién viene?
Ella viene / *Viene

(Pérez-Leroux y Glass, 1999:226)

Similar al estudio de Pérez-Leroux y Glass (1999), Montrul y Rodríguez Louro (2006) consideran la distribución de los pronombres de sujeto nulos y explícitos según si el sujeto es tópico, si introduce un cambio de referente o si establece foco o contraste.

Presentan los siguientes ejemplos:

- (18) Mismo referente: Pepe no vino hoy a trabajar. *Pepe/?él/ *pro* estará enfermo.
(19) Cambio de referente: Hoy no fui a trabajar. Pepe/él/* *pro* pensó que estaba enferma.

(20) Tópico: ¿Quién vino? Él/Mario/* *pro* vino.

(21) Foco: El periodista_i dijo que él_i no había escrito ese reporte.

(Montrul y Rodríguez Louro, 2006:404)

Según las autoras estos ejemplos sirven para demostrar el hecho de que en las lenguas [+pro drop], como el español y el italiano, además de factores morfosintácticos, la distribución del sujeto nulo vs el sujeto explícito se ve afectada por factores pragmáticos, gobernados por el discurso. También encontramos evidencia de esto en los resultados de mi investigación, donde la forma del sujeto se ve afectada por los factores pragmáticos de continuidad de tópico y cambio de tópico. Veremos que los hablantes tienden a usar sujetos nulos para marcar la continuidad de tópico en el discurso, y seleccionarán formas plenas para marcar foco o contraste. Los estudios de Pérez-Leroux y Glass (1999) y de Montrul y Rodríguez Louro (2006) se enfocan en la adquisición de segunda lengua, y por lo tanto los retomaré en el Capítulo 3 para examinar las implicaciones de estos factores en los procesos de la adquisición.

Otro uso pragmático de un referente explícito es el foco contrastivo, en cuyo caso el pronombre explícito va acompañado de énfasis prosódico, como en el caso de (22). Las mayúsculas indican énfasis prosódico:

(22) El periodista_i dijo que él_i/ÉL_i no había escrito ese reporte.

(Montrul, 2004:128)

Aquí vemos que, sin énfasis prosódico, el pronombre *él* es tópico, pero con énfasis prosódico, *ÉL* también es contrastivo, es decir, la referencia es solo *él mismo*, y no otro periodista. Lubbers Quesada (2015:35) aclara que, aunque esto es cierto, en algunos contextos puede darse el caso de que la referencia sea con otro antecedente. Zubizarreta

(1998:76) dice que, en español, el estrés prosódico es bastante rígido e imposible en posición preverbal, como en los ejemplos de (23), a menos de que la interpretación sea contrastiva, como en (24):

- (23) a. *El SOL salió (vs. El sol SALIÓ).
 b. *El CORREO llegó (vs El correo LLEGÓ).
- (24) a. JUAN llamó por teléfono (no Pedro).
 b. MARÍA se comió el pastel (no Marta).

(Zubizarreta, 1998:76)

Para resumir, la Tabla 2.3.1 más abajo, muestra la distribución entre los sujetos nulos y explícitos, según las nociones pragmáticas de foco y tópico:

Tabla 2.3.1 - Distribución en español de los sujetos nulos y explícitos, según las nociones pragmáticas de tópico y foco

Tópico/Foco		Tipo de sujeto
Tópico (información conocida/compartida/dada)	• Continuidad de tópico	nulo
	• Cambio de tópico	explícito
Foco (información nueva)	• Introducción de información	explícito
	• Foco contrastivo	explícito

Es importante añadir que una diferencia significativa entre las propiedades lingüísticas sintácticas y aquellas pragmáticas es el hecho de que, en estas últimas, se consideran todas las expresiones referenciales de las que dispone una lengua, por lo tanto, en la tabla anterior, un sujeto explícito puede referirse no solamente a un pronombre (*él/ella*), sino también a nombres propios (*María*), a frases nominales (*el sol, la mujer*) o

a demostrativos (*eso, esto, aquello*). En mi investigación, se consideran todas las formas referenciales de sujeto, en línea con los estudios pragmáticos que hemos visto.

Además, el análisis pragmático no desestima las consideraciones sintácticas y se tienen en cuenta en la presente investigación. Principalmente, vimos que la prominencia del referente según su posición sintáctica en la oración es un factor que afecta las nociones pragmáticas de tópico y foco, así como la distribución de los referentes nulos y explícitos, tanto anafóricos como no anafóricos (Carminati, 2002; Alonso-Ovalle et ál., 2002).

Las propiedades pragmáticas están íntimamente relacionadas con las propiedades cognitivas que describiré en la siguiente sección, ya que estas también se ven afectadas por el contexto discursivo, aunque van más allá al tener en cuenta también factores cognitivos como lo son la memoria y la atención del oyente. Tomasello (2005:200) dice que las nociones de accesibilidad, topicalidad y de “lo dado” (*givenness*) son habilidades cognitivas que el hablante debe considerar en el proceso referencial para el oyente. Por lo tanto, para esta investigación es relevante determinar la interacción entre los factores pragmáticos (tópico/foco) que hemos visto y aquellos cognitivos (memoria/atención).

2.4 Propiedades cognitivas

El estudio de las expresiones de sujeto desde una perspectiva cognitivista se basa en gran medida en el concepto de la accesibilidad del referente y su prominencia cognitiva. Lo importante a reconocer con este tipo de interpretación, es que la selección de un sujeto no depende solamente de propiedades sintácticas, sino también de factores cognitivos como la atención, la memoria y la información compartida. En este trabajo los

factores cognitivos son de particular importancia debido al enfoque teórico adoptado para esta investigación, sobretodo en la comparación de los datos orales vs los datos escritos debido a los recursos limitados del sistema de memoria que pueden afectar la modalidad discursiva. En estos análisis también se tienen en cuenta las propiedades discursivo-pragmáticas de foco y tópico, así como otras propiedades lingüísticas, por ejemplo, los factores de prominencia sintáctica.

La noción de la accesibilidad es cognitiva, ya que se activa un proceso de comprensión en relación directa con la “facilidad o dificultad para recuperar el material relevante para la interpretación” (Figueras, 2002:54). Dicho proceso está sujeto a la capacidad de la memoria, de tal manera que una expresión referencial codificará la menor o mayor accesibilidad que se tenga en la memoria de un determinado referente. Por ejemplo, un sujeto nulo codificará un alto nivel de accesibilidad en la memoria a corto plazo, mientras que un sujeto pleno codificará un bajo nivel de accesibilidad en la memoria a largo plazo, como veremos con ejemplos de Ariel (1988, 1990) en esta misma sección.

Varios autores han propuesto diferentes escalas, o jerarquías, que representan el nivel de accesibilidad a los referentes en el discurso; entre ellos se cuentan Prince (1981), Givón (1983), Ariel (1988) y Gundel, Hedberg y Zacharski (1993). Dichas jerarquías nacen originalmente de la diferenciación entre información nueva e información dada. Chafe (1976) hace esta distinción al hablar de dos tipos de información: información nueva (aquella no activada en la mente del oyente), e información dada (aquella activada en la mente del oyente; es decir, que el oyente ya conoce). Esta es una cuestión cognitiva, puesto que la información dada se encuentra almacenada en la mente y el hablante va a

acomodar su lenguaje con respecto al supuesto estado mental de la información en la mente del oyente (p. 32). Esto último es importante, puesto que Chafe afirma que es el hablante el que decide el estado cognitivo (EC) (p.30); por lo tanto, es el hablante el que hace presuposiciones acerca del EC del oyente y acomoda así la elección de las formas lingüísticas que utilizará para hacer referencia a las entidades en el discurso. Chafe añade que una mención previa es la base más común para determinar si algo está activado en la mente y da el siguiente ejemplo, donde el uso del pronombre *it* es adecuado, gracias a la activación del referente *a painting* en el enunciado anterior:

(25) I'd like to show you *a painting* (new). I bought *it* (given) last week
 (Chafe, 1976:32, mis cursivas)

Igualmente, Chafe señala la capacidad limitada de la memoria, y por lo tanto resalta la noción de recuperabilidad (*recoverability*) (p. 32). Esta noción se verá elaborada subsecuentemente en los trabajos de accesibilidad de Ariel (1990) y con la Givenness Hierarchy de Gundel, Hedberg y Zacharski (1993), entre otros. Sin embargo, la primera propuesta de análisis discursivo que tiene en cuenta los factores de familiaridad, prominencia y TC es la jerarquía de familiaridad de Prince (1981). La taxonomía propuesta por esta autora se categoriza según la información sea dada (conocida/familiar) o nueva (no familiar), ver Figura 2.4.1.

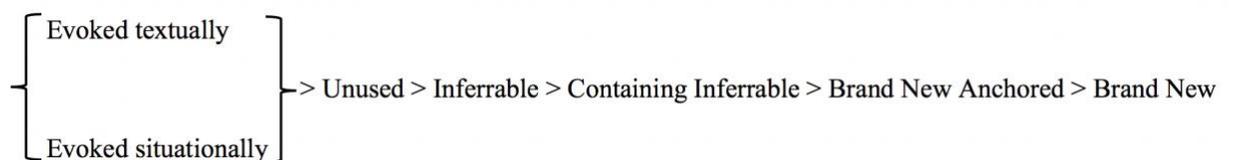


Figure 2.4.1 - Escala de familiaridad (adaptada de Prince, 1981:245)

El hablante asume un nivel de familiaridad que tiene el oyente con una entidad determinada y selecciona una expresión referencial específica que refleja ese nivel de familiaridad. Por ejemplo, en (26a), el pronombre de primera persona *I* se califica como evocado contextualmente (*Evoked situationally*), mientras que *A person* en (26b) se califica como totalmente nuevo (*Brand New*).

- (26) a. *I* bought a Toyota
 b. *A person* bought a Toyota

(Prince, 1981:245)

Las expresiones referenciales utilizadas por el hablante no son intercambiables, en consecuencia, si utilizamos una expresión que no corresponde con el nivel de familiaridad que tiene el oyente con respecto al referente, el acto comunicativo no podrá considerarse adecuado o cooperativo, según los principios de cooperación de Grice (Prince, 1981:245). Aunque estos principios son pragmáticos, la jerarquía propuesta por Prince puede considerarse cognitiva ya que implica la accesibilidad que se tenga a un determinado referente en la mente del oyente y, como anteriormente mencionado, la accesibilidad es una noción cognitiva.

Aunque Prince presenta una taxonomía, no da una lista de cuáles expresiones se suelen usar con cada categoría como sí lo hacen los autores que veremos a continuación. Lo que sí hace esta autora, es introducir la noción de inferibles (*Inferrables*) al interpretar las expresiones anafóricas. En palabras de Prince, “A discourse entity is Inferrable if the speaker assumes the hearer can infer it, via logical, or, more commonly, plausible, reasoning, from discourse entities already Evoked or from other Inferrables” (1981:236). Esto quiere decir que el referente del pronombre puede ser identificado ya sea a través de

una mención explícita previa o al haber sido simplemente evocado en el discurso anterior, ver ejemplo (27):

(27) I got on *a bus* yesterday and *the driver* was drunk.

(Prince, 1981:233)

La frase nominal *the driver* es inferible ya que el referente *a bus* ha sido mencionado en el discurso anterior, y el hecho de que todos los buses tienen conductores forma parte del conocimiento mutuo. También puede ocurrir que el referente no haya sido mencionado, como es el caso en (28):

(28) [Talking about how the kids across the street threw paint in their yard]
Those kids are just – And *she's* pregnant with another one.

(Zulaica-Hernández, 2009:649)

El referente del pronombre *she* no se menciona explícitamente en este enunciado, pero puede inferirse (ser evocado) que los niños tienen una madre, a través de la mención de la frase nominal en el enunciado anterior *those kids*. Este fenómeno tiene que ver con los esquemas y con el concepto de enlace (*bridging*), acerca de los cuales hablaré en más detalle al final de esta misma sección.

Varios autores han relacionado las diferentes expresiones referenciales con las nociones pragmáticas de continuidad y cambio de tópico. Por ejemplo, Givón (1983:9) propone una escala de implicación basada en niveles de accesibilidad y de continuidad. La Figura 2.4.2 es la escala propuesta por Givón (1983:17), la cual es de carácter funcionalista, ya que relaciona una determinada forma con un uso (función):

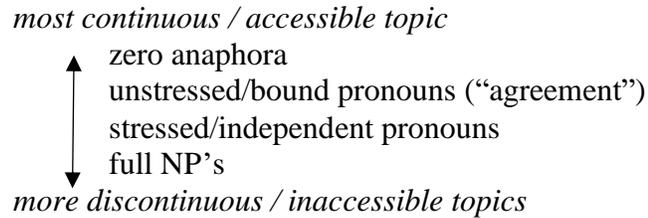


Figura 2.4.2 – Escala de continuidad de tópicos (Givón, 1983:18)

Esta escala representa una relación inversa en cuanto a que, entre mayor continuidad, más mínima será la expresión referencial, lo cual en español corresponde con el sujeto nulo; mientras que a menor continuidad (como en el caso de un cambio de tópicos), menos mínima será la expresión, lo cual corresponde con expresiones léxicas, por ejemplo, expresiones definidas o indefinidas. La propuesta de Givón relaciona la familiaridad del referente entre los interlocutores, así como su accesibilidad en la memoria de dicho referente.

Ariel (1990) desarrolla la noción de la accesibilidad más allá de la distinción entre información nueva y dada y asocia diferentes esfuerzos de procesamiento cognitivo con diferentes referentes lingüísticos, en donde el oyente parte del grado de accesibilidad codificado en la expresión referencial y realiza un proceso inferencial que le permite interpretarla (Figueras, 2002:55). La perspectiva de Ariel difiere con la de Prince (1981) en que, en vez de basarse en el contexto para explicar las expresiones referenciales (es decir, en la fuente del conocimiento), proviene del lenguaje natural mismo y, a su vez, asocia la accesibilidad con el contexto. Por contexto, Ariel se refiere al conocimiento general o enciclopédico, al entorno físico, y al material lingüístico anterior (Ariel, 1990:10). Para ilustrar lo anterior con un ejemplo, en (29) Ariel arguye que el contexto no puede explicar la preferencia por (29b), en donde el contexto del referente es el

conocimiento general, sobre (29a), en donde el hablante lleva al oyente hacia el contexto físico a través de una expresión referencial demostrativa:

- (29) a. *That woman over there* is very intelligent.
 b. *Rachel* is very intelligent.

Según Ariel, una expresión referencial debe asociarse con el nivel de accesibilidad del referente, y no con el tipo de contexto (1990:11), lo cual explica la preferencia por (29b), ya que, para la introducción de un referente en un discurso, se prefiere el uso de un nombre propio en vez de una expresión demostrativa (Ariel, 1988:68). Así mismo, la autora arguye que hay una relación directa entre el nivel informativo de la anáfora y la accesibilidad del referente y esta codificación se encuentra en un continuo de marcadores de baja, media y alta accesibilidad (Ariel, 1988, 1990). En la Figura 2.4.3, más abajo, vemos dos ejes: el eje vertical corresponde con la memoria y el eje horizontal con la accesibilidad:

		Accessibility					
		Low				High	
MEMORY ↑ ↓	Long-term	KC:	Joan Smith the president	Joan Smith	The president	Smith	Joan
		PC:	This/that hat we bought last year	That hat	This hat	I/You/That	This
	Short-term	LC:	SHE she	Hers	The former/latter	Herself	Ø

KC: General Knowledge Context PC: Physical Context LC: Linguistic Context

Figura 2.4.3 - Initial Accessibility Marking (adaptado de Ariel, 1990:70)

En esta figura vemos que expresiones de baja accesibilidad y por lo tanto altamente informativas, como *Joan Smith the president*, se encuentran almacenadas en la memoria a

largo plazo; un demostrativo como *that hat* es de accesibilidad media y se encuentra en la memoria a medio plazo; y en el otro extremo se encuentra el pronombre nulo, de alta accesibilidad y en consecuencia de baja informatividad, el cual se encuentra en la memoria operativa (a corto plazo).

Dentro de cada nivel de accesibilidad Ariel distingue lo que podríamos llamar subniveles en donde, por ejemplo, una frase nominal definida es menos accesible que un nombre, y dentro de los nombres un nombre completo (nombre y apellido) codifica menor accesibilidad que un nombre parcial (solo el nombre o solo el apellido), y así sucesivamente con todos los niveles⁶. Estos niveles de accesibilidad de Ariel (1990) son relevantes en el análisis de los nombres propios de mi investigación, ya que el análisis de Ariel ayuda a explicar en algunos casos la selección que hacen los participantes del estudio entre el primer nombre del personaje (Charlie) vs el nombre completo (Charlie Chaplin).

Es importante resaltar que la memoria está relacionada con la distancia en el discurso entre la anáfora y el antecedente. De hecho, Ariel identifica la distancia como uno de los factores relevantes en el proceso de interpretación de las expresiones anafóricas, siendo los otros factores el número de competidores por el papel de antecedente, la topicalidad del antecedente y el papel de los marcos en la identificación del referente (1988:65). Los marcadores de baja accesibilidad se encuentran en la memoria a largo plazo porque tienen antecedentes distantes, y por lo tanto contienen una mayor información léxica, lo cual conlleva a un mayor costo de procesamiento cognitivo en el oyente, como lo son las frases nominales definidas y los nombres propios,

⁶ Para una revisión completa y detallada de los marcadores de accesibilidad, ver Ariel 1988 y 1990.

ejemplificado en (30a y b). En los marcadores de accesibilidad media Ariel incluye los demostrativos y, para el español, da el ejemplo de *este* y *ese* (no solamente de uso deíctico), visto en (31). Por último, los marcadores de alta accesibilidad son los pronombres y los huecos (*gaps*), que son marcadores semánticamente vacíos con un antecedente saliente, por ejemplo, en (32a y b). El símbolo \emptyset representa un hueco:

- (30) a. *The party* is scheduled to announce its nuclear policy this afternoon.
b. *Ghandi* is a real man.
- (31) *This stupid neighbor* is getting on my nerves.
- (32) a. Sherlock Holmes to Watson: The butler did *it* (the murder, etc.)
b. Here is some syrup for you. Shake \emptyset before using.

(Ariel, 1990:7-8)

La distancia entre el elemento anafórico y el antecedente se tiene en cuenta en la presente investigación porque ayuda a determinar la categorización del EC. Como hemos visto, la distancia afecta la distinción entre si un referente se encuentra activado en la memoria a corto plazo, o si es familiar y almacenado en la memoria a largo plazo, por consiguiente, afectando la forma referencial que seleccionará el hablante/escritor.

Gundel, Hedberg y Zacharski (1993) proponen una jerarquía basada en seis ECs que denominan la Givenness Hierarchy (GH). En esta propuesta, la forma de la expresión referencial depende del EC que se asume del referente intencionado por parte del hablante. Los factores que intervienen son el nivel de atención y memoria con respecto a dicho referente. Gundel et ál. mantienen que la GH es universal; es decir, aplica a todas las lenguas, aunque no todas las lenguas utilizan las mismas formas para señalar los distintos estados. En su estudio, los autores presentan una propuesta para el inglés, el

español, el japonés, el ruso, y el chino mandarín (Gundel et ál., 1993:284) y se basan en una combinación de datos de diferentes registros y géneros, desde el oral, el escrito, el informal e informal, sin hacer una distinción específica en los resultados obtenidos. En el Capítulo 3 veremos una propuesta de Blackwell y Lubbers Quesada (2012) en la cual se modifica esta jerarquía para el español, y se demuestra su aplicabilidad no solo para los hablantes nativos del español, sino también para los estudiantes de español como L2. Los trabajos de Gundel et ál. (1993) y de Blackwell y Lubbers Quesada (2012) son relevantes al presente estudio debido a que se analizan los datos de este último dentro del mismo marco teórico.

Los ECs propuestos por Gundel et ál. son los siguientes (ver Figura 2.4.4, más abajo): de tipo identificable (type identifiable - TYP), referencial (referential - REF), identificable de modo único (uniquely identifiable - UNI), familiar (FAM), activado (activated - ACT), referente en foco (in focus - INF)⁷:

in focus >	activated >	familiar >	uniquely identifiable >	referential >	type identifiable
<i>it</i>	<i>that</i> <i>this</i> <i>this N</i>	<i>that N</i>	<i>the N</i>	indefinite <i>this N</i>	<i>a N</i>

Figura 2.4.4 - The Givenness Hierarchy (Gundel, Hedberg y Zacharski, 1993:275)

A continuación, presento una descripción con ejemplos de cada uno de los estados representados en la Figura 2.4.4 más arriba. Todos los ejemplos son tomados de Gundel et ál. (1993:276-280):

⁷ Utilizo las traducciones de los estatus propuestas por Figueras (2002:61), pero mantengo los acrónimos de Gundel et ál. (1993) para facilitar la comparación en el análisis de mis resultados en el Capítulo 3.

- TYP: se asume que el oyente accede a la representación mental del referente *dog*, solo si comprende la palabra *dog*. En este caso, una expresión indefinida es suficiente: “I couldn’t sleep last night. *A dog* (next door) kept me awake”.
- REF: el inglés es la única lengua de aquellas presentadas por los autores que distingue una forma lingüística diferente entre los estados TYP y REF. En este estado, el oyente no solo debe poder acceder a una representación mental, sino que debe reconstruirla a medida que se procesa el enunciado: “I couldn’t sleep last night. *This dog* (next door) kept me awake”. Como explicaré más adelante, este estado no es relevante para el español.
- UNI: el interlocutor puede identificar el referente por medio de la expresión lingüística del nominal mismo. Este estado es el mínimo requerido para el uso de expresiones definidas: “I couldn’t sleep last night. *The dog* (next door) kept me awake”. Esta expresión definida es adecuada (*felicitous*) aun si no estamos familiarizados con el referente, ya que podemos acceder a él mentalmente a través de su expresión lingüística.
- FAM: ya se tiene una representación en la memoria de la entidad o entidades, por lo tanto, el uso de demostrativos es suficiente para su identificación: “I couldn’t sleep last night. *That dog* (next door) kept me awake”.
- ACT: el referente puede accederse en la memoria a corto plazo: “I couldn’t sleep last night. *That* kept me awake”.
- INF: en este estado también ya se tiene una representación en la memoria a corto plazo y además, el referente es el centro de atención en el enunciado actual, y generalmente se encuentra en posición de sujeto u objeto directo: “My neighbor’s

bull mastiff bit a girl on a bike / *It's* the same dog that bit Mary Ben last Summer”.

Según Gundel et ál., los estados más accesibles son aquellos de INF y ACT, puesto que el antecedente del referente es más inmediato en la memoria a corto plazo. A diferencia de las escalas descritas más arriba, esta es una jerarquía implicativa; es decir, cada estado está contenido en el siguiente y no son mutuamente excluyentes. Por ejemplo, si un referente se puede codificar como INF, es porque también es ACT, FAM, UNI, y así sucesivamente. Además, es unidireccional, lo cual significa que una entidad INF también es ACT pero no lo contrario.

Posteriormente, Gundel (2010) aclara que los estados de la GH no codifican un nivel o grado de accesibilidad, como sí lo hacen las escalas propuestas por Prince (1981), Givón (1983) y Ariel (1990), entre otros, sino que codifican la manera. Por grado de accesibilidad, Gundel se refiere al esfuerzo de procesamiento cognitivo asociado con las diferentes expresiones referenciales, (Gundel, 2010:148) y la manera se refiere a cómo y a dónde se accede a la representación mental para interpretar el intencionado referente, de tal modo que el hablante guía al oyente hacia una interpretación restringida por la forma seleccionada de expresión lingüística (Gundel, 2010:152). Por ejemplo, el pronombre *it* está asociado con una representación que se encuentra en el foco de atención del oyente; un demostrativo como *this/that* está asociado a una representación que se encuentra en la memoria operativa del oyente; mientras que una frase nominal indefinida como *a N* está asociada a una representación de tipo de entidad expresada por el sustantivo.

Utilizo la propuesta de Gundel et ál. como base teórica para mi investigación porque las investigaciones de estos autores tienen una base cognitivista. Gundel

(2010:154) dice que muchos de los criterios utilizados para determinar los ECs tienen un apoyo en investigaciones psicolingüísticas, por ejemplo, el criterio de que entidades introducidas en una posición sintáctica prominente en la oración inmediatamente anterior pueden asumirse en el foco de atención es una consecuencia automática de la manera como los humanos procesan las estructuras sintácticas en una oración. Tomasello (2005:267) dice que los niños, ya a partir de los 24 meses, tienen la capacidad lingüística de especificar en los turnos de sus conversaciones el tópico del discurso, así como el comentario que desean hacer sobre el tópico, el cual generalmente incluye información nueva. Dentro de este marco de atención compartido (*joint attentional frame*) entre los interlocutores, el tópico está dentro del marco, es accesible y contiene información compartida; desde la perspectiva de la GH, el tópico hace parte del foco de atención. El comentario (o foco⁸) se encuentra fuera del marco de atención y se requiere de un cierto esfuerzo cognitivo para ser llevados lingüísticamente al foco de atención (Tomasello, 2005:267).

Zulaica-Hernández (2009) hace un análisis de los demostrativos en español *eso/esto/aquello* y su EC propuesto como ACT por la GH. El autor, en contra de Gundel et ál. (1993), arguye por un estado INF de los demostrativos debido a la alta prominencia del referente. Este es un estudio de corpus utilizando el CREA, en el cual se midió la distancia entre el antecedente y la anáfora. Zulaica-Hernández analizó nueve tipos de eventos con verbos como *suced*, *ocurrir*, *terminar*, etc., combinados con los tres demostrativos en español arriba mencionados, por ejemplo: *esto ocurrió*, *eso ocurrió*,

⁸ No confundir *en foco*, lo cual se refiere al estado de atención actual relacionado con el EC INF de Gundel et ál. y *foco* o *comentario*, relacionado con la información nueva que se añade acerca del tópico.

aquello ocurrió. En total, se analizaron 193 casos y se midió la distancia en número de cláusulas, con el antecedente.

La primera conclusión del estudio es que los tres demostrativos prefieren antecedentes inmediatamente anteriores (p.660): *esto* (78%), *eso* (78.2%) y *aquello* (76%). En segundo lugar, el 80% de los demostrativos están dentro del EC ACT, pero, según Zulaica-Hernández, deben tratarse como INF por lo que son lo suficientemente prominentes ya que acaban de ser nombrados en la cláusula anterior.

Este autor también menciona otros factores que resaltan la prominencia de los demostrativos: son sintácticamente prominentes ya que usualmente son cláusulas enteras, y por lo tanto es difícil degradarlos a posiciones secundarias y, además, generalmente constituyen el tópico de la conversación.

Zulaica-Hernández no proporciona ejemplos del corpus, lo cual es una debilidad de este artículo, pero principalmente este autor no parece tener en cuenta otros factores que deben considerarse al establecer el EC INF, como la prominencia pragmática. Tomando un ejemplo directamente de Gundel (2010), vemos que la forma demostrativa *that* es ambigua ya que puede referirse tanto a *my linguistic class* como a todo el enunciado *I failed my linguistic class*, mientras que el pronombre *it* no es ambiguo, y solo puede referirse a *my linguistic class*:

(33) A: I failed my linguistics class.

B: Can you repeat *that*?

B: Can you repeat *it*?

(Gundel, 2010:156)

Esto se debe a que, según Gundel, el hecho mismo de enunciar una oración lleva a la oración al estado de ACT, pero no a INF, mientras que el tópico pragmático de la

oración, en este caso *my linguistic class*, está en foco debido a su importancia dentro del enunciado, su relativa prominencia sintáctica y su mención reciente.

Otro ejemplo en contra del argumento de Zulaica-Hernández también se puede extraer de Gundel (2010):

(34) The kitchen is across the den. *It's/That's* my favourite room.

(Gundel, 2010:158)

En este caso, el pronombre *it* solo puede referirse a *the kitchen*, la entidad que se encuentra en foco al encontrarse en posición prominente de sujeto sintáctico, mientras que el demostrativo *that* es ambiguo y puede referirse tanto a *the kitchen* como a *the den*, este último meramente activado pero no en foco. Por lo tanto, lo más natural es una correferencia entre *that* y *the den*, debido a que, a fin de cuentas, *that* normalmente codifica el referente que no está en foco.

A través de estos ejemplos concluyo que los demostrativos, ya sea en español o en inglés, codifican un referente activado pero no en foco, en línea con el análisis de la GH de Gundel et ál (1993), y así serán clasificados mis datos.

Por último, quiero hacer una breve mención al concepto de *bridging* (enlace/enlazamiento), fenómeno que se conoce en la literatura también como anáfora indirecta, entre otros nombres. Este concepto es relevante en situaciones en las cuales una expresión referencial tiene como antecedente un referente que no está presente explícitamente en el discurso. Clark (1975:169) define *bridging* como el proceso por el cual un oyente vincula una expresión referencial a un antecedente cuando este no se encuentra en la primera etapa de búsqueda. En otras palabras, *bridging* es un proceso que enlaza un referente que no ha sido explícitamente mencionado con algo que sí lo ha sido

a través de inferencias. Veamos por ejemplo la oración en (35), donde, en una primera búsqueda no se encuentra un antecedente directo para la expresión definida *the ceiling*:

(35) I looked into the room. *The ceiling* was very high.

(Clark, 1975:171)

Sin embargo, la expresión es adecuada, ya que todas las habitaciones tienen techos y, a través de un proceso inferencial, se puede identificar *the room* como el antecedente de *the ceiling*. A su vez, este concepto está ligado a la noción de los marcos como base para el TC en algunas expresiones referenciales que veremos en la Sección 2.5. Hay varios ejemplos en los datos de mi investigación donde vemos que el uso de una cláusula definida es adecuado, aun siendo la primera mención en la narrativa, debido a que puede enlazarse a un esquema cognitivo a través del proceso de bridging. Por ejemplo, los hablantes tienden a introducir a la familia de la película del estudio con la cláusula indefinida *una familia*, para luego mencionar a los componentes individuales de la misma a través de cláusulas definidas como *el padre, la madre, etc.* Generalmente, estas instancias se categorizan dentro del EC UNI ya que el referente puede ser identificado por la expresión misma.

En conclusión, las propiedades cognitivas, al ser compatibles con las propiedades sintácticas y pragmático-discursivas, proporcionan una perspectiva holística en cuanto al proceso de selección, uso e interpretación de las expresiones referenciales de sujeto. Los modelos teóricos presentados permiten explicar los procesos cognitivos llevados a cabo por los hablantes y por lo tanto proporcionan una base teórica para el análisis discursivo de la presente investigación. En la siguiente y última sección de este capítulo, voy a

describir algunos análisis relacionados con el TC como factor influyente en este proceso de resolución interpretativa de las expresiones referenciales.

2.5 El terreno común (TC)

En esta sección voy a presentar cómo se ha venido estudiando el concepto del TC dentro del análisis del discurso para determinar su asociación con el uso de formas lingüísticas específicas. Los diferentes estudios han abarcado el uso de las descripciones definidas e indefinidas, el TC como factor en las oraciones escindidas, con la prosodia (como es el caso del inglés) y el orden de las palabras (en el caso del español). Para este trabajo, nos vamos a concentrar específicamente en los estudios llevados a cabo con respecto a las cláusulas definidas por lo que en nuestras hipótesis predecimos un aumento de estas formas referenciales cuando se puede asumir TC entre escritor y lector. Esta discusión es de particular importancia debido a que en la presente investigación se toma en cuenta el papel del TC como variable independiente para determinar hasta qué punto afecta la elección del referente, principalmente en el momento de introducción de un referente en la narrativa.

Carlson (1992:60) define el TC como aquella parte compartida entre el hablante y el oyente de sus conocimientos, creencias y suposiciones. Según este autor, para que un oyente pueda comprender a un hablante, el primero puede limitarse a ciertos dominios de la información que pertenecen al TC; es decir, a aquellas partes del conocimiento, creencias y suposiciones que son compartidos entre el oyente y el hablante (p. 61). El TC es dependiente del contexto, el cual Carlson define como aquella información disponible

a una persona en particular, con un proceso particular para una ocasión en particular⁹ (p.65). Esta definición es importante porque enfatiza la relatividad del contexto, por lo tanto, el TC que haya entre los interlocutores es relativo a las personas que participan, al momento en el que lo hagan, al lugar, a la ocasión, a la situación y a la información en sí. En la Figura 2.5.1, más abajo, Carlson ofrece una definición para determinar el conocimiento compartido de una proposición p .

Clark y Marshall (1981:11) se refieren a la anterior definición como la paradoja del TC, la cual podemos explicar de la siguiente manera: a A le debe surgir la pregunta de cómo saber que B, a su vez, sabe de lo que A está hablando (p). Para llegar a esa conclusión, en su mente A debe hacer un cálculo infinito de posibilidades. De la misma manera, B debe hacer un número infinito de suposiciones que lo llevan a interpretar correctamente la proposición de A.

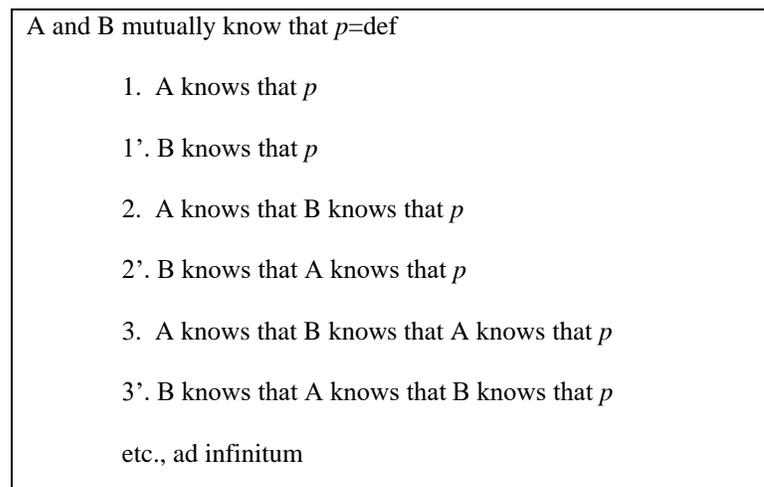


Figura 2.5.1 - Definición de "conocimiento compartido" (Schiffer, 1972, citado en Carlson, 1992:68)

⁹ Originalmente, esta definición proviene de Clark y Carlson (1981:318).

La paradoja surge porque no obstante se pase por, supuestamente, un número infinito de cálculos mentales: los interlocutores llegan a sus conclusiones en cuanto al TC que hay entre ellos en una cantidad finita de tiempo, de hecho, es una cantidad de tiempo muy pequeña. Para estos autores, la resolución de la paradoja tiene repercusiones en el uso del artículo definido, ya que este requiere que el referente le sea familiar tanto al oyente como al hablante.

Para resolver la paradoja, Clark y Marshall (1981:32) proponen procesos heurísticos de copresencia, ya que los consideran como los únicos a través de los cuales se puede establecer el TC (1981:32). Para que un referente definido pueda usarse de manera adecuada, se necesita de una triple copresencia, a saber: (1) la pertenencia a una comunidad, lo cual se refiere a que todos en una misma comunidad asumen que todos en esa comunidad saben que *p*; (2) la copresencia física, que se considera la evidencia más fuerte de TC e indica una presencia física simultánea ante un evento, y (3) la copresencia lingüística, en donde se excluye la simultaneidad y se refiere a eventos o elementos que han sido nombrados en la conversación.

Cada una de estas variedades de TC incluye una serie de presuposiciones que hacen posible los procesos heurísticos. En el caso de la pertenencia a una comunidad, las dos presuposiciones son la copertenencia, (tanto A como B saben que ambos pertenecen a una comunidad en particular), y la universalidad del conocimiento (todos en esa comunidad saben que *p*). Por ejemplo, dos personas educadas, pertenecientes a la comunidad de los Estados Unidos de América, pueden presuponer que ambos saben que George Washington fue el primer presidente de la nación, o que el estado de Colorado se encuentra al sureste de Dakota del Sur (Clark y Marshall, 1981:35-38). En el caso de la

copresencia física, hay tres presuposiciones. La primera es la inmediatez física, en donde hay una simultaneidad y atención mutua hacia algo en particular; por ejemplo, si A se refiere a una vela con la expresión *esta vela*, A está segura de que dicha expresión lingüística será suficiente para que B sepa a cuál vela se está refiriendo, siempre y cuando ambos estén mirando a la vela en cuestión al mismo tiempo. La segunda presuposición es el potencial de la copresencia física; es decir, si B no está prestando atención a la vela, pero al enunciar *esta vela*, A sabe que B prestará atención, así completa el TC basado en la copresencia física. La tercera presuposición es la copresencia física a priori. Por ejemplo, ambos han dejado de mirar la vela, pero tienen la capacidad de traerla a la memoria solo con que A enuncie *esa vela* (p. 38-39). Por último, en cuanto a la heurística de la copresencia lingüística, las presuposiciones básicas son dos: la copresencia potencial y la copresencia a priori. En primer lugar, tanto A como B deben prestar atención simultáneamente y de forma racional al enunciado de A, *una vela*. La atención simultánea es la copresencia física potencial. Además de esto, se debe poder asumir que hay un entendimiento entre A y B; es decir, se presta atención racional al enunciado. La copresencia física a priori se refiere a que, al enunciar *una vela*, B debe poder recordar y mentalmente localizar el objeto físico al cual la expresión hace referencia (p.39-40). Los autores añaden que lo más común en una situación comunicativa es el establecer TC a través de una combinación de estos procesos heurísticos (p. 40-42).

Este trabajo seminal de Clark y Marshall (1981) es teórico y se ha prestado como base para subsecuentes investigaciones empíricas. Sin embargo, existe un número limitado de estudios que han intentado aislar específicamente el TC como variable para determinar su influencia en el uso e interpretación de las expresiones referenciales en el

discurso. Los más relevantes para la presente investigación son los trabajos de Schreuder y Buttrick (1992), Smith y Jucker (1998) y Smith, Noda, Andrews y Jucker (2005). Estos trabajos son los únicos, a mi saber, cuyas investigaciones se relacionan con expresiones referenciales y donde se pretende determinar el comportamiento de los hablantes en cuanto a la selección de las formas, manipulando el TC que se asume entre los participantes.

En el estudio de Schreuder y Buttrick (1992), los investigadores analizan el TC con respecto al uso de las expresiones demostrativas referenciales en inglés. El objetivo del estudio es el de caracterizar cómo las personas entienden las expresiones demostrativas en general, pero sobre todo cuando el referente es indeterminado; o sea, cuando hay dos o más posibles referentes, como en el ejemplo (36), a continuación:

(36) Could I look at *that newspaper*.

(Schreuder y Buttrick, 1992:79)

Para entender a cuál periódico se refiere el hablante (asumiendo que haya más de uno en el momento del enunciado), no solo el oyente tiene que poder entender las palabras *that newspaper*, sino que también debe registrar el gesto que haga el hablante para indicar a cuál de ellos se está refiriendo. Este gesto puede ser una inclinación de la cabeza, una mirada, el movimiento de una mano, etc. (p.78).

Schreuder y Buttrick (1992) utilizaron dos variables en este estudio: en primer lugar, la prominencia física de un objeto, la cual se refiere a sus características físicas como lo pueden ser su tamaño, su posición, su color, etc. En segundo lugar, el TC, el cual se establece de diferentes maneras, entre ellas, la de mostrar una foto de alguien mutuamente conocido por ambos participantes vs una foto de alguien desconocido, o la

de pedirle a un participante que escoja un regalo basándose en la personalidad de una persona conocida vs una persona desconocida. Se llevaron a cabo cuatro experimentos, de los cuales solamente dos y tres son relevantes para la presente investigación; a continuación paso a resaltar estos últimos.

En el experimento dos, el cual se enfocó en la prominencia de los objetos, de las 360 respuestas que los estudiantes dieron al contestar la razón por la cual habían escogido un determinado objeto, el 75% estaba relacionado con la prominencia de este; sin embargo, en el experimento tres, en el cual el enfoque era el TC, el 90% de las respuestas estaba relacionado con el TC como motivo para escoger un objeto (p.94). Los resultados demuestran que las personas utilizan el TC para identificar un referente y, en ausencia de este, recurren a la prominencia. En otras palabras, al introducir el TC como variable en los experimentos, este tomó prevalencia sobre otros factores, incluyendo la prominencia física, y los participantes seleccionaron objetos basándose en el conocimiento compartido que asumían existía entre ellos y no con respecto a una prominencia física de los objetos proporcionados. En palabras de Schreuder y Buttrick, “all the information the listener should ever appeal to is the speaker’s and addressees’ common ground” (1992:99), con relación a la desambiguación de una referencia demostrativa indeterminada.

Aunque este estudio se enfoca exclusivamente en los demostrativos, es un trabajo en cuya metodología vemos un intento exitoso en el aislamiento del TC como factor. Por lo tanto, en relación con la presente investigación, sirve de soporte a la hipótesis cuatro del estudio que predice la posibilidad de aislamiento de esta variable.

El estudio de Smith y Jucker (1998) es relevante porque demuestra el papel de los marcos cognitivos en la selección de las expresiones referenciales, tomando en cuenta el

TC. Los autores analizan expresiones referenciales con respecto al TC y arguyen que, en una conversación, los participantes recurren al TC para manipular la accesibilidad a referentes problemáticos y analizan las estrategias que utilizan los participantes para determinar el TC. En primer lugar, un hablante selecciona un referente juzgando el nivel de TC que hay entre este y su interlocutor. Sin embargo, si hay dudas con respecto al nivel del TC, el hablante intentará negociar la información, como se ve en (37):

(37) we went to an island uhm.. *Corfu*,
 I don't know if you ever heard of it and went there,

(adaptado de Smith y Jucker, 1998:171)

En este ejemplo, el hablante demuestra inseguridad acerca del uso del referente *Corfu*, debido a una posible falta de TC entre los interlocutores. Al expresar esta duda explícitamente, el hablante se asegura de que el uso de esta determinada expresión lingüística es adecuada y accesible para el oyente. La conclusión a la que llegan estos autores es que el uso de una expresión referencial debe incluir un modelo interactivo en una conversación, en donde tanto el hablante como el oyente negocian la accesibilidad e identificación del referente, con respecto al TC que se establece entre ellos.

En este estudio se demuestra que los marcos cognitivos también juegan un papel en el correcto uso e interpretación de una expresión referencial. La idea original de los marcos proviene del concepto de “escena” de Bartlett en el 1932 y se desarrolló de especial manera en el ámbito de la inteligencia artificial con Minsky, a partir del 1974 (Fillmore, 1975:124). Fue Fillmore (1975) quien primero asoció los marcos con la lingüística y, según este autor, los marcos son escenarios estándar ligados a la cultura de la comunidad de habla y se encuentran asociados con escenas prototípicas. Los marcos se

activan con la mención de un elemento que permite las inferencias, como es el caso de la expresión *the room* en el ejemplo (33) de la Sección 2.4.

Como mencionado en la Sección 2.4, hay varios datos en el presente estudio que demuestran el uso de marcos cognitivos o esquemas en la selección adecuada de cláusulas definidas, como en el caso de los miembros individuales de la familia de la película. Pero también ocurre en otras instancias, por ejemplo, con “película” como marco cognitivo, introducido en las narrativas como *una película/un video* etc. y el subsecuente uso de las cláusulas definidas *el hilo conductor/el personaje principal/los subtítulos*. Según Smith, Jucker y Müller (2000:534), determinadas expresiones lingüísticas activan un determinado marco en la memoria, algunas frases nominales definidas provienen de tener un marco o esquema que provee la base del TC, como podemos ver en el ejemplo (38), donde el participante utiliza la expresión definida *the waiter*, ya que está narrando una escena que ocurre en un restaurante, y es esta escena lo que activa el marco que a su vez permite el uso adecuado de la frase nominal definida seleccionada por el hablante:

(38) *the waiter* came over and *the waiter* was this really big guy,

(Smith, Jucker y Müller, 2000:537)

Tal vez, el trabajo más relevante para mi investigación con respecto al TC es el de Smith, Noda, Andrews y Jucker (2005) ya que es el único en el cual se analiza cómo se introducen personajes nuevos en una narrativa cuando hay y cuando no hay presencia de TC, en línea con las hipótesis cuatro y cinco de la presente investigación. Se comparan diálogos, donde existe la posibilidad de retroalimentación que confirme el TC, y

monólogos, donde no se puede confirmar el TC, ofreciendo así una base de comparación con las narrativas escritas, donde tampoco podemos confirmar el TC con el lector.

Los participantes del estudio de Smith et ál. (2005) debían narrar parte de la película muda de Charlie Chaplin *The immigrant* (1917). Para establecer el TC, se formaron parejas y tanto el hablante como el oyente vieron juntos la primera parte de una película, luego, el hablante termina de ver la película solo y le hace un recuento al oyente de toda la película. En la primera parte del videoclip aparecen Chaplin y una mujer, y estos personajes hacen parte del TC, mientras que, en la segunda parte, el resto de los personajes no hacen parte de la información compartida, entre ellos un artista. Para los monólogos, el hablante le debe hacer un recuento a un compañero “imaginario”, por ejemplo, como si estuviera haciendo una grabación en un contestador telefónico. Esta metodología se asemeja a la del presente trabajo, donde se les pidió a los participantes que le hicieran un recuento de la película a un amigo ficticio.

Con los resultados de Smith et ál (2005), se notaron varias diferencias en la introducción de los personajes según la prominencia del TC. Para introducir a Charlie Chaplin, quien hace parte del TC, se usa una estrategia incremental donde no hay una introducción clara y discreta, sino que se hacen una serie de pre-introducciones, para luego pasar a nombrar al personaje con el uso del nombre propio o de un pronombre. En (39) vemos que la presencia de Chaplin se da como dada, al usar estrategias de pre-introducción:

(39) they were tied on the ship
and everybody one by one was given a ticket and
and let out of the ship to big America
but *Charlie Chaplin* doesn't have any money left

(Smith et ál, 2005:1875)

Para personajes sin TC, como un artista que aparece a un cierto momento de la película, la primera mención es a través de expresiones indefinidas generales:

(40) so then *some other guy*
 he is kind of like chubby

(Smith et ál, 2005:1881)

En este mismo ejemplo (40), cinco enunciados más adelante se hace referencia a este personaje como *an artist* y seis enunciados después, una vez establecido el TC, se utiliza la expresión referencial definida *the artist*.

En el caso de los monólogos, el hablante no tiene oportunidad de confirmar el TC con el oyente, acercándose esta situación a las narraciones escritas del presente trabajo. Los participantes resuelven este problema a través de diferentes estrategias, ya sea creando una audiencia imaginaria o asumiendo una audiencia informada. Varios de los resultados obtenidos por Smith et ál. (2005) se repiten en la presente investigación, principalmente, el hecho de que los participantes crean una audiencia imaginaria durante una especie de preintroducción en las narrativas. En estas preintroducciones, el escritor reconoce el TC o falta de este antes de pasar a narrar los eventos de la película. Un ejemplo de audiencia imaginaria de Smith et ál. (2005) lo vemos en (39), más arriba, en donde el participante, a través de un prólogo introductorio, le habla directamente a una audiencia imaginaria con un lenguaje interactivo a través del uso de pronombres personales que incluyen a la audiencia (*you*) y de marcadores del discurso (*well*). Más adelante en (41), este mismo participante pasa a pre-introducir a Chaplin, el personaje principal, como el director/productor de la película:

(41) *well* it's too bad that *you* missed the movie,
 it's actually pretty famous as most *Charlie Chaplin* movies are,

Es solo más adelante que el participante introduce a los personajes como tales en la narrativa:

(42) ...erm basically the major characters, there was *himself* and then
this girl,

(Smith, Noda, Andrews y Jucker, 2005:1885)

En este estudio vemos que las expresiones referenciales que el hablante escoge dependen del TC entre los interlocutores. También dependen de la prominencia que el hablante le quiera dar a ese referente: los personajes principales tienen expresiones referenciales diferentes a los personajes secundarios. A esta misma conclusión llegó Saunders (1999) en su tesis doctoral con respecto a la importancia de los personajes en la narrativa y cómo esta prominencia afecta la selección de las expresiones referenciales. El estudio de Saunders no está basado en el aislamiento del TC específicamente y, al ser un estudio de adquisición de lengua, lo retomaré en el Capítulo 3.

En conclusión, hay algunos estudios que han demostrado que el TC juega un papel fundamental en el uso adecuado y en la interpretación de expresiones referenciales. Sin embargo, todos utilizan exclusivamente datos orales y, a mi saber, no hay ninguno hasta el momento que se haya enfocado específicamente en las expresiones referenciales de sujeto para el español, así como tampoco sé de ninguno que compare HNs con hablantes no-nativos (HNNs). De hecho, el trabajo de Smith et ál. (2005) utiliza 191 participantes entre HNs de inglés y estudiantes de inglés como L2, pero estos autores no especifican el número de participantes en cada grupo, así como tampoco diferencian entre sus respuestas. Además, no describen el nivel de adquisición de lengua de los estudiantes, lo cual forzosamente tiene un efecto en los resultados del estudio, el cual no se establece en

su investigación. El estudio de Gundel et ál. (1993) sobre el cual se basa el marco teórico de esta investigación, también tiene un hueco significativo en su metodología. En ese trabajo, los autores utilizaron datos provenientes de diferentes fuentes, entre datos orales y escritos, de diversos grados de planeación y géneros, así como datos que estos autores habían obtenido para otros estudios (ver nota a pie de página en Gundel et ál., 1993:290). Es de esperar que, al mezclar variables como modalidad, género y nivel de planeación, no podamos saber si los diferentes esfuerzos cognitivos por parte de los participantes afectan los resultados. Además, en el trabajo de Gundel et ál. (1993) aun cuando se reconoce el TC como factor pragmático-cognitivo, no se aísla para entender su posible efecto en la selección y uso de expresiones referenciales. En el presente estudio se pretende llenar estos huecos al diferenciar claramente entre la competencia lingüística de los participantes, así como la modalidad (escrita u oral) utilizada y al considerar el TC como variable independiente.

En el Capítulo 3, a continuación, voy a analizar los conceptos introducidos en las secciones anteriores desde la perspectiva de la adquisición de segunda lengua, de nuevo con un énfasis en la adquisición del español como L2.

CAPÍTULO 3

3 EL ESTUDIO DE LAS EXPRESIONES REFERENCIALES DE SUJETO EN LA ADQUISICIÓN DE SEGUNDA LENGUA.

3.1 Introducción

Durante el proceso de adquisición de una segunda lengua (ASL), el estudiante desarrolla e incorpora de manera gradual las características de la lengua meta. Según Ellis (1985:58), durante este proceso el estudiante aprende a construir un discurso con diferentes niveles de organización y en su desarrollo el estudiante internaliza y almacena nuevas formas y crea relaciones de forma-función en su interlengua. Entendemos por interlengua un sistema lingüístico distinto tanto de la L1 como de la L2 de un hablante adulto, el cual consta de su propia gramática, conocida como *interlanguage grammar* (Ritchie y Bhatia, 1996:8).

La adquisición de las expresiones referenciales de sujeto forma parte del proceso para la construcción del discurso en una segunda lengua. Su estudio ha sido extensivo en el español y se ha llevado a cabo desde diferentes perspectivas lingüísticas que reflejan las propiedades sintácticas, pragmáticas y cognitivas que vimos en el Capítulo 2 de este trabajo. Los modelos teóricos presentados en el anterior capítulo informan el estudio de la ASL de manera integral ya que consideran los diferentes factores que afectan la selección de las expresiones referenciales de sujeto. Principalmente, vimos que la Givenness Hierarchy (GH) de Gundel, Hedberg y Zacharski (1993) permite llevar a cabo un análisis con un enfoque funcional-cognitivista, teniendo en cuenta factores pragmáticos como el

tópico y el foco, así como factores sintácticos de prominencia relacionados con la posición del sujeto en la oración, planteados por la hipótesis de la posición del antecedente (PAH) de Carminati (2002). Subsecuentes investigaciones han intentado aplicar estos modelos al estudio de la ASL, y de esta manera, el estudio de la adquisición del español como L2 ha sido objeto de varias investigaciones relevantes para el presente trabajo.

En este capítulo voy a revisar algunos de los estudios que reflejan los diferentes enfoques lingüísticos dentro de la ASL, con un énfasis en la adquisición de sujetos en el español como L2. En la primera sección, hablaré de la adquisición de los sujetos desde la perspectiva del generativismo, los cuales están relacionados con las propiedades sintácticas revisadas en la sección 2.1. El generativismo surge primeramente como una teoría que pretende explicar la adquisición de la primera lengua en los niños (L1). Este enfoque tiene dos objetivos principales: en primer lugar, intenta describir las propiedades que ofrece la gramática universal (GU) en los estados iniciales de la adquisición (Rothman y Pascual y Cabo, 2014:47). Este enfoque considera la complejidad de la adquisición y su eventual éxito a pesar del input ambiguo e incompleto que reciben los niños; a esto se le denominó *la pobreza de estímulos o el problema lógico del aprendizaje* (Sánchez y Toribio, 2003:190). En segundo lugar, la lingüística generativa intenta explicar los pasos por los cuales pasa un individuo desde el estado inicial hasta la adquisición final o el problema del desarrollo de la adquisición (Rothman y Pascual y Cabo, 2014:47). Aquí, los investigadores buscan entender cómo las habilidades cognitivas de los niños interactúan con las muestras de lenguaje recibidas del entorno durante el proceso de adquisición (Sánchez y Toribio, 2003:190). Es importante aclarar

una idea errónea que se tiene acerca del generativismo: la GU no rechaza el entorno como fuente de adquisición; de hecho, tanto el léxico como algunas formas específicas del lenguaje derivan de este en su totalidad. Sin embargo, y contrario a otros enfoques lingüísticos, la GU mantiene que los humanos vienen equipados con el material lingüístico necesario que permite el desarrollo del lenguaje en un individuo (Rothman y Pascual y Cabo, 2014:47) y es de aquí que se deriva la idea del innatismo del lenguaje. En el estudio de la adquisición de la L2, el generativismo reexamina el problema lógico y el problema del desarrollo con respecto al acceso total, parcial o falta de acceso a la GU por parte del adulto (Sánchez y Toribio, 2003:190).

En la segunda parte describiré la adquisición de sujetos en el español como L2 desde el enfoque variacionista, el cual es un modelo inspirado por el sociolingüista William Labov en los años 60 con su trabajo sobre el inglés de hablantes nativos (HNs) y que hoy en día abarca otras lenguas, diferentes comunidades de habla (tanto L1 como L2 e inclusive bilingües) y varios fenómenos lingüísticos (Gudmestad, 2014:80). En el modelo variacionista, el planteamiento va más allá del nivel oracional, lo cual significa que se incluyen factores extralingüísticos como el nivel socioeconómico del hablante, la distancia social, el contexto social, el género, la L1 del hablante, entre otros factores que veremos más adelante. Estos estudios se basan en investigaciones sociolingüísticas en donde se tienen en cuenta aquellos factores que restringen las elecciones lingüísticas del hablante. En cuanto a los sujetos en español, las selecciones pueden estar motivadas por la referencia a la forma del verbo, la continuidad de tópico, la perseverancia (*priming*), la especificidad del sujeto, etc. La hipótesis principal de varios de estos estudios consiste en que la variabilidad en la L2 está restringida por muchos (si no todos) de los mismos

factores que restringen la elección en la L1 (Lubbers Quesada, 2015:265-266). Aunque en el presente estudio no se toman en cuenta variables sociolingüísticas de manera independiente, es importante reconocer que hay otros factores que motivan las elecciones de los hablantes, como el aspecto, tiempo y modo, entre otros, además de los factores pragmático-cognitivos que usamos como marco de referencia para esta investigación.

En la tercera parte hablaré del funcionalismo. Zyzik (2014:30) describe al funcionalismo no como a una teoría lingüística en sí, sino como una orientación, y añade que, a pesar de ello, los objetivos de los lingüistas funcionalistas generalmente contrastan con los generativistas desde varios puntos de vista. El primero de estos contrastes es la inseparabilidad de la sintaxis del significado, del discurso y del lenguaje en uso; en segundo lugar, para los funcionalistas la estructura del lenguaje se basa en la función del lenguaje, lo cual implica que varios de los fenómenos lingüísticos con los cuales nos encontramos, pueden ser explicados por fuerzas externas al mismo, como por ejemplo la eficiencia conversacional, el esfuerzo del procesamiento y la frecuencia de uso. Esto equivale a decir que el uso del lenguaje afecta la estructura del mismo, y los lingüistas funcionalistas, contrariamente a los lingüistas con un enfoque formal, estudian cómo se derivan las estructuras gramaticales, sintácticas y textuales a partir de la manera cómo se usa el lenguaje.

En la cuarta y última sección veremos cómo la lingüística cognitiva (LC) ha llevado a cabo los estudios de adquisición de sujeto. La LC surge en los años 80 como reacción al paradigma chomskiano debido a que varios de sus proponentes (estudiantes de Chomsky), entre los cuales sobresalen George Lakoff y Ronald Langacker, no estaban de acuerdo con el alcance limitado del modelo generativo y el hecho de que descuidara o

desatendiera a los aspectos cognitivos y sociales de la comunicación lingüística (Tomasello, 1998:8). La LC también se caracteriza por reconocer a la lengua como un fenómeno mental, pero se distancia de la autonomía de la sintaxis del generativismo (Cadierno y Pedersen, 2014:16-17). Además, y también contrario al generativismo y en línea con el funcionalismo, según la LC los niños adquieren el lenguaje a través del uso, concepto conocido como lenguaje basado en el uso (*usage-based language*); es decir, el lenguaje no se aprende de un conocimiento abstracto e innato, sino de enunciados particulares en contextos particulares (Lieven y Tomasello, 2008:168).

Es importante enfocarse en estos cuatro enfoques porque históricamente las investigaciones sobre la adquisición y uso de sujetos en español se han fundamentado mayormente en una de estas perspectivas. Aunque las metas generales y las indagaciones específicas de cada enfoque son distintas, cada uno aporta elementos sustanciales acerca de la expresión de sujetos en español y del proceso de su adquisición.

El estudio de Gundel et ál. (1993) que se usa como punto de partida para la presente investigación, puede caracterizarse como funcionalista o cognitivista, según se entiendan estos conceptos, pero en línea con el enfoque teórico adoptado para este trabajo, también podemos considerarlo cognitivo-funcionalista. Esto se puede justificar porque la GH se basa en los factores de memoria y atención, que como veremos en este capítulo son de naturaleza cognitivista, y al mismo tiempo, en la GH se comienza con la noción de *referente* y se examinan las formas (p. ej. cláusulas léxicas, pronombres explícitos, pronombres nulos) usados por los hablantes, metodología de carácter funcionalista.

3.2 La adquisición de sujetos en el generativismo

El reduccionismo del enfoque generativista no permite, en mi opinión, un análisis completo en cuanto a la selección y uso de las expresiones referenciales, principalmente porque no tiene en cuenta el carácter comunicativo del lenguaje y cómo el uso afecta las elecciones de los hablantes. Sin embargo, el estudio de las interfaces que se deriva del generativismo ha contribuido a un mayor entendimiento de los problemas en la adquisición de una segunda lengua. En las interfaces se valoran las explicaciones semánticas, pragmáticas y discursivas al igual que las sintácticas, y es a partir de estos trabajos que podemos encontrar un aporte substancial para este estudio, en particular con respecto a las dificultades de adquisición de las propiedades pragmáticas sobre las sintácticas y las explicaciones que estos estudios han aportado. Para entender las interfaces, sin embargo, hay que empezar yendo más atrás con las investigaciones relacionadas con el acceso a la GU y los estudios paramétricos del sujeto nulo en español dentro del campo de la adquisición de la L2.

Mientras la adquisición de la L1 se caracteriza por una adquisición final uniforme entre individuos, en la L2 se ve variabilidad en el éxito final. Rothman y Pascual y Cabo (2014:48) dicen que a pesar de algunas diferencias obvias entre la L1 y la L2 (distintos estados iniciales, variación en la madurez cognitiva y en los factores de procesamiento), la tarea en sí de la adquisición de los adultos no es demasiado diferente de aquella de los niños. En general, en el generativismo estas diferencias entre L1 y L2 se suelen entender en términos de la accesibilidad (o falta de) a la GU (Sánchez y Toribio, 2003:190); es decir, si después del periodo crítico el adulto tiene acceso o no al denominado *dispositivo de adquisición del lenguaje* (DAL), entendido en la tradición chomskiana como lo

equivalente a la capacidad y mecanismo de adquisición del lenguaje (Liceras y Díaz, 2000:42). En un principio, los investigadores daban respuestas dicótomas al respecto: “either there was (e.g., Flynn 1987; Schwartz 1986; White 1985) or was not adult UG accessibility (e.g., Bley-Vroman 1989; Clahsen and Muysken 1986)” (Rothman y Pascual y Cabo, 2014:50-51). Sin embargo, tanto en Rothman y Pascual (2014:50) como en Rothman e Iverson (2007:329) se afirma que se han hecho subsecuentes estudios que abogan por respuestas menos binarias ante preguntas complejas como esta, dando cabida a explicaciones en torno a un acceso parcial a la GU, esto es, los principios de la GU están disponibles pero no los parámetros, o modelos que describen problemas en el proceso de transferencia de la L1 (por ejemplo del sistema prosódico), problemas específicos de aprendizaje (problemas de *mapping*), o problemas en las interfaces (p. ej. la interfaz sintaxis-pragmática).

Los estudios generativistas buscan evidencia acerca de la adquisición del parámetro del sujeto nulo (PSN) y se basan principalmente en la distribución entre el sujeto nulo y el sujeto explícito, ignorando otras formas como por ejemplo las frases nominales (FNs), los nombres propios y los demostrativos. En la Sección 2.2 mencioné el trabajo de Rothman e Iverson (2007) al describir las propiedades sintácticas de los sujetos. En este estudio los autores investigan si periodos prolongados e intensos de exposición a la L2 con nativos ayuda al reinicio del PSN y proponen un número más restringido de parámetros. Estos autores demuestran que no todas las propiedades se subsumen en la primera (el sujeto referencial puede ser nulo o explícito), como es la hipótesis común. En

cambio, proponen que el PSN consta de un número más restringido de propiedades, dos (números 1 y 2, más abajo) o máximo tres (OPC)¹⁰:

1. El sujeto referencial puede ser nulo o explícito.
2. El sujeto expletivo es obligatoriamente nulo.
3. Posibilidad de inversión sujeto-verbo.
4. No violación de la huella *-qu*.
5. La restricción manifiesta de pronombre (OPC).

(Adaptado de Rothman e Iverson, 2007:329)

Rothman e Iverson (2007) explican que en los estudios paramétricos se planteaba la hipótesis de que los aprendientes de una lengua solo tenían que adquirir una de estas propiedades para adquirir el PSN en su totalidad, ya que cada propiedad subsume las otras. Estudios posteriores arguyeron por una adquisición gradual de las propiedades. Se parte de la base de que la gramática está regida por la GU y de que es posible reiniciar los parámetros, pero que no siempre se ha demostrado que haya adquisición de todas las propiedades como un conjunto; es decir, no todos a la vez. En su estudio con HNs de inglés y aprendientes de español como L2, Rothman e Iverson (2007) encontraron evidencia de que los aprendientes de español como L2 tienen la capacidad de reiniciar el PSN ya que en los datos obtenidos los participantes demostraron el uso del sujeto nulo, mientras que el resto de las propiedades se van adquiriendo poco a poco, como la posibilidad de inversión sujeto-verbo y la no violación de la huella *-qu*, demostrando de esta manera que no hay *clustering* de propiedades que se subsumen a la primera

¹⁰ Ver sección 2.1 para una descripción de las propiedades del PSN.

propiedad; por lo tanto, no solamente se aprenden de manera gradual, sino también en un orden específico.

La propuesta de Rothman e Iverson (2007) tiene repercusiones en los estudios de adquisición de lengua, entre otras razones porque animan el debate de si los estudiantes de L2 llegan a adquirir completamente las propiedades sintácticas. Este debate ya había comenzado con anterioridad debido a la observación por parte de algunos autores de que, incluso a niveles avanzados de adquisición, se notan déficits residuales en la producción del lenguaje que no parecen provenir de una falta de adquisición de las propiedades sintácticas. De esta manera, se motivan los subsecuentes trabajos relacionados con las interfaces sintáctico-discursivo/pragmáticas. Algunos de los estudios más prominentes acerca de estas cuestiones son los de Pérez-Leroux y Glass (1999), Montrul (2004), Montrul y Rodríguez Louro (2006), Sorace y Filiaci (2006), Lozano (2008), Rothman (2009), Filiaci (2010), Jegerski, VanPatten y Keating (2011) y, más recientemente, Lozano (2016). Estos estudios son relevantes en mi investigación porque sirven de apoyo al hecho de que se necesita de un enfoque que va más allá del análisis puramente sintáctico, el cual se ve limitado en sus explicaciones dentro del análisis del discurso. A continuación, vamos a revisar los trabajos más significativos para entender el planteamiento de las interfaces en el estudio de las expresiones referenciales.

La hipótesis de la interfaz (HI) fue inicialmente propuesta por Sorace y Filiaci (2006) en un estudio con estudiantes avanzados de italiano como L2, y postula que “narrow syntactic properties are completely acquirable in a second language, even though they may exhibit significant developmental delays, whereas interface properties involving syntax and another cognitive domain may not be fully acquirable” (p.340). Lo anterior

significa que, aun a niveles avanzados de adquisición, se ven errores o déficits residuales que son comportamientos no nativos los cuales se derivan del costo de integrar propiedades sintácticas y discursivas simultáneamente. De tal manera, la HI plantea que las propiedades sintácticas son más fáciles de adquirir que aquellas discursivo-pragmáticas.

Aunque el estudio de Pérez-Leroux y Glass (1999) es anterior a la formulación de la HI, es relevante porque demuestra que las propiedades sintácticas se adquieren más temprano en el proceso de adquisición que las propiedades discursivas. La metodología de su estudio consistió en crear dos contextos: uno de ellos tuvo en cuenta la hipótesis de la OPC, el cual incluyó una serie de ocho historias seguidas por oraciones que los participantes debían traducir, y el segundo contexto tuvo en cuenta los factores discursivo-pragmáticos de tópico/foco (ver sección 2.3 para una definición y descripción de estas nociones). La hipótesis de la OPC se presta de manera ideal para estudios sintácticos puesto que epitomiza la pobreza de estímulos desde la perspectiva de la gramática generativa. Esto se debe a que es una propiedad que no se aprende a base de input: a pesar de que los niños no escuchan evidencia negativa en contra de la OPC, adquieren este parámetro (Zyzik, 2017:41). Esto tiene consecuencias para las investigaciones de ASL, ya que se considera que, si los estudiantes demuestran adquisición de este parámetro, es evidencia de que tienen acceso a los principios innatos de la GU (Rothman, 2009:952).

En el contexto de la OPC hubo 98 participantes distribuidos entre un grupo de nativos y tres grupos de estudiantes (principiantes, intermedios y avanzados). Pérez-Leroux y Glass utilizaron dos tipos de historias: las primeras cuatro tenían sujetos

referenciales (en donde hay un referente prominente que sirve de antecedente al pronombre de sujeto) y las otras cuatro de tipo variable ligada (*bound variable*) con una interpretación obligatoria de correferencia entre el pronombre nulo de la cláusula subordinada y el sujeto de la cláusula principal (p. 232). En el primer tipo de historias puede ocurrir un pronombre explícito o léxico en la traducción, y en el segundo tipo de historias se espera el uso de un pronombre nulo. En ambos casos se usaron cuantificadores del tipo *nobody*, *everybody* o *no journalist*, en el test de traducción de inglés a español.

Un ejemplo de este tipo de traducción de una oración variable ligada se ve en (1), y en (2) vemos un ejemplo de traducción de las historias referenciales. En ambos casos, antes de la oración a traducir, se les presentó una pequeña historia que proporcionaba el contexto adecuado:

(1) A traducir: No journalist admitted that *he* had talked to the jurors.

Traducción: Ningún periodista admitió que *pro* había hablado a los jurados.

(2) A traducir: But no journalist said that *he* is guilty.

Traducción: Ningún periodista dijo que *él* era culpable.

(Adaptado de Pérez-Leroux y Glass, 1999:232-233)

El estudio de tópico/foco incluyó a 101 participantes (mismos niveles que la prueba de la OPC), y consistió en historias ilustradas seguidas por dos tipos de preguntas, cuyo propósito era el de guiar la interpretación. El ejemplo (3) que sigue muestra que la pregunta de sujeto se relaciona con el foco de la oración, y por lo tanto requiere un sujeto explícito, mientras que la pregunta de objeto se relaciona con el tópico y permite el sujeto nulo para referirse al referente *ella*:

(3) Hace calor y la familia va al jardín.

- Pregunta de sujeto:
¿Quién piensa la abuela que regará las plantas?
- Respuesta esperada de foco:
La abuela piensa que **ella** regará las plantas.

- Pregunta de objeto:
¿Qué piensa la abuela que hará en el jardín?
- Respuesta esperada de tópico:
La abuela piensa que **pro** regará las plantas.

(Adaptado de Pérez-Leroux y Glass, 1999:236)

Los resultados muestran que la OPC opera a todos los niveles de adquisición: todos los grupos, tanto estudiantes como nativos, discriminaron entre sujetos nulos y explícitos según el tipo de historia. Como esperado, en las historias de tipo variable ligada se ve una mayor proporción de pronombres nulos que en las historias referenciales, además, la producción de sujetos nulos en las historias de tipo variable ligada entre los grupos incrementa a medida que aumenta la competencia lingüística, como vemos en la Tabla 3.2.1:

Tabla 3.2.1 - Porcentaje de sujetos nulos en historias de tipo variable ligada y en historias referenciales en el estudio de la OPC (Adaptado de Pérez-Leroux y Glass, 1999:234)

	Sujeto nulo Historias variable ligada	Sujeto nulo Historias referenciales
Principiantes	57.7%	21.2%
Intermedios	73.8%	35.7%
Avanzados	93.1%	58.3%
Nativos	85.0%	31.3%

Esto demuestra que el conocimiento sintáctico de la distribución del sujeto nulo existe a partir de los niveles principiantes, y que el input nativo no tiene un efecto significativo en su adquisición incremental, resultado en línea con los datos obtenidos por

Rothman e Iverson (2007). Sin embargo, la prueba de ANOVA reveló que el efecto del grupo en la diferencia proporcional en los sujetos nulos entre cada contexto no es significativo ($p=0.1151$), mientras que sí lo es en el estudio de tópico/foco ($p.240$), como veremos a continuación.

En el contexto de tópico/foco los resultados muestran que todos los grupos discriminaron en el uso de pronombres entre las historias de foco y tópico, y a medida que avanza la competencia lingüística también aumenta la producción de sujetos nulos en el contexto de tópico y disminuye en el contexto de foco. En las historias de tópico el incremento de sujetos nulos es como sigue: 30.8% principiantes, 36.3% intermedios, 57.1% avanzados y 47.4% nativos. Contrariamente al experimento de la OPC, el efecto de grupo entre contextos es significativo ($p=0.0017$), lo cual demuestra una adquisición incremental del sujeto nulo a medida que aumenta la competencia lingüística: “the effect of groups was significant, suggesting that substantial learning takes place” (Pérez-Leroux y Glass, 1999:240).

Las autoras concluyen que, a raíz del hecho de que en los contextos de la OPC la discriminación de las formas de sujeto ha sido mayor a partir de los niveles iniciales, hay evidencia de que las restricciones sintácticas se aprenden más temprano que aquellas reguladas por el discurso; esto se debe a que se ve el efecto de la OPC a partir de los niveles iniciales sin necesidad de input, mientras que la adquisición de los factores discursivo-pragmáticos son propiedades que deben aprenderse y que se desarrollan de manera incremental a medida que aumenta la competencia lingüística en la distinción entre los contextos de tópico y foco.

Es de destacar el trabajo de Montrul y Rodríguez Louro (2006) porque estas autoras investigan específicamente la interfaz sintáctico-discursiva y tienen en cuenta las nociones de cambio de tópico y de continuidad de tópico que, como hemos resaltado con anterioridad, son parte integral del análisis en la presente investigación. El estudio incluyó 48 estudiantes de español L2 de niveles intermedio y avanzado y 20 hablantes nativos. A través de una prueba de producción oral, las autoras analizaron los niveles de adquisición de las propiedades morfosintácticas y de las discursivo-pragmáticas en la adquisición de sujetos en el español. La pregunta principal era si estas propiedades se adquieren al mismo tiempo durante el desarrollo de la interlengua. Similar al estudio de Pérez-Leroux y Glass (1999), las autoras consideraron la distribución de los pronombres de sujeto nulos y plenos según si el sujeto es tópico, si introduce un cambio de referente o si establece foco o contraste. Presentan los siguientes ejemplos:

(4) Mismo referente:

Pepe no vino hoy a trabajar. *Pepe/?él/*Pro* estará enfermo.

(5) Cambio de referente:

Hoy no fui a trabajar. Pepe/él/**Pro* pensó que estaba enferma.

(6) Tópico:

¿Quién vino? Él/Mario/**Pro* vino.

(7) Foco:

El periodista_i dijo que él_i no había escrito ese reporte.

(Montrul y Rodríguez Louro, 2006:404)

Los resultados indican que los estudiantes más avanzados demuestran una adquisición casi nativa con respecto tanto de las propiedades morfosintácticas, como de

aquellas discursivo-pragmáticas; mientras que en los estudiantes intermedios hay una mayor influencia de la L1, ya que se ve una adquisición de las primeras, pero resultados incorrectos de las segundas. Con sujetos explícitos, el grupo intermedio produjo 22.9% de sujetos redundantes, porcentaje que disminuye a 7.6% en el grupo avanzado y a 0.3% en el grupo casi-nativo (p.412). Las autoras concluyen que la adquisición es gradual: hay evidencia de reinicio del PSN del inglés al español ya que los estudiantes adquieren las propiedades morfosintácticas, y que una vez que esto sucede, los estudiantes comienzan a adquirir también las propiedades discursivo-pragmáticas. Esto quiere decir que la adquisición de ambas propiedades no es paralela. Además, aunque los estudiantes llegan a adquirir completamente las propiedades sintácticas, se notan déficits residuales de las propiedades sintáctico-discursivas incluso en niveles avanzados (p.413), evidencia de que éstas últimas son más difíciles de adquirir.

El estudio de Lozano (2008) es particularmente revelador dentro de este análisis porque utiliza participantes de dos lenguas [+pro-drop], el español y el griego, por lo tanto, no deberían encontrarse problemas de adquisición del español como L2 por parte de los griegos. El autor se centró en los déficits que afectan la representación lingüística (la sintaxis) y los déficits de procesamiento (interfaz sintáctico-discursiva). Su hipótesis fue que si los déficits afectan la sintaxis, se espera un comportamiento nativo de los estudiantes, al ser la L1 y la L2 idénticas en cuanto a su propiedad de [+pro-drop], pero, si los déficits son de procesamiento, se esperan divergencias.

En el estudio participaron 12 hablantes nativos de español y 85 griegos nativos aprendices de español L2 de tres niveles diferentes de competencia: intermedio, avanzado-bajo y avanzado-alto. El autor utilizó dos condiciones en pruebas de juicio de

aceptabilidad contextualizadas: la condición de tópico, donde se espera un mayor uso del pronombre nulo, y la condición de foco contrastivo, donde se espera un mayor uso del pronombre explícito. En la condición de foco contrastivo los estudiantes de nivel más avanzado muestran un comportamiento casi nativo; sin embargo, en la condición de tópico se notan déficits residuales al utilizar un pronombre explícito, lo cual Lozano (2008:863) atribuye a una relajación de la PAH (ver Sección 2.2 para una descripción completa de esta hipótesis) y explica que los déficits son residuales ya que solo afectan a los pronombres plenos.

Los resultados del estudio demuestran que los déficits se encuentran en la interfaz sintáctico-discursiva, puesto que los problemas son de procesamiento (p.862): a pesar de que los estudiantes son capaces de utilizar tanto sujetos nulos como explícitos desde niveles principiantes, aun en el nivel avanzado se ven déficits residuales de sobreuso del sujeto explícito en la condición de tópico. En otras palabras, la sintaxis se adquiere antes que los aspectos discursivos y, por lo tanto, los aspectos morfosintácticos son de adquisición más fácil y temprana.

Más recientemente, es notable el trabajo de Lozano (2016) ya que utiliza un corpus escrito (Corpus Escrito del Español L2, CEDEL2) creado a través de la narración de una película. Debido a que no es común encontrar estudios que tengan en cuenta datos escritos, este trabajo me permite tener un punto de comparación, sobre todo a nivel metodológico y de resultados; de hecho, Lozano justifica su metodología al decir que la resolución anafórica ha sido observada tanto en corpus orales como en aquellos escritos (Bel, Perera y Salas, 2010, Perales y Portillo, 2007, citados en Lozano 2016:243). Además, el estudio de Lozano (2016) es relevante porque no solamente considera los

pronombres nulos y explícitos, sino que también incluye sintagmas nominales (SNs), que, de nuevo, no es común en estudios con un enfoque sintáctico. Según Lozano (2016:236), la resolución anafórica está relacionada a cómo una expresión anafórica, ya sea esta un SN, un pronombre explícito o un pronombre nulo, hace correferencia con su antecedente en el discurso y dice que hay factores sintácticos y discursivos que restringen la forma de la anáfora. Además, arguye que no todas las propiedades sintáctico-discursivas de la interfaz son igualmente problemáticas en términos de resolución anafórica: situaciones de foco contrastivo llegan a una adquisición nativa, mientras que aquellas de continuidad de tópico muestran déficits residuales (p.238), como también lo demostró en Lozano (2008).

Los participantes del estudio fueron 10 nativos de español, y 10 estudiantes muy avanzados de español L2, inglés L1. Los resultados demuestran que tanto los nativos como los aprendientes utilizan (de manera casi idéntica) SNs para marcar cambios de tópico (69.9% L1 vs 74.1% L2) y para introducir nuevas entidades en el discurso (13.4% L1 vs 27.7% L2). En este contexto, los SNs son la forma más utilizada, seguidas por sujetos explícitos. En contextos de continuidad de tópico, el sujeto nulo es el más utilizado (93.3% L1 vs 80.4% L2), como es de esperar. De crucial importancia son los resultados que evidencian el uso pragmático de manera lícita: 96.4% de los nativos hacen un uso pragmáticamente adecuado de los sujetos vs 82.7% de los participantes avanzados (p.251-252).

Desafortunadamente, al haber utilizado únicamente un corpus escrito, Lozano (2016:262) reconoce que no se pueden hacer comparaciones con datos orales, y por lo tanto no se puede llegar a conclusiones con respecto a la posible influencia que puede tener el modo de producción (escrito vs oral) en la selección de una expresión referencial

de sujeto. Este es un hueco que la presente investigación intentará llenar, al incluir en la metodología tanto datos orales como escritos. De esta manera vamos a poder observar directamente si la modalidad afecta de manera significativa las elecciones de los hablantes.

En conclusión, vemos que las investigaciones con carácter generativista han ampliado su enfoque desde los estudios más restringidos de la adquisición/reinicio del PSN, hasta incluir consideraciones discursivo-pragmáticas a través del estudio de las interfaces. Estas aportaciones son importantes para la presente investigación porque demuestran que las propiedades sintácticas y las propiedades pragmático-discursivas deben considerarse de manera integral. Los estudios que hemos descrito confirman que hay abundante evidencia de que los aprendientes de una L2 tienen la capacidad de adquirir las propiedades sintácticas, pero aquellas pragmático-discursivas se adquieren con un mayor grado de dificultad. Sin embargo, en su gran mayoría estos estudios se siguen limitando a la distribución binaria entre los sujetos nulos y explícitos, aunque como vemos con Lozano (2016), hay unos esfuerzos recientes por incluir otras formas referenciales, acercándose a los estudios con un enfoque cognitivo-funcionalista.

3.3 La adquisición de sujetos en el variacionismo

Al igual que el generativismo, el variacionismo se originó con el estudio de la L1 en el inglés de los nativo-hablantes, pero hoy en día abarca otras lenguas, diferentes comunidades de habla y varios fenómenos lingüísticos (Gudmestad, 2014:80). También se ha aplicado al estudio de la ASL con trabajos como los de Gudmestad y Geeslin (2010), Abreu (2012), Geeslin, Linford, Fafulas, Long, y Díaz-Campos (2013),

Gudmestad, House y Geeslin (2013) y Otheguy (2014), entre otros. Según Gudmestad (2014:81), el variacionismo es útil en el estudio de una L2 porque esta comparte con la L1 tres principios: (1) un sistema lingüístico que está gobernado por reglas, (2) experimenta cambios y (3) se ve impactado por factores sociales.

Desde esta perspectiva se busca identificar aquellos factores tanto internos como externos que influyen la frecuencia de uso de una variante. Algunos de los factores externos que se tienen en cuenta en los estudios variacionistas sobre la adquisición de sujetos en español L2 son el género, el estatus social del hablante, la etnicidad, el nivel de educación, la edad, la L1 del hablante y el efecto individual del participante. Entre los factores lingüísticos relacionados con sujetos, el variacionismo tiene en cuenta variables como el cambio del referente, la especificidad, el tiempo/aspecto/modo (TAM) del verbo, la perseverancia, entre otros. Además, se usan análisis estadísticos para determinar la variación diacrónica y sincrónica, lo cual ayuda a predecir cuándo una variante llega a ocurrir. Aunque estos no son factores que se tienen en cuenta en el presente estudio, es importante reconocer que todas las anteriores son propiedades válidas de estudio para un entendimiento lingüístico completo de las elecciones de los hablantes.

Una contribución importante de los estudios variacionistas han sido las metodologías que se han introducido. Las investigaciones dentro de este enfoque han llevado al desarrollo de un método estadístico muy sofisticado para el análisis lingüístico cuantitativo de la producción del lenguaje, y se ha convertido en una fuente importante de información acerca del cambio lingüístico en progreso (Ritchie y Bhatia, 1996:14). Las técnicas investigativas desarrolladas permiten determinar la frecuencia en la cual ocurre una determinada forma lingüística en un individuo o en una comunidad de habla. Estas

técnicas han sido útiles en los estudios de ASL, incluyendo el presente trabajo, ya que nos ayudan a entender los cambios que ocurren durante el proceso de adquisición. Los modelos de regresión requieren de una cooperación entre el lingüista y el estadístico lo cual ha fortalecido las colaboraciones interdisciplinarias. La metodología en esta disertación se ha beneficiado de un trabajo conjunto entre la investigadora principal del proyecto y un equipo de estadísticos que han diseñado modelos de análisis específicos para este trabajo (ver Capítulo 4).

A diferencia del generativismo, en el variacionismo se reconoce que la interlengua es un sistema dinámico condicionado por el entorno y, por lo tanto, por factores sociales. En las interfaces, la variación de la L2 se ve afectada tanto por factores lingüísticos como extralingüísticos y la variación cambia a medida que progresa la adquisición (Gudmestad, 2014:80). Al igual que otros autores que estudian las interfaces, Gudmestad (2014:87) dice que la variación en el español L2 apoya los resultados con respecto a la tardía adquisición de las estructuras lingüísticas que están gobernadas por las interfaces; es decir, entre módulos (morfología, sintaxis, semántica, discurso, pragmática).

Una de las diferencias entre el estudio variacionista de una L1 y una L2 es el tipo de variabilidad que se estudia (Gudmestad, 2014:81), la cual es de dos tipos:

1. Variabilidad tipo I: variación vertical > variación entre formas nativas y formas no-nativas (*target-like and non-target like forms*). Por ejemplo, en el francés, un estudio del tipo I analiza la alternancia entre el uso no-nativo de *a le* y la contracción nativa *au*.

2. Variabilidad tipo II: variación horizontal > variación entre diferentes formas nativas de la misma lengua. Por ejemplo, el uso o ausencia entre nativos del francés de la negación pre-verbal *ne*.

Gudmestad (2014:81) dice que, para el español, los estudios se han centrado en el tipo II y señala, entre otros, el trabajo de Geeslin, Linford, Fafulas, Long y Díaz-Campos (2013), el cual analiza la adquisición de la forma del sujeto en español y considera la importancia del individuo como factor extralingüístico, al comparar los datos del individuo con los del grupo.

Los participantes del estudio de Geeslin, Linford, Fafulas, Long y Díaz-Campos (2013) fueron 180 estudiantes de español como L2 (inglés L1), desde el primer año de estudios hasta el nivel graduado, y 27 nativos de español de diversos países de habla hispana, todos estudiantes graduados. El instrumento utilizado fue una tarea escrita contextualizada dentro de un diálogo entre HNs de español hablando de la vida universitaria. Las dos respuestas posibles se diferenciaban solo en cuanto a que una contenía un pronombre de sujeto nulo y la otra un pronombre de sujeto explícito. Se manipularon las siguientes variables lingüísticas, y se codificaron para cada instancia de sujeto: cambio de referente, TAM, continuidad de TAM, persona de la forma del verbo. La variable extralingüística fue la variable “participante”, la cual se utilizó para examinar hasta qué grado las diferencias individuales influyen el poder de predicción del modelo estadístico. La variable dependiente en este estudio fue la forma seleccionada y las dos categorías son el sujeto nulo y el sujeto explícito.

Se encontró variación individual del siguiente tipo: de los seis estudiantes de español L2, por lo menos uno seleccionó el sujeto nulo en todos los contextos. Entre los

nativos, también hubo uno que seleccionó solo sujetos nulos en la tarea escrita del estudio. En contraste, hubo estudiantes que nunca seleccionaron el sujeto nulo en los grupos menos avanzados, lo cual puede ser un indicio de que el estudiante no ha adquirido el PSN en su gramática. Esto no ocurrió en ningún caso con los participantes nativos. Los resultados entre los nativo-hablantes y los estudiantes graduados muestran la menor desviación estándar, lo cual indica que hay evidencia de que el “participante” afecta los resultados de manera decreciente a medida que aumenta la competencia lingüística. Esto se demuestra, por ejemplo, en que a medida que aumenta la competencia lingüística, disminuye el número de participantes que selecciona únicamente una sola forma de sujeto (en los primeros niveles hay una tendencia a seleccionar mayormente sujetos nulos, y en algunos casos de manera exclusiva, como mencionado anteriormente).

Los autores concluyen que el factor extralingüístico “participante” ayuda a mejorar el poder de predicción del modelo (p.168) y, de hecho, el no incluir este factor oscurece los resultados de otras variables, como es el caso del efecto TAM en el grupo del tercer semestre. Aunque en mi investigación no se lleva a cabo un análisis cuantitativo teniendo en cuenta el “participante” como variable independiente, sí se conduce un análisis cualitativo en el cual se demuestra que, en algunas ocasiones, todas (o muchas) de las instancias de una determinada forma, provienen de un solo participante. Es importante tener esto en cuenta al momento de trazar conclusiones con respecto a los resultados cuantitativos.

En el estudio de las expresiones de sujeto, un trabajo reciente que destaca por su carácter multidisciplinario entre la lingüística y la estadística es el de Gudmestad, House y Geeslin (2013). El estudio utiliza un modelo probabilístico multinomial (el modelo

bayesiano) en el uso de expresiones de sujeto de tercera persona singular, y tiene en cuenta todas las expresiones referenciales de sujeto (frases nominales léxicas, sujetos nulos, pronombres personales y otros pronombres). El modelo bayesiano permite examinar la relación entre una expresión referencial (variable dependiente) y múltiples variables independientes al mismo tiempo. Estas variables son: la L1 de los participantes (inglés/español), la perseverancia (forma del referente en la mención anterior), el cambio de referente (continuidad/cambio), el número del verbo (singular/plural), la especificidad del referente (identificable por nombre/no identificable/no aplicable), el TAM, la negación verbal (ausente/presente), el pronombre de objeto (ausente/presente), la cohesión del referente (distancia y tipo/función del referente en la mención previa), el sexo del participante y participante (categoría individual para cada participante en cuanto a sus efectos aleatorios).

El modelo estadístico también incluyó la interacción entre el cambio de referente y la cohesión del referente. De manera colectiva, a esta interacción y a las variables independientes se les denominó predictores. Adicionalmente, el modelo incluyó la interacción entre la L1 y los predictores, de tal manera que se pudiera hacer una comparación entre la expresión de sujeto de nativos y no nativos.

Los participantes fueron 16 estudiantes graduados avanzados y nativos del inglés y 16 nativos de español, también estudiantes graduados viviendo en los Estados Unidos, pero originarios de diversos países de habla hispana. Se recogieron un total de 6,342 instancias de sujeto de tercera persona a través de entrevistas sociolingüísticas.

La complejidad del modelo demanda una fuerte colaboración con un estadístico, aunque el poder interpretar los resultados no requiere de un conocimiento profundo de las

fórmulas utilizadas. Al incluir las anteriores variables e interacciones de manera simultánea, el modelo permite identificar las variables que predicen el uso de una expresión referencial sobre otra. Por ejemplo, con la excepción de la negación verbal y el sexo, todas las variables tienen algún efecto en la selección del sujeto nulo sobre otra opción: el cambio del referente explica la variabilidad sobre la preferencia por la variable “otros pronombres” y frases nominales léxicas, pero no por pronombres personales, por encima de sujetos nulos. La L1 solo es importante cuando interactúa con otras variables, lo cual quiere decir que la relación entre la expresión de sujeto y el número, la especificidad, el TAM y el pronombre de objeto difiere entre nativos y no nativos (p.387).

Este tipo de modelo estadístico beneficia el estudio de la ASL, siendo especialmente útil en el campo de las interfaces, ya que tiene la capacidad de revelar el efecto que pueden llegar a tener ciertas variables entre sí, para el uso de una expresión referencial sobre otra. El hecho de que este estudio haya demostrado que el efecto de la L1 solo es visible con la interacción de ciertas variables lingüísticas, evidencia detalles que de otra manera habrían permanecido en la oscuridad.

Por último, voy a resaltar brevemente el trabajo sociolingüístico de Tarone (1983) ya que proporciona información acerca del efecto de la modalidad, relevante para esta investigación. El estudio de Tarone (1983) no se enfoca en la adquisición de expresiones referenciales sino en la variabilidad de la interlengua, pero tiene en cuenta datos orales y escritos y dice que la variabilidad que se observa en los aprendientes de una lengua como L2 puede verse afectada por la tarea utilizada para extraer los datos. Tarone estudia la competencia lingüística desde el punto de vista del conocimiento lingüístico y no

solamente de la producción, esto es de crucial importancia porque indica que la capacidad lingüística de un aprendiente incluye tanto la producción como la percepción, la escritura, la lectura y también la capacidad de juzgar la gramaticalidad de las formas, argumentos a favor del uso de datos escritos en las investigaciones lingüísticas. El continuum de la interlengua sobre el que se basa esta autora es el propuesto por Labov (1969, citado en Tarone, 1983:152), donde en un extremo se encuentra el lenguaje no-cuidado (estilo vernáculo) y en el otro extremo el estilo cuidadoso, aquel donde el hablante presta más atención a la forma del lenguaje y en el cual podemos incluir la producción del lenguaje escrito. Lo más importante a resaltar del trabajo de Tarone son sus conclusiones con respecto a la permeabilidad de la interlengua por parte de la L2: según esta autora, el estilo más cuidado muestra más permeabilidad que el estilo no-cuidado, en otras palabras, se acerca más al habla nativa que el lenguaje oral, esto debido a la mayor atención a la forma por parte de los aprendientes. En la presente investigación veremos que se llega a la misma conclusión con respecto al acercamiento al habla nativa por parte de los HNNs en la modalidad escrita.

En conclusión, vemos que los estudios con un enfoque variacionista incluyen una gama más amplia que el generativismo de las expresiones referenciales de sujeto. Son estudios predominantemente cuantitativos, cuyos objetivos y metodología requieren un análisis estadístico capaz de analizar los posibles efectos de varios factores tanto lingüísticos como extralingüísticos. En el presente estudio, a través de las herramientas de la encuesta (ver Capítulo 4), se recopilieron datos extralingüísticos acerca de los participantes como lo son el sexo, la edad y el lugar de nacimiento. Estos factores no son el enfoque del estudio, sin embargo, el tener a disposición dicha información permite

ampliar al análisis en un trabajo futuro donde se pueda determinar si, por ejemplo, las variedades del español de los participantes afectaron la selección de las formas referenciales.

Como mencionado anteriormente, este tipo de análisis ha contribuido a investigaciones multidisciplinarias de colaboración entre la lingüística y la estadística, así como a la conexión entre la sociolingüística y los estudios de ASL. El poder hacer comparaciones con estudios variacionistas con la gran variedad de factores y de variables que se incluyen en ellos, es una dificultad que presenta este tipo de metodología; sin embargo, es importante tener en cuenta estas investigaciones en este estudio por dos razones principales: en primer lugar, resaltan el efecto que puede tener el individuo sobre los resultados del grupo. Este efecto puede ser considerable y, a pesar de no ser una variable independiente en mi estudio, sí se tiene en cuenta en el análisis cualitativo. En segundo lugar, aunque el modelo bayesiano utilizado por Gudmestad et ál. (2013) no es el único modelo de regresión disponible, la inclusión de un análisis estadístico similar, en cooperación con estadísticos, permite llevar a cabo observaciones en cuanto al comportamiento de las variables del presente estudio que, de otra manera, no serían evidentes o donde no podría asegurarse su nivel de significancia.

3.4 La adquisición de sujetos en el funcionalismo

El funcionalismo gana impulso en los años 80 como una forma de reacción en contra de los estudios de los años 60 y 70, que se enfocaban en el análisis del error y en el orden de la adquisición de los morfemas (Zyzik, 2014:30). Los estudios bajo esta perspectiva que investigan las expresiones referenciales frecuentemente tienen en cuenta

los factores pragmáticos de tópico y foco, los cuales hacen parte del análisis del presente trabajo.

Las investigaciones funcionalistas de ASL buscan generalizaciones empíricas acerca de la estructura de la lengua nativa de los adultos, y una de las mayores preguntas es si la interlengua obedece estas generalizaciones (Ritchie y Bhatia, 1996:13). Durante los años 80, Andersen propone una de las generalizaciones más notables para explicar algunos principios operativos de la interlengua: el principio 1:1 (*One to One Principle*). Este principio está basado en estudios que condujo Andersen en la adquisición del español como L2 y puede resumirse de la siguiente manera: el estudiante traza mapas entre forma y significado de manera consistente, aun cuando diverja de la norma (Zyzik, 2014:31). Zyzik da el siguiente ejemplo (p. 33): si un estudiante le adjudica el significado semántico de “objetivo” a la preposición *para*, es posible que subsuma otras funciones como “propósito” y/o “beneficiario” dentro de este significado, y que construya este mapa en su interlengua.

Las dos metodologías predominantes que se utilizan en el funcionalismo son las siguientes:

1. Enfoque función-a-forma: se empieza con un concepto, por ejemplo, tiempo o número y se analiza qué forma lingüística utilizan los estudiantes para expresarla. Las preguntas de investigación se centran en cómo ciertas formas se van desarrollando con el tiempo.
2. Enfoque forma-a-función: se empieza con una forma y se analiza para cuáles conceptos/funciones se utiliza. Este tipo de análisis permite ver posibles patrones de uso en lo que previamente parecía una variación aleatoria. Un

ejemplo es el estudio en español de las preposiciones *por* y *para*, que son problemáticas para el estudiante de español como L2: dada una preposición, se examina para cuáles funciones es utilizada, y cómo se va desarrollando con mayor competencia lingüística (Zyzik, 2014:33).

Una de las fortalezas de la perspectiva funcionalista es el hecho de que describe la adquisición de lengua incluyendo los estados iniciales y estudia todos los niveles de adquisición, además, incluye los diferentes niveles del lenguaje: el léxico, la morfología, la sintaxis, el discurso. Un problema metodológico que tiene este enfoque es el de cómo obtener una determinada forma en un estudio. Por ejemplo, si estamos investigando los clíticos, cómo instigamos a que un estudiante los produzca, en vez de producir una FN (Zyzik, 2014:36).

En el campo de la adquisición de sujetos, Zyzik (2014:37) dice que han sido asombrosamente pocos los estudios con un enfoque funcionalista y que los existentes tienen una metodología de función-a-forma; es decir, se comienzan con la noción de “sujeto” y se examinan cuáles son las formas (FNs, pronombres explícitos, sujetos nulos) que el estudiante de L2 utiliza para expresar esta noción en las diferentes etapas de la adquisición. Resalta que de lo más importante de estos estudios es el análisis de los datos con una perspectiva discursivo-pragmática, entre los cuales destaca el trabajo de Lubbers Quesada y Blackwell (2009), que describiré más adelante en esta sección.

Lubbers Quesada (2015:168) afirma que los trabajos con una perspectiva funcionalista se han enfocado mayormente en la capacidad de las diferentes formas de sujeto de crear cohesión en el discurso. Halliday y Hasan (1976:4) explican que la cohesión ocurre cuando la interpretación de algún elemento en el discurso es dependiente

de otro; en otras palabras, es una noción discursiva que se forma a través de marcadores lingüísticos explícitos (p. ej.: expresiones anafóricas) que indican la relación que existe entre oraciones. Lubbers Quesada (2015:168) señala que, en el funcionalismo, la cohesión textual se ha estudiado teniendo en cuenta el énfasis, el contraste, la desambiguación verbal (p. ej. entre las formas de primera y tercera persona del imperfecto: *yo hablaba/ella hablaba*), la accesibilidad de tópico (p. ej., los referentes correspondientes al uso de los sujetos nulos son más accesibles que las frases nominales léxicas (ver Sección 2.4), y el *grounding* (p. ej. los sujetos explícitos tienden a usarse más en el primer plano y a marcar acciones o un cambio en la estructura de la narrativa). En la L2, la cohesión es importante para construir un discurso ya que cuando tenemos una falta de cohesión, por ejemplo, en una situación de adquisición parcial de la lengua, puede haber una ruptura en la comunicación (Lubbers Quesada, 2015:36). En los resultados de este estudio veremos que esta es una situación que suele ocurrir cuando un hablante utiliza un sujeto nulo que no tenga un antecedente accesible¹¹ en la memoria del lector debido, en ocasiones, a la distancia entre el elemento anafórico y el referente, a problemas de ambigüedad verbal o cuando hay más de un referente en una escena, entre otras razones.

Saunders (1999), en su disertación doctoral, dice que la cohesión en el discurso se forma a través de marcadores lingüísticos explícitos que señalan la relación entre oraciones o proposiciones, mientras que la coherencia se refiere a si un texto es lógico e interpretable (1999:63). Según Saunders (1999), las expresiones anafóricas son elementos de cohesión y aquello a lo que se refieren son los nodos (conceptos introducidos por

¹¹ Ver sección 2.4 para la noción de *accesibilidad*.

primera vez por Halliday y Hasan, 1976). Por ejemplo, en una frase como *Arturo y su esposa salen a pasear y él coquetea con una desconocida*, Arturo es el nodo y el pronombre *él* es el elemento cohesivo. Saunders afirma que un texto puede contener elementos cohesivos y sin embargo resultar incoherente, por ejemplo, cuando no hay correferencia entre un elemento anafórico y el tópico de la oración:

(8) Abraham Lincoln lived in Illinois. This Midwestern State_i rejected the Equal Rights Amendment to the U.S. Constitution_j. It_i is the basis for the laws of our land.

(Saunders, 1999:63)

El enlace que crea cohesión entre la primera y la segunda oración en (8) es la FN *This Midwestern State* y su nodo es *Illinois*. El enlace entre la tercera y segunda oración es el elemento cohesivo *it* y el nodo es la FN *the U.S. Constitution*. Este texto resulta incoherente porque, a pesar de haber elementos cohesivos, estos no hacen referencia al mismo tópico. Un ejemplo de falta de coherencia en español es cuando hay un sobreuso de elementos cohesivos como pronombres explícitos, en vez de usar pronombres nulos cuando hay continuación de tópico. Sin embargo, en el presente trabajo vamos a argüir que el uso de sujetos explícitos cuando un sujeto nulo es posible o esperado a veces favorece la cohesión cuando, cognitivamente, el sujeto nulo resulta menos accesible, resultando en un mayor costo de procesamiento.

La cohesión es universal ya que todas las lenguas tienen herramientas para crearla, pero las lenguas se diferencian según una jerarquía anafórica, propuesta por Saunders (1999), la cual depende del nivel de prominencia del referente. Saunders encontró que la

competencia discursiva se ve afectada por el uso apropiado de los sujetos nulos en el español, por parte de estudiantes con inglés como L1. Esta jerarquía es como sigue:

indefinite NPs >> definite NPs >> proper nouns >> overt pronouns >> null pronouns

(Saunders, 1999:51)

El estudio de Saunders incluyó 60 estudiantes de español como L2 en cinco niveles de adquisición, desde principiantes hasta avanzados, y ocho nativos de español como grupo de control. Se analizaron 135 narrativas orales basadas en dos historias ilustradas. Estas historias incluyen un personaje principal (un cliente cuya compra de mercancía es confundida en una tienda), un personaje secundario (otro cliente) y otro terciario (la vendedora, que intenta resolver el problema del primer cliente). Saunders predice que se usarán diferentes formas de sujeto según la función o prominencia de cada personaje y cómo es introducido, mantenido y reintroducido en la narrativa. La autora dice que se considera que un personaje interviene en la historia solo si es introducido en posición de sujeto y no en posición de objeto indirecto; por ejemplo, si aparece un segundo personaje en la narrativa, y este se introduce en posición de objeto indirecto, el hablante continuará usando un sujeto nulo para referirse al personaje principal (p. 72). Saunders encuentra diferencias según el personaje se encuentre en posición de sujeto o en otra posición sintáctica. A continuación, me concentro en los resultados de la posición de sujeto.

Los resultados revelan que los hablantes nativos siguen la jerarquía anafórica para crear cohesión en el discurso. Generalmente, los personajes son introducidos en la narrativa con un elemento alto en la jerarquía (hacia la izquierda), es decir, una frase nominal; mientras que subsecuentes menciones del mismo personaje se indican con un

elemento bajo en la jerarquía, por ejemplo, sujetos nulos, a menos de que el uso del sujeto nulo cause ambigüedad. Además, la exacta manifestación de la jerarquía depende del papel del personaje. Por ejemplo, al introducir al personaje principal, los hablantes utilizan el nombre propio el 53.3% de las veces, como en ejemplo (9), o un nombre propio como parte de una frase indefinida el 46.7% de las veces, como en ejemplo (10):

(9) Te voy a platicar lo que le pasó a **José**.

(10) Algo muy gracioso e interesante le pasó ayer a **mi amigo Arturo**.

(Saunders, 1999:81)

En ningún momento los hablantes nativos utilizan una frase nominal definida para introducir al personaje principal, ya que las frases nominales definidas son siempre anafóricas, según Saunders (1999:83). Es de notar que a los participantes se les pidió que narraran lo que le había sucedido a “su amigo Arturo”, introduciendo de esta manera a Arturo en el discurso, por lo cual no se presentaron casos de frases nominales indefinidas ya que se asumía previo conocimiento de Arturo, así como tampoco se utilizaron pronombres o sujetos nulos, ya que el uso exofórico¹² de los pronombres es común en la narrativa escrita pero no en la oral (Saunders, 1999:82-83).

Por parte de los nativos (80%), el personaje secundario se ve introducido casi exclusivamente por frases nominales indefinidas, y el resto por demostrativos (6.7%) y por nombre propio (13.3%). Los deícticos no hacen parte de la jerarquía anafórica; sin embargo, Saunders dice que su uso es equivalente a las frases nominales indefinidas, ya que en ambos casos se supone falta de conocimiento previo por parte del oyente

¹² El uso exofórico se refiere a un tipo de referencia que, según Halliday y Hasan (1976:18), lleva al lector/hablante fuera del texto y no se considera cohesiva puesto que no une dos elementos en el texto. En este caso, el referente se encuentra en las instrucciones dadas a los participantes.

(1999:97). Esta diferencia de las formas utilizadas para la introducción entre el personaje principal y el secundario demuestra que existe un nivel de familiaridad menos pronunciado con este último, o que el personaje principal suele ser el tópico del enunciado.

La vendedora, la cual es el personaje terciario, se introduce de manera muy diferente en cuanto a que no se considera un personaje dentro de la narrativa, sino parte del escenario. Por lo tanto, vemos un uso frecuente de frases nominales definidas (p. ej.: *la vendedora, la señorita en la caja*), o incluso FNs que hacen una inferencia a la vendedora, pero no está explícitamente mencionada (p. ej.: *la tienda*). Estos son usos apropiados de las frases nominales definidas cuando una entidad forma parte de un esquema compartido (ver Sección 2.4 para una descripción de los esquemas).

Las formas de sujeto son diferentes según el papel del personaje, como acabamos de ver, pero también según el personaje ya haya sido introducido en la narrativa. Los HNs mantienen al personaje principal en la narrativa a través del uso del sujeto nulo casi en la totalidad de los casos (92.6%), y es reintroducido en la narrativa principalmente a través de una FN (68.2%), pero en algunas ocasiones también a través de un pronombre explícito (31.8%), como en (11):

(11) ...y entonces eh...*pro*_i llevó sus cosas. Pero *la señorita en la caja registradora*_j pues solamente le_i dio la bolsa. Bueno él...*él*_i se fue a la casa...

(Saunders, 1999:93)

Saunders no encontró muchos casos de mantenimiento y de reintroducción del personaje secundario. En posición de sujeto, el segundo cliente se mantiene a través de una frase nominal definida, un pronombre explícito o un pronombre nulo, mientras que las reintroducciones ocurren exclusivamente con frases nominales definidas. Al haber

pocas instancias, Saunders no calculó porcentajes y los resultados son únicamente ilustrativos y no concluyentes. Sin embargo, Saunders hace algunas observaciones (1999:99-100): los HNs tienen una mayor tendencia a usar el sujeto nulo en el mantenimiento del personaje principal que en el secundario; todas las instancias de mantenimiento del personaje secundario siguen la jerarquía anafórica (se usa un elemento más bajo en la jerarquía que en la introducción inicial del personaje); las reintroducciones del personaje secundario ocurren exclusivamente con una frase nominal definida.

En cuanto a la vendedora, solo en tres narrativas nativas (de dos hablantes), se menciona a este personaje más de una vez. Las tres instancias de mantenimiento ocurren en posición de no-sujeto y con un pronombre explícito (*su, le*) y, al igual que con el personaje secundario, su reintroducción ocurre exclusivamente con una FN definida (*la señorita encargada de la tintorería*). En conclusión, este personaje terciario generalmente suele introducirse, mantenerse y reintroducirse en posición de objeto indirecto u objeto preposicional, en vez de en posición de sujeto. En posición de sujeto solo ocurre en un 16.7% de los casos, lo cual contribuye a la cohesión de la narrativa ya que le permite al oyente enfocarse en los otros dos personajes principales (Saunders, 1999:106).

Los resultados de los estudiantes de español L2 indican variación entre los diferentes niveles de adquisición, y no es sino hasta los niveles más avanzados que las narrativas de los estudiantes se acercan a aquellas de los hablantes nativos. Los datos revelan que, en general, los estudiantes principiantes e intermedios no hacen una distinción entre el sujeto nulo y el sujeto explícito, y tampoco distinguen de manera clara entre los diferentes tipos de sujetos explícitos (pronombres, nombres propios, frases nominales definidas e indefinidas). Saunders considera que solo a partir del nivel

intermedio-alto, se considera que los estudiantes han adquirido el sujeto nulo (la adquisición es definida por un 80% de uso correcto del sujeto nulo en los contextos obligatorios). Por ejemplo, en este estudio el mantenimiento del personaje principal es uno de los contextos obligatorios para el sujeto nulo y mientras en este contexto los nativos utilizan esta forma el 92.6% de las veces vs 32% de los estudiantes principiantes, es solo al llegar al nivel avanzado que los estudiantes llegan al 81.4%.

En la introducción del personaje terciario, al igual que los nativos, los estudiantes utilizan correctamente una frase nominal definida, por ejemplo, *la señora que les vendió esto; la vendedora; la señorita de la caja* (Saunders, 1999:102), sin embargo, a nivel principiante no se puede concluir si se ha adquirido la distinción entre sujetos definidos e indefinidos, ya que el 83% de los participantes no mencionan a este personaje en sus narrativas. Solo los estudiantes avanzados se acercan al uso nativo.

En la reintroducción de los personajes, los hablantes nativos utilizan tanto nombres propios (68.2%) como pronombres explícitos (31.8%), mientras que los estudiantes usan nombres propios casi exclusivamente en todos los niveles de adquisición, específicamente los estudiantes avanzados el 85.7% de las veces.

En conclusión, el estudio de Saunders analiza las referencias nulas y explícitas según la función que sirvan en la narrativa, y se demuestra que el análisis puramente sintáctico no es suficiente ya que los referentes deben analizarse en el discurso. Establece que la selección de la forma de los nativos depende de los siguientes factores (p.172):

1. El papel, el nivel de familiaridad y la relativa importancia del personaje en el discurso. Por ejemplo, el personaje principal generalmente se introduce por

nombre debido a su alto nivel de familiaridad, mientras que al personaje secundario suele introducirse como *otro hombre*.

2. La función del personaje según su introducción, mantenimiento o reintroducción. Por ejemplo, para el mantenimiento, se puede usar un sujeto más hacia la derecha de la jerarquía (sujetos más mínimos) con respecto a la primera introducción.
3. La posición sintáctica del referente en la oración, es decir, según este se encuentre en posición de sujeto o en posición de objeto indirecto u objeto preposicional.

El extenso estudio de Saunders es relevante a varios niveles. En primer lugar, demuestra que el parámetro [+/-pro-drop] de la GU no es suficiente para el análisis de la adquisición de los sujetos en la ASL. Este estudio pone de manifiesto el hecho de que aun cuando los estudiantes usan sujetos nulos, no necesariamente esto es evidencia de su adquisición ya que los estudiantes también tienen que demostrar que entienden las diferencias entre las diversas funciones de los referentes, así como las diferencias entre las distintas opciones lingüísticas que se tienen para los referentes explícitos. Por lo tanto, así como los estudios de las interfaces, vemos que los factores discursivos son más difíciles de adquirir y deben considerarse en las investigaciones enfocadas en ASL. Saunders, como en la mayoría de los casos con otros autores, utiliza únicamente narrativas orales, por lo cual no es posible llegar a conclusiones con respecto a una posible influencia de la modalidad sobre las elecciones de los hablantes en relación con las funciones de los referentes en el discurso. Es posible que la modalidad escrita, así como lo demuestra Tarone (1983), al igual que el presente estudio, muestre un acercamiento al uso nativo en el uso de las expresiones referenciales.

Como mencionado anteriormente, los estudios funcionalistas, así como aquellos que investigan las interfaces, y los estudios cognitivistas que veremos en la siguiente sección, también tienen en cuenta factores discursivo-pragmáticos. Uno de estos estudios es el de Lubbers Quesada y Blackwell (2009), en donde se proponen reglas pragmáticas muy específicas que rigen el uso del pronombre nulo y explícito en el español, tanto en nativo-hablantes como en estudiantes de español como L2, en narrativas personales orales. Las autoras se centran específicamente en la alternancia entre el *yo* y el *zero* en la primera persona singular.

Este es un estudio de función-a-forma, donde se comienza con la noción de “sujeto” y se examinan las formas usadas por los estudiantes en las diversas etapas de adquisición. Las autoras comienzan identificando cinco reglas pragmáticas que describen las tendencias de uso entre los nativos de español. En sus resultados, las autoras encontraron que el sujeto nulo se ve favorecido en el contexto de (1) la prominencia del referente, la cual puede darse a través del conocimiento mutuo, de factores sintácticos o de accesibilidad, y (2) en el contexto de epistémicos parentéticos, donde se usa un sujeto nulo con el propósito de evaluar un enunciado anterior o posterior o para mitigar la fuerza de un enunciado, o para hacer una evaluación acerca de un enunciado (p. ej. *pro* no sé, *pro* digo). Las otras tres reglas pragmáticas predicen el uso de un sujeto explícito: (3) cambio de foco de atención entre un referente y otro. En este caso se puede usar un sujeto nulo siempre que haya otros factores que permitan la desambiguación, como por ejemplo por el contexto, el conocimiento mutuo o la morfología del verbo; (4) foco contrastivo, donde se usa un sujeto explícito al introducir información no dada/nueva y cuando esta información se encuentra en oposición con otro referente en el discurso, y (5) peso

pragmático, regla según la cual un nativo de español usa el pronombre de sujeto *yo* de manera explícita, para que el enunciado sea más relevante (tenga más peso) a nivel personal. Su uso ocurre con actos de habla con verbos de sustentación (*claiming*), creencia, opinión, emoción o conocimiento (p. ej. *yo creo*) (Lubbers Quesada y Blackwell, 2009:119-122).

Una vez identificadas estas reglas pragmáticas en las narrativas de los HNs de español del estudio, las autoras pasan a analizar si los aprendientes de español L2 siguen las mismas reglas, y cuáles se adquieren más temprano/más tarde en las etapas de adquisición. La base de datos incluyó 130 narrativas orales producidas por 30 estudiantes de cinco niveles diferentes y por 20 nativo-hablantes de México como grupo de control. Es de notar que, en general, la frecuencia de uso de formas referenciales es muy parecida en todos los grupos, con una preferencia por los sujetos nulos. Incluso a niveles principiantes, los aprendientes utilizan el sujeto nulo un 67% de las veces, comparado con el 65% de los nativos. Las autoras atribuyen esto al hecho de que el parámetro del sujeto nulo (PSN) puede reiniciarse¹³; sin embargo, en los niveles principiantes no se encontró ningún uso del pronombre explícito con peso pragmático, lo cual puede atribuirse a que no han incorporado este uso específico en su sistema. Al mismo tiempo, los resultados muestran que los estudiantes avanzados usan el peso pragmático aun sin haber tenido instrucción explícita, por ejemplo: “en el aspecto que *yo* quiero de un hombre...que *yo* quiero de mi vida” (2009:128, mis itálicas).

El hecho de que los hablantes nativos seleccionan un sujeto nulo en el 93% de los contextos con un referente prominente y seleccionan un sujeto explícito en un 100% de

¹³ Ver sección 2.1 para una descripción de las propiedades del PSN.

los casos de foco contrastivo y de peso pragmático, indica que estas dos formas (sujeto nulo y sujeto explícito) no se encuentran en variación libre, sino que dependen del contexto.

Otro resultado de este estudio que cabe mencionar es el hecho de que todos los estudiantes, incluso los avanzados, siguen usando el pronombre pleno para referentes ya sobresalientes en el discurso; es decir, en el foco de atención. Esto puede atribuirse a que la morfología no es lo suficientemente prominente para los aprendices de manera que puedan usar un pronombre nulo (los nativos utilizan pronombres nulos el 93% de las veces en este contexto, vs 80% del grupo avanzado).

No es el objetivo de los estudios funcionalistas el de mostrar la adquisición de las propiedades sintácticas que rigen el uso del sujeto nulo, por lo que la importancia de este trabajo recae en mostrar la distribución entre el sujeto nulo y explícito según los diferentes niveles de competencia lingüística de los aprendientes de español como L2 y cómo las elecciones de los hablantes pueden verse restringidas por las reglas pragmáticas identificadas por las autoras. Por lo tanto, a pesar de centrarse en el uso del pronombre de primera persona, los resultados informan la actual investigación con respecto a los factores pragmáticos según el nivel de adquisición de los HNNs de español.

En conclusión, el funcionalismo es un enfoque basado en el uso, en donde la función comunicativa de un enunciado es lo que impulsa la adquisición de una lengua (Zyzik, 2017:34). Vemos que los estudios funcionalistas tienen en cuenta factores discursivo-pragmáticos como el tópico y el foco, centrándose en la capacidad de las diferentes formas referenciales de crear cohesión en una narrativa. El trabajo de Saunders (1999) estudia las formas referenciales utilizadas por los hablantes según su función en la

narrativa, y propone una jerarquía anafórica en la cual la forma de sujeto está condicionada por el papel del personaje en la narrativa, así como su importancia, el nivel de familiaridad y la función del referente. Mientras que el estudio de Lubbers Quesada y Blackwell (2009) analiza la alternancia entre el sujeto nulo y el sujeto explícito en los pronombres de primera persona en varios niveles durante el proceso de adquisición de lengua, bajo la perspectiva pragmático-discursiva del análisis del discurso.

Estos trabajos informan la presente investigación ya que demuestran, de nuevo, el importante papel que juegan los factores pragmáticos en el análisis del discurso. Además, presentan evidencia acerca del efecto del nivel de prominencia de un referente en la narrativa en la distribución de los sujetos nulos y explícitos.

Las investigaciones realizadas bajo esta perspectiva, al igual que la mayoría de los estudios variacionistas, dependen fundamentalmente de datos provenientes de la producción oral (Zyzik, 2017:35), así que en el presente trabajo, como hemos mencionado anteriormente, buscamos llenar este hueco al presentar datos provenientes de narrativas escritas que puedan compararse con datos orales y así poder determinar si los factores que se han encontrado en los estudios funcionalistas tienen el mismo peso según la modalidad discursiva escrita utilizada para la producción lingüística.

3.5 La adquisición de sujetos en el cognitivismo

Como vimos en la introducción de esta disertación, aunque existen diferencias entre el funcionalismo y el cognitivismo, estas perspectivas lingüísticas no son claramente, o necesariamente, separables en todo momento y, por lo tanto, en este trabajo el enfoque se considera cognitivo-funcionalista. En primer lugar, ambas perspectivas no se presentan como teorías en sí sino como enfoques, y los estudios que se llevan a cabo

tienen en cuenta tanto los factores cognitivos que veremos a continuación, así como aquellos pragmático-discursivos, aún sin desestimar propiedades sintácticas, como lo son aquellas relacionadas con la prominencia del referente según su posición en la oración (pre o postverbal). Estos enfoques se alejan de la visión modularista del generativismo al aseverar que en la ASL participan los mismos sistemas cognitivos involucrados en otros procesos de aprendizaje, entre ellos la percepción, la memorización, el procesamiento de la información, etc. (Dussias, 2003:233). Sin embargo, Cadierno y Pedersen (2014:33-34) dicen que, aunque el proceso de aprender una L2 es similar al de la L1, la L2 es un proceso más complejo porque “el aprendiente tiene que aprender a utilizar las estrategias de estructuración conceptual apropiadas y específicas de la L2”. Estos autores añaden que la influencia de la L1 sobre la L2 puede entrar en competición durante el desarrollo de la interlengua, ya que la L1 se encuentra fuertemente asentada.

Finalmente, al igual que el funcionalismo, el cognitvismo se aleja del concepto del innatismo de propiedades estructurales específicas de la lengua y se basa mayormente en el lenguaje en uso. Según Tomasello, en los modelos del lenguaje basados en el uso, “all things flow from the actual usage events in which people communicate linguistically with one another” (2006:439); esto significa que el conocimiento acerca del lenguaje surge de su propio uso y no de una gramática universal innata. Tomasello añade que el uso repetitivo de ciertas estructuras es lo que permite el proceso de adquisición de las expresiones lingüísticas.

Las investigaciones enfocadas en la ASL desde la perspectiva de la lingüística cognitiva (LC), han examinado temas como la inseparabilidad de la gramática y el léxico en el aprendizaje de una L2; la medida en la que las construcciones hacen parte de las

representaciones mentales de los estudiantes; el papel que juega la frecuencia en la adquisición de la L2; el papel del nivel básico de la categorización en la adquisición de vocabulario, y el papel de la L1 en el desarrollo de los patrones de la L2 (Cadierno e Hijazo-Gascón, 2013:103). Específicamente en el estudio de las expresiones de sujeto, los análisis se basan en gran medida en el concepto de la accesibilidad del referente y su prominencia, así como en los estados de atención, memoria y nivel de conocimiento mutuo entre los interlocutores con respecto al referente. Estos son los factores principales que informan la presente investigación.

En el Capítulo 2 de este trabajo se revisaron las jerarquías más prominentes que se han propuesto en la L1 para el análisis de la selección y el uso de sujetos referenciales, tanto anafóricos como no anafóricos, entre ellas la Givenness Hierarchy (GH) de Gundel, Hedberg y Zacharski (1993). A continuación, paso a describir el trabajo de Blackwell y Lubbers Quesada (2012), en el cual se propone una jerarquía adaptada al español, en donde las autoras prueban el poder predictivo de la GH en estudiantes de español como L2. Aparte de la contribución que estas autoras hacen al estudio de la ASL del español desde este enfoque, la importancia del trabajo de Blackwell y Lubbers Quesada (2012) reside en que se usarán los datos orales recogidos para ese estudio y vendrán comparados con los datos obtenidos de narrativas escritas para la presente disertación. Es por este motivo que, a continuación, el trabajo de Blackwell y Lubbers Quesada (2012) será descrito en mayor detalle.

Las autoras analizaron 40 narrativas orales, previamente recogidas, entre tres grupos de estudiantes de español como L2 (inglés como L1): principiantes, intermedios y avanzados. También se analizaron 10 narrativas de un grupo de nativos. Los participantes

habían visto los primeros cinco minutos de la película muda de Charlie Chaplin, “A Woman” (1915) y luego narraron lo que vieron a un entrevistador hablante nativo que supuestamente no había visto la película, aunque este hecho no se hizo explícito de manera consistente con todos los participantes por parte del entrevistador (M. Lubbers Quesada, comunicación personal, 28 de marzo 2019).

En este estudio, las autoras proponen una distribución diferente para el español, en donde identifican tres niveles en el estado cognitivo (EC) activado, dependiendo del grado de recuperabilidad del referente. Esto representa un avance con respecto a la GH original de Gundel et ál. (1993) ya que se reconoce que, aunque el referente está activado en la memoria a corto plazo, se necesitará de una forma más o menos elaborada para evitar ambigüedad y seleccionar el antecedente intencionado. Estos niveles son:

- Activado y recuperable (ACTR): donde el referente es recuperable a través del uso de sujetos nulos debido al almacenamiento del referente en la memoria a corto plazo, al ser el tópico de la secuencia inmediatamente anterior. En el ejemplo (12) vemos que los primeros dos sujetos nulos se refieren a Charlie, el tópico que está en foco en el momento de usar la expresión referencial. Al intervenir *el esposo*, este pasa a estar en foco y el siguiente sujeto nulo se refiere al nuevo personaje; sin embargo, Charlie es recuperable con un sujeto nulo en el siguiente enunciado:

(12) ...y *pro_j* llega y todo y *pro_j* está probando a ver en donde *pro* es más profundo. Mientras tanto el esposo_i se da cuenta mientras *pro_i* está pensando que *pro_j* era la muchacha,...

(Blackwell y Lubbers Quesada, 2012:149)

- Activado (ACT): el referente está en foco en la secuencia anterior, pero no es recuperable con un sujeto nulo, por lo tanto, se requiere, o se espera, el uso de un pronombre explícito. En el ejemplo (13) el sujeto nulo seleccionaría a Charlie (el cual es el antecedente del primer sujeto nulo en este enunciado); por lo tanto, se requiere del pronombre explícito *él* para seleccionar al *señor de los ojos vendados*, el cual se encuentra activado en la memoria, pero no es el foco de atención:

(13) Entonces *pro*i jala al señor de los ojos vendadosj, y cuando *él**i/j se da cuenta de que...

(Blackwell y Lubbers Quesada, 2012:149)

- Activado y no recuperable (ACTNR): en este caso el pronombre no es suficiente para identificar el referente y se necesita de una frase nominal general (p. ej.: una de las dos, las dos mujeres, uno de ellos). El ejemplo (14) muestra que, aunque el referente *otros dos señores* está activado en la memoria a corto plazo, un sujeto nulo o un pronombre incorrectamente recuperarían a Chaplin como antecedente:

(14) Aparecen otros dos señores y se sientan con Chaplin en la banca, pero Chaplin empieza a tomar de su bebida mientras ellos están distraídos. *Uno de ellos* se da cuenta y empieza otra vez otro pleito.

(Blackwell y Lubbers Quesada, 2012:150)

En la jerarquía propuesta por Blackwell y Lubbers Quesada (ver Figura 3.5.1, más abajo), se unen los estados referencial (REF) y de tipo identificable (TYP), ya que en español corresponden con las mismas formas lingüísticas, como también lo muestran Gundel et ál. (1993:284):

in focus > {pro/él}	activated > {él/éste/ése/aquel/este N} activated + recoverable {pro} activated + non-recoverable {el/los/uno de N/pronoun}	familiar > {ese/aquel}	uniquely identifiable > {el N}	referential / type identifiable {∅ N/un(a) N}
---------------------------	---	---------------------------	--------------------------------------	---

Figura 3.5.1 - Givenness Hierarchy para el español (Lubbers Quesada, 2015:196)

En la Tabla 3.5.1 vemos descripciones más amplias de cada estado, junto con las formas que la jerarquía predice para cada cual.

Tabla 3.5.1: Codificación y descripción de los estados cognitivos

Estado cognitivo y descripción	Codificación	Ejemplos
1. En foco: - El referente es el centro de atención - La entidad más sobresaliente de la declaración previa - Foco contrastivo	INF	Nulo / él ∅
2A. Activado + Recuperable con sujeto nulo: - El referente está en la memoria a corto plazo ya que es el tópico inmediatamente precedente en la secuencia de eventos. - Es recuperable con un sujeto nulo ya que puede ser inferido	ACTR	Nulo ∅
2B. Activado: - El referente está en la memoria operativa pero no puede ser recuperado con un sujeto nulo.	ACT	pronombre explícito él/ella Demostrativos este/ese/aquel
2C. Activado + No recuperable: - El referente está en la memoria a corto plazo pero NO puede ser recuperado con un pronombre explícito	ACTNR	GenNP (General NP) el/los/uno de N/pronombre Expresiones generales no pronominales. Ej.: los dos, las dos mujeres, el otro, uno de los hombres, uno de ellos
3. Familiar: - El referente puede ser identificado porque el interlocutor tiene una representación en la memoria operativa debido a una mención en el discurso previo.	FAM	DefNP / PropN el N/ese/aquel Ej.: la muchacha, Charlie

<p>4. Identificable de modo único: - Las entidades están en la memoria del interlocutor, y el referente puede ser identificado basado en (1) la sola expresión, o (2) por asociación a través de inferencia con otra entidad activada en el discurso reciente.</p>	UNI	DefNP el N Ej.: la mujer, los actores, Charlie Chaplin
<p>5. Referencial/De tipo identificable: - Se introduce la entidad por primera vez - Se asume que el interlocutor entiende a qué se refiere la descripción de la frase pronominal (NP).</p>	TYP	IndefNP un(a) N/∅N Ej.: una muchacha Puede ser un DefNP para referirse a uno o todos los miembros de un grupo, ej.: la policía

En Blackwell y Lubbers Quesada (2012) se formularon cinco hipótesis con respecto a cada uno de los ECs. Estas hipótesis reflejan la jerarquía revisada propuesta por las autoras. Los resultados muestran que tanto los nativos como los estudiantes hicieron la selección restringida por los cinco ECs del español, como pronosticado por la GH, pero lo hicieron de manera diferente y usando los referentes con diferente frecuencia; además, los nativos se adhirieron con mayor fuerza a la jerarquía. Los resultados generales muestran que los patrones más claros se encuentran en las selecciones de los referentes a los extremos de la jerarquía: sujetos nulos en foco y frases nominales indefinidas para el estado cognitivo de tipo identificable. Las mayores diferencias se ven con los estados en el medio: ACTR/ACT/ACTNR, FAM y UNI. A continuación, describo los resultados principales, por estado cognitivo.

En el EC en foco (INF), el referente es el centro de atención y corresponde con la entidad más prominente en el discurso. Se predice el uso del sujeto nulo y los resultados confirman esta hipótesis: los HNs seleccionaron los sujetos nulos el 91% de las veces, mientras que los aprendientes entre el 61% y el 65% (2012:154). En el siguiente ejemplo

de un HN, el sujeto *Charles* es llevado al foco de atención al ser mencionado, sus subsecuentes menciones anafóricas utilizan un sujeto nulo:

- (15) ...pues, Charles llegó, *pro* vio a la muchacha que estaba sola; *pro* se sienta platicar con ella pero... bueno, pues casi, casi *pro* tropezó con ella.

(Blackwell y Lubbers Quesada, 2012:148)

Como mencionado anteriormente, Blackwell y Lubbers Quesada (2012) proponen tres ECs activados, según la recuperabilidad y prominencia del referente. Estos tres estados son mutuamente excluyentes. Las autoras proponen el estado activado-recuperable (ACTR) para aquellos casos en los cuales el referente no es el más prominente (no es el foco de atención); aun así, es recuperable con un sujeto nulo gracias a factores gramaticales (p. ej.: concordancia verbal de género), el conocimiento mutuo y las condiciones semánticas del verbo. Los resultados revelan que los nativos tienen una tendencia a utilizar el sujeto nulo en un 82% de los casos, mientras que los estudiantes avanzados lo hacen en un 59% de los casos. Los estudiantes intermedios prefieren un pronombre explícito (41%) y los principiantes un nombre propio (52%). En el ejemplo (16), el estudiante avanzado utiliza el pronombre nulo para identificar a Charlie, aunque en la oración anterior el tópico sean Charlie y el hombre (*ellos*), pero las propiedades gramaticales de número del verbo sirven para identificar correctamente al referente, por lo tanto, las autoras lo codifican como ACTR:

- (16) y Charlie_i usa el un el bastón para jugar más con el hombre_j...ellos_{ij} van a caminar...*pro*_i usa el bastón para como engañar, burlar el hombre

(Blackwell y Lubbers Quesada, 2012:156)

El EC ACT se propone para aquellas entidades activadas en la memoria a corto plazo (debido a su mención inmediatamente anterior), pero en donde el contexto del

discurso no permite un sujeto nulo para su recuperabilidad. Las autoras plantean como hipótesis el uso del pronombre explícito ya que un sujeto nulo podría causar ambigüedad. Los datos revelan unos resultados interesantes, puesto que los hablantes nativos seleccionan el pronombre explícito solamente en un 41% de los casos, seguido por frases nominales definidas (26%) y nombres propios (13%). También, seleccionaron sujetos nulos un 12% de las veces, pero en este caso, las autoras arguyen que esta selección conlleva a un problema de comunicación, al no poderse identificar el referente a través del contexto. Para el EC ACT, todos los grupos de no nativos prefirieron FNs definidas, seguidas por el pronombre explícito. Además, aunque el resultado no es estadísticamente significativo, se ve que los estudiantes menos avanzados tienden a utilizar una menor variedad de formas referenciales. En el ejemplo (17), un sujeto nulo seleccionaría erróneamente a Charles, el tópico de la oración precedente, por lo tanto, se requiere de un pronombre explícito (en negrilla) adecuadamente seleccionado por un participante de nivel principiante para referirse a otro personaje, el cual ha sido introducido en el enunciado inmediatamente anterior, en posición de objeto preposicional con el pronombre *él*:

(17) cuando *él*_i no vio, Charlie Chaplin_j bebió la soda de *él*_i, y *él*_i vio

(Blackwell y Lubbers Quesada, 2012:157)

En el EC ACTNR, las autoras dicen que los hablantes utilizan formas referenciales semánticamente generales en contextos donde el referente está activado en la memoria a corto plazo, pero no es recuperable con el solo uso de un pronombre explícito definido. Ejemplos de estas expresiones son *los dos*, *las dos mujeres*, *uno de ellos*, etc. A estas expresiones las autoras las codifican como GenNP. Sin embargo, esta codificación es

debatible, puesto que las expresiones pueden fácilmente calificarse como FNs definidas o indefinidas. Presentaré este argumento en el Capítulo 4, Sección 4.5. Los resultados del estudio demuestran que la hipótesis se cumple para los HNs en un 44% de los casos, mientras que para el grupo de HNNs, únicamente se cumple para el grupo intermedio (67%). Los estudiantes avanzados solo seleccionan esta forma en un 10% de los casos, prefiriendo el nombre propio en un 59% de las instancias. Las autoras notan que, entre los estudiantes, hay una menor variedad de expresiones generales en comparación con los nativos, lo cual puede deberse a una mayor complejidad de las expresiones referenciales disponibles de este tipo en el repertorio de los HNNs y/o a que los nativos muestran un mayor número de sujetos en sus narrativas (2102:158). En el ejemplo (18), el HN selecciona la FN general *el otro*, la cual es semánticamente de baja informatividad:

(18) *el otro* queda inconsciente

(Blackwell y Lubbers Quesada, 2012:159)

En el EC familiar (FAM), las autoras predicen el uso de frases nominales definidas o de nombres propios para identificar entidades que no se encuentran en la memoria a corto plazo, pero que han sido introducidas previamente en la narrativa, lo cual incluye menciones en las instrucciones dadas a los participantes durante la encuesta. Los resultados confirman esta hipótesis en el 90% de los casos, para todos los grupos combinados, encontrándose las FNs definidas en un primer lugar. Solo los nativos utilizaron sujetos nulos (4%), lo cual conlleva a malentendidos y ambigüedad, como en el siguiente ejemplo en donde solo al haber visto la película, puede identificarse a Chaplin como al antecedente del pronombre nulo:

- (19) El otro_i se le pone al brinco pero cuando ve_i que Chaplin_j se le va a enfrentar, sale_i corriendo. Mientras el hombre_k está tirado en el piso, *pro*_k queda inconsciente, y *pro*_j camina, *pro*_j pasa sobre él_k

(Blackwell y Lubbers Quesada, 2012:160)

El ejemplo (20) muestra el uso de una FN definida, *la muchacha*, para seleccionar a un referente que se encuentra en la memoria a largo plazo, al haber sido mencionado con bastante anterioridad:

- (20) ...pasa una muchacha_i que le coquetea al esposo_j, este señor_j se va atrás de la muchacha_i ... (varias escenas después)... pero **la muchacha**_i le pide que *pro*_j se tape los ojos.

(Blackwell y Lubbers Quesada, 2012:150)

En el estado identificable de modo único (UNI), las autoras proponen como hipótesis el uso de frases nominales definidas en donde el referente se encuentra en la memoria del oyente y puede identificarse, bien por medio de la expresión referencial en sí (ej.: Charlie Chaplin, el sol, el perro del vecino, la película), bien porque la entidad es inferible (por ejemplo en el caso de que pueda hacerse una asociación con otra entidad anteriormente mencionada en el discurso, como puede ser el caso del uso de *padre*, si se ha anteriormente mencionado a *una familia*) (para una discusión acerca de los *inferibles*, ver sección 2.3) o bien a través de marcos cognitivos (ver sección 2.4). La hipótesis se cumple al seleccionar expresiones definidas (FNs o nombres propios) casi en el 100% de los casos. El ejemplo (21) muestra un caso de inferencia a través de un marco cognitivo, ya que tanto el narrador (no nativo) así como el entrevistador pueden identificar la expresión *el título* como perteneciente al video que se está narrando, ya que se puede inferir que los videos/películas tienen títulos:

(21) Entrevistador: ¿Qué pasó en el video?

Participante: Ok. Pues, *el título* es “Una Mujer”...

(Blackwell y Lubbers Quesada, 2012:160)

Por último, en el estado de tipo identificable (TYP), las autoras predicen el uso de FNs indefinidas, ya que las entidades se introducen en la narrativa por primera vez. La hipótesis se cumple para todos los grupos, aunque no en un 100% de los casos. Las diferencias entre los varios grupos no son estadísticamente significativas. En algunos casos, los grupos utilizan frases nominales definidas, lo cual sucede 7% de las veces en los nativos y entre 11% y 23% de las veces entre los estudiantes. La diferencia principal en el uso de FNs definidas entre nativos y no nativos es que las expresiones utilizadas por los nativos son de carácter general (*la mujer ideal*), ver ejemplo (22), mientras que los no nativos introducen personajes específicos (*la mujer, el policía, la coqueta*), como ocurre en el ejemplo (23):

(22) y bueno, aparece *la mujer ideal*; para eso, aparece una chica; con una actitud

(23) y luego *la coqueta*; camina enfrente de él; y ella; coqueta con el hombre;.

(Blackwell y Lubbers Quesada, 2012:161)

En conclusión, este estudio revela que el uso de las expresiones de sujeto de tercera persona está restringido por el EC de la entidad en el discurso. Esta tendencia aplica tanto a los HNs como a los HNNs, aunque los datos revelan que, a mayor competencia lingüística, hay una mayor sensibilidad al EC del referente; es decir, en la mayoría de los casos hay un acercamiento al uso de las formas nativas a medida que avanza la competencia lingüística. Aunque las formas referenciales no son siempre predecibles, los datos muestran que hay una mayor adherencia a la GH en los extremos (estados INF y

TYP), mientras que hay una mayor variedad de formas en los estados intermedios, sobre todo en los estados activados.

Aunque la GH fue propuesta en el 1993 y ha servido de base para subsecuentes estudios, Gudmestad et ál. dicen que, hasta la fecha de publicación de su estudio, no se han llevado a cabo trabajos aplicando la GH a investigaciones de ASL:

...the only L2 study we know of which focuses exclusively on third-person contexts, as the current study does, Blackwell and Quesada (2012) applied Gundel, Hedberg, and Zacharski's (1993) givenness hierarchy to subject expression in Spanish, revealing that use of subject-expression forms is constrained by such pragmatic factors (Gudmestad et ál., 2013:375).

Han pasado varios años desde que Gudmestad et ál. hicieran esta aseveración, pero, a mi saber, y después de haber llevado a cabo una búsqueda exhaustiva, el estudio de Blackwell y Lubbers Quesada (2012) descrito aquí sigue siendo el único que considera específicamente la jerarquía propuesta por Gundel et ál. (1993) en un estudio dedicado a la ASL de las expresiones referenciales de sujeto, específicamente para el español.

Los estudios que hemos analizado hasta el momento, de diferentes maneras han contribuido a un mayor entendimiento de los procesos que llevan a cabo los hablantes al seleccionar expresiones referenciales de sujeto y en los procesos de adquisición de una segunda lengua por parte de HNNs. Hemos visto que en estos procesos influyen factores sintácticos, sociales, pragmáticos y cognitivos. El análisis llevado a cabo por Blackwell y Lubbers Quesada (2012) toma en cuenta varios de estos factores y propone hipótesis que explican las diferencias entre los HNs y los HNNs de español. Los factores más relevantes en su estudio son la prominencia sintáctica codificada por la PAH de

Carminati (2002), los factores pragmáticos de tópico y foco y los factores cognitivos de memoria y atención. Al integrar estos factores, las autoras ofrecen un análisis discursivo que permite determinar cómo los hablantes hacen una evaluación de la disponibilidad cognitiva (accesibilidad, topicalidad y *givenness*) del referente con respecto al oyente.

Partiendo de lo anterior como marco de referencia para el presente trabajo, también hemos visto cuáles son algunos de los huecos que presentan las investigaciones hechas hasta el momento, incluyendo los trabajos de Gundel et ál. (1993) y de Blackwell y Lubbers Quesada (2012). Principalmente, los vacíos son dos: en un primer lugar, estos estudios no se hacen la pregunta si los hablantes tienen diferentes cargas cognitivas según la modalidad que se utilice para la producción del lenguaje; en otras palabras, no se tienen en cuenta las posibles diferencias de procesamiento entre el lenguaje escrito y el lenguaje oral. Esto puede deberse al hecho de que el lenguaje escrito históricamente se ha considerado secundario al lenguaje oral. Sin embargo, para que un HNN pueda considerarse competente en una segunda lengua, debe poder demostrar conocimientos y adquisición lingüísticos en todas las habilidades que comprende una lengua, esto incluye la capacidad de expresarse oralmente, por escrito, en sus capacidades de lectura, y en su capacidad de juzgar la gramaticalidad de las formas.

En segundo lugar, aunque tanto Gundel et ál. como Blackwell y Lubbers Quesada discuten el TC como factor que influye en las elecciones de los hablantes, en ninguno de los dos estudios se consideró hasta qué punto este factor tiene un efecto directo en la selección de las formas referenciales. Por lo tanto, dentro del marco que provee la propuesta para el español de Blackwell y Lubbers Quesada (ByLQ) (2012), esta investigación tiene como objetivo llenar estos huecos para de esta manera contribuir con

los estudios de ASL y proporcionar un avance a la jerarquía planteada por Gundel et ál. (1993).

En los capítulos que siguen voy a describir la estructura general de esta disertación, la metodología llevada a cabo, incluyendo una explicación de cómo la metodología se compara y difiere de ByLQ (2012). Presentaré los resultados obtenidos, varios de los cuales sirven de evidencia a la aplicabilidad de la GH en español en el análisis de las formas de referenciales de sujeto. Las cinco hipótesis planteadas por Blackwell y Lubbers Quesada (2012) servirán de base inicial para describir cómo se comparan los resultados de los datos orales con los datos escritos de este trabajo.

CAPÍTULO 4

4 METODOLOGÍA

4.1 Introducción

A través de diferentes enfoques y metodologías, los estudios analizados en los Capítulos 2 y 3 han aportado un mayor entendimiento acerca de los procesos y factores involucrados en la selección, uso e interpretación de las expresiones referenciales de sujeto, tanto por parte de HNs como de HNNs durante la adquisición de una segunda lengua. Las diferentes perspectivas lingüísticas que hemos revisado (el generativismo, el variacionismo, el funcionalismo y el cognitivismo) tienen en cuenta propiedades sintácticas, pragmáticas y cognitivas, que hacen parte del análisis de la presente investigación. En particular, el análisis llevado a cabo por ByLQ (2012), basado en la propuesta de la GH de Gundel, Hedberg y Zacharski (1993), considera varios de estos factores y propone hipótesis que explican las diferencias entre los HNs y los HNNs de español. Los factores más relevantes son la prominencia sintáctica codificada por la PAH de Carminati (2002), los factores pragmáticos de tópico y foco y los factores cognitivos de memoria y atención, al mismo tiempo que se tiene en cuenta el papel que juega el TC en las elecciones de los hablantes.

En la primera sección de este capítulo se presentan las preguntas de investigación y las hipótesis planteadas para este estudio. En las subsecuentes secciones se describen los participantes del estudio, los instrumentos utilizados para la recolección de datos, así como los procesos de codificación y la metodología para el análisis de los mismos. En

esta última sección se describen los procesos para el análisis cuantitativo a través de modelos de regresión estadísticos y los procesos para el análisis cualitativo que se deriva de los resultados obtenidos.

4.2 Preguntas de investigación e hipótesis

Este estudio tiene un doble objetivo: en primer lugar, se busca determinar si la GH revisada de ByLQ (2012) es igualmente válida para predecir la adquisición de las formas de sujeto en español L2 en textos escritos, donde un escritor no tiene las mismas posibilidades de determinar los niveles de atención y memoria en el lector con respecto al referente. Además, queremos saber si los HNs de español y los HNNs de español de nivel avanzado siguen los mismos patrones que utilizan para el lenguaje oral.

En segundo lugar, se busca aislar la variable “terreno común ” (TC) para determinar hasta qué punto afecta la elección del referente, principalmente en el momento de introducción de un referente en la narrativa. Dicha variable no fue considerada por Gundel et ál. (1993) ni por ByLQ (2012), aunque en ambos trabajos se menciona que el conocimiento compartido es un factor determinante en la selección de las formas lingüísticas. Se busca comparar la posible influencia de este factor entre los HNs y los estudiantes de español como L2 en las narrativas escritas. Por lo tanto, el propósito es probar las predicciones de la GH para las formas referenciales de sujeto en el discurso narrativo escrito, buscando una manera de aumentar o disminuir el nivel de atención y memoria al establecer o no un TC entre escritor y lector.

El estudio busca contestar las siguientes dos preguntas de investigación, con sus respectivas sub-preguntas:

1. ¿Cómo se comparan las narrativas derivadas de dos medios de producción distintos: producción oral y producción escrita?
 - 1.1. ¿Hay diferencias entre los HNs de español y los HNNs de español de nivel avanzado, en sus composiciones escritas?
 - 1.2. ¿Cómo se comparan los resultados del presente estudio sobre el lenguaje escrito con los resultados de ByLQ (2012) del lenguaje oral entre ambos grupos de participantes?
2. ¿Es posible manipular el TC para aumentar o disminuir el nivel de atención y memoria de los participantes en las narraciones escritas?
 - 2.1. ¿Cuáles estados cognitivos (Ecs) y correspondientes expresiones referenciales se verán afectados al establecer diferentes niveles de TC entre los participantes?
 - 2.2. ¿Cómo se diferencian las narrativas escritas entre nativos de español y estudiantes avanzados de español como L2, dependiendo del factor TC?

A continuación, se plantean las siguientes hipótesis para contestar a priori las preguntas de investigación:

- Hipótesis 1: Debido a que la GH es universal, tanto en las narrativas orales como en las escritas, los participantes seguirán la GH, con mayor adherencia a los estados extremos de INF y TYP; es decir, los participantes seleccionarán un sujeto nulo para un referente con EC INF y una FN indefinida para un sujeto con EC TYP.
- Hipótesis 2: Se verá un acercamiento mayor al HN por parte de los HNNs, ya que en el discurso escrito los hablantes tienen más oportunidad de pensar en sus

respuestas y corregirlas; más específicamente, los HNNs seleccionarán las formas de sujeto de una manera más semejante a las selecciones de los HNs en las narrativas escritas que en las narrativas orales de ByLQ (2012).

- Hipótesis 3: Al ser el discurso escrito más cuidado que el discurso oral, y al no poder verificar el nivel de atención y memoria, los participantes, tanto HNs como HNNs, usarán expresiones referenciales más elaboradas en las narrativas escritas que en las orales.
- Hipótesis 4: Es posible aislar el TC como factor en la selección de expresiones referenciales y las suposiciones de los participantes con respecto al nivel de atención y memoria se verán afectados. Más específicamente, al introducir los personajes en el discurso que supone previa familiaridad con el referente, el EC de TYP corresponderá mayormente con expresiones definidas en las narrativas con un TC establecido que en las narrativas sin un TC determinado.
- Hipótesis 5: El factor TC desempeñará un papel más importante en las narrativas escritas de los HNs que de los HNNs. Se observará una codificación más notable del TC por parte de los HNs a través de un mayor uso de expresiones definidas que los HNNs en las narrativas con TC, y en un mayor uso de cláusulas con información de fondo en narrativas sin TC (p. ej. “El/un hombre con bastón”), según el personaje esté siendo introducido o reintroducido en la narrativa.

4.3 Participantes

En esta investigación participaron un total de 40 personas: 20 HNs de español y 20 HNNs de español de nivel muy avanzado. Ninguno de los participantes recibió una

compensación por su participación, la cual se llevó a cabo de manera totalmente voluntaria y anónima. Además del español, algunos de los HNs, así como algunos de los HNN, hablan otras lenguas con diversos niveles de competencia: catalán, italiano, francés, portugués, coreano, alemán y gallego.

Los HNs fueron doce mujeres y ocho hombres, mayores de 25 años. El 50% proviene de España y el otro 50% de países latinoamericanos: Colombia (8), Ecuador (1) y Venezuela (1) (ver la Tabla 4.3.1, abajo).

Tabla 4.3.1 - Hablantes nativos de español (L1)

L1						Total
Género	F 12	M 8				20
Edad	18-24 0	25-34 5	35-44 9	45-55 5	55+ 1	20
Lenguas además del español	Inglés 18	Catalán 2	Italiano 2	Portugués 2		
País de nacimiento	Colombia 8	Ecuador 1	España 10	Venezuela 1		20
País de residencia	Colombia 3	China 1	EEUU 12	España 3	Suiza 1	20

Los HNNs (ver la Tabla 4.3.2, más abajo) provienen en su totalidad de los Estados Unidos y todos residían en los Estados Unidos en el momento de participar en el estudio. Participaron doce mujeres y ocho hombres, mayores de 18 años. Todos los participantes no-nativos son estudiantes de posgrado de programas de español (maestrías y estudios doctorales relacionados con la lingüística) de diversas universidades de los Estados Unidos y/o profesores de español. Únicamente dos de los HNNs no estudiaron español en un país de habla hispana, pero todos han continuado con sus estudios de español por un mínimo de tres años después de la escuela secundaria. Para verificar el nivel de

competencia lingüística, los HNNs completaron una prueba adaptada del DELE del Instituto Cervantes¹⁴, la cual incluyó 20 preguntas (ver Apéndice A). Aunque todos los HNNs pertenecen a grupos que se consideran hablantes avanzados de español (profesores de español y estudiantes de maestría/doctorados en español), solamente se seleccionaron aquellos hablantes que completaron la prueba con un 70% de precisión con el objetivo de utilizar los participantes más avanzados entre los que respondieron a la encuesta¹⁵.

Tabla 4.3.2 - Hablantes no nativos de español (L2)

L2							Total
Género	F	M					20
	12	8					
Edad	18-24	25-34	35-44				20
	1	14	5				
Lenguas además del español y el inglés			Francés	Portugués	Alemán	Gallego	Coreano
			2	5	1	1	1
País de nacimiento	EEUU						20
	20						
País de residencia	EEUU						20
	20						
Números de años de estudio de español después de la secundaria							
	De 3 a 4 años	Más de 5 años					20
	2	18					
Años de estudio de español en un país de habla hispana							
	Nunca	Menos de 1 año	Entre 1 y 3 años	Más de 3 años			20
	2	5	9	4			
Ocupación actual	Estudiante graduado	Profesor de español					20
	10	10					

4.4 Instrumentos y recolección de datos

La recolección de datos se hizo a través de una encuesta¹⁶ en línea distribuida por Qualtrics que incluyó seis secciones para todos los participantes, con excepción de la prueba de competencia lingüística que se utilizó únicamente para los HNNs. Las secciones fueron las siguientes:

¹⁴ Diploma de español: nivel superior. Instituto Cervantes, 11 de mayo de 2007.

¹⁵ De los 29 participantes no nativos que participaron en la encuesta, se eliminaron nueve.

¹⁶ IBR ID#STUDY00006201

- Sección 1: introducción con explicación de la encuesta e inclusión de la carta de consentimiento.
- Sección 2: datos biográficos (edad, sexo, tiempo de estudio del español, otras lenguas habladas, nacionalidad, país de residencia y ocupación).
- Sección 3: prueba DELE para verificar el nivel avanzado del participante no-nativo y seleccionar los más avanzados entre ellos (ver Apéndice A).
- Sección 4: instrucciones que indican que el participante verá un videoclip de una película muda. En esta sección se establece el TC de la siguiente manera:
 - Con TC: “Una vez que lo hayas visto, cuenta por escrito los eventos de la película a un amigo que también la ha visto. Al final, dile a tu amigo lo que más te gustó de ella”.
 - Sin TC: “Una vez que lo hayas visto, cuenta por escrito los eventos de la película a un amigo que no la ha visto aún. Al final, dile a tu amigo lo que más te gustó de ella”.

El objetivo de pedirles a los participantes que incluyeran información acerca de lo que más les gustó se hizo para darles una razón para narrar una película, de esta manera se buscó eliminar parcialmente la artificialidad de la tarea, aunque se reconoce la imposibilidad de eliminarla por completo.

- Sección 5: se incluye un videoclip de la película muda *A Woman* de Charlie Chaplin (1915), el cual tiene una duración de 5’44”. Se utiliza este clip ya que es el mismo utilizado en el trabajo de ByLQ (2012) y, al tener múltiples personajes, ofrece la oportunidad de varias expresiones de sujeto referenciales de tercera persona.

- Sección 6: instrucciones que indican cómo narrar la película de forma escrita. Se les pide que vean el video solo una vez y se les indica que el panel de escritura se programó para cerrarse automáticamente después de 15 minutos, esto con el propósito de obtener datos de escritura espontánea, en la cual el participante tiene menos oportunidad de revisión que en la escritura planeada.
- Sección 7: los participantes pasan a escribir su narración. A los HNNs se les da un glosario de palabras útiles y se les indica que no busquen palabras adicionales en el diccionario (este es el mismo glosario utilizado en el estudio oral de ByLQ). Se prepararon cuatro encuestas con algunas diferencias en las instrucciones, para reflejar los siguientes grupos:
 - Grupo 1 [L1TC]: 10 x HNs con instrucciones que indican existencia de TC entre el escritor y el lector.
 - Grupo 2 [L1NTC]: 10 x HNs con instrucciones que indican falta de TC.
 - Grupo 3 [L2TC]: 10 x HNNs con instrucciones que indican existencia de TC.
 - Grupo 4 [L2NTC]: 10 x HNNs con instrucciones que indican falta de TC.

Para los dos grupos de hablantes (HNs y HNNs), el tiempo promedio para completar la encuesta fue de 25 minutos. Los HNNs recibieron la encuesta en inglés y los HNs en español.

4.5 Codificación y metodología en el análisis de datos

La codificación de las expresiones referenciales, así como el análisis cualitativo y cuantitativo se llevó a cabo en línea con el trabajo de ByLQ (2012), ya que de esta manera podemos comparar los resultados de los datos orales con los escritos. En un primer lugar, el análisis fue hecho por la investigadora y subsecuentemente revisado y corroborado por otra analista con experiencia con este tipo de investigación.

Se codificaron todas las expresiones de sujeto referencial de tercera persona plural y singular, tanto animados como no animados. Así como en el estudio del 2012, se eliminaron los siguientes sujetos referenciales: nulos obligatorios (p. ej.: \emptyset es muy extraño), pronombres indefinidos (p. ej.: no pasó *nada*) y sujetos dativos de verbos dativos psicológicos, aunque estos últimos activan un referente en la memoria y por lo tanto influyen la selección lingüística en el discurso siguiente (p. ej.: a Charlie le gusta *la mujer*) (ejemplos tomados de ByLQ, 2012:152).

Se realizó un análisis completo al incluir todas las formas referenciales de sujeto disponible para los hablantes en español: sujetos nulos (Null)¹⁷, pronombres personales explícitos (Overt), demostrativos (Dem), FNs definidas (DefNP), nombres propios (PrpN) y FNs indefinidas (IndNP). Después de clasificar las formas (anafóricas y no anafóricas), estas se codificaron como referente singular/plural y animado/no-animado. A continuación, se pasó a la codificación según su EC en el discurso (INF / ACTR / ACT / ACTNR / FAM / UNI / TYP). Para este proceso se tienen en cuenta varios factores, entre ellos la prominencia del referente según su posición sintáctica pre o posverbal; la función

¹⁷ Las abreviaciones para las formas referenciales y para los ECs se mantienen en inglés y siguiendo las mismas abreviaturas que en el trabajo de Blackwell y Lubbers Quesada (2012), para simplificar la comparación entre los dos estudios.

pragmática del referente según sea tópico o foco; la continuidad de tópico / cambio de referente, y la distancia entre la actual mención del referente y su mención inmediatamente previa (o entre el elemento anafórico y el antecedente) para determinar el nivel de memoria y atención con respecto a la entidad. A continuación, voy a detallar el proceso para determinar cada EC, el cual es el mismo seguido en el estudio oral¹⁸. Las formas que codifican los ECs proveen información procedimental acerca de la *manera* de accesibilidad cognitiva (Gundel, 2010:152); es decir, dónde y cómo se accede a una representación mental de la interpretación intencionada por el hablante/escritor. De este modo, el hablante/escritor guía al oyente/lector restringiendo las posibles interpretaciones hacia aquellas que tienen el EC explícitamente codificado por la forma seleccionada. Los ejemplos utilizados a continuación hacen parte de los datos escritos de la presente investigación:

- INF: el referente ha sido introducido en una posición sintáctica prominente en la oración inmediatamente anterior. Es el tópico de la oración, por lo tanto, se encuentra en el centro de atención.

(1) Luego que *el hombre* (DefNP/INF) se encuentra con la dama, *él* (Overt/INF) le ofrece una bebida [P1.L1¹⁹]

(2) El primer señor regresa y [\emptyset] (Null/INF) está muy molesto que [P35.L2]
- ACTR/ACT/ACTNR: el referente no es el tópico del actual enunciado, pero se encuentra en la memoria a corto plazo ya que ha sido mencionado en la oración/escena inmediatamente anterior. Su clasificación entre ACTR/ACT/ACTNR depende del nivel de recuperabilidad del referente.

¹⁸ Ver Gundel (2010:154) para las pautas de codificación según el protocolo de los estados cognitivos.

¹⁹ P = Participante #; L1 = HN / L2 = HNN

Por ejemplo, si el referente es el t3pico de la oraci3n anterior, este es recuperable con un sujeto nulo y por lo tanto se considera ACTR (3), de otra manera se necesitar3 de un sujeto expl3cito, o al haber m3s de un referente posible en el enunciado anterior que pueda causar ambigüedad (4 y 5):

(3) ...espera a que el adúltero le traiga la bebida. [\emptyset] (Null/ACTR) Se sienta con ella y [P10.L1]

(4) ...Charlie bebe un poco de la bebida de uno de los hombres. Él (*Overt/ACT*) se enoja y [P23.L2]

(5) Bueno, la pel3cula empieza con una familia...La mamá, el pap3, y a lo mejor su hija. *La mamá* (DefNP/ACTNR) est3 durmiendo y [P30.L2]

- FAM: el referente es identificable ya que una representaci3n de este se encuentra en la memoria del oyente/lector debido a una previa menci3n en la narrativa, aunque no necesariamente en el discurso inmediatamente anterior. En el ejemplo (6), *el otro hombre* fue mencionado varias escenas m3s atr3s, por lo cual ya existe una conceptualizaci3n de esta entidad en la memoria:

(6) pero justo entonces *el otro hombre* (DefNP/FAM) volvi3 con la bebida para la coqueta [P22.L2]

- UNI: existe una representaci3n 3nica del referente en la mente del oyente/lector, por lo que es identificable a trav3s de su sola menci3n; por ejemplo, el personaje de Chaplin (7^a) o las pel3culas (7b); la representaci3n tambi3n puede ser inferible por asociaci3n con otra entidad mencionada en el discurso reciente al haber sido activado a trav3s del proceso de enlace

(*bridging*); por ejemplo, al mencionar el referente *la familia*, es posible identificar una subsecuente mención de sus miembros (7):

(7) a. Pues le cuento que vi uno de los varios cortos que hizo *Chaplin* (PrpN/UNI) durante su carrera [P11.L1]

b. Pues [\emptyset] (NULL/UNI) empieza con el hombre principal con su mujer y su amiga, [P21.L2]

(8) ...está la familia en un banco de un parque, *el marido* (DefNP/UNI), *la mujer* (DefNP/UNI)... [P10.L1]

- TYP: el referente es introducido en el discurso por primera vez. A través de su mención puede asumirse que el oyente/lector puede acceder al tipo de objeto descrito por la expresión referencial utilizada por el hablante/escritor:

(9) a. Acabo de ver *una película* (IndNP/TYP) de Charles Chaplin de los años 20 o 30 en blanco y negro [P7.L1]

b. Pues el resumen es que hay *un hombre aburrido* (IndNP/TYP) que ve pasar a una chica [P9.L1]

Además de las formas referenciales y de los Ecs, las otras variables esenciales se denotaron como sigue:

- Grupo: atributos del hablante en cuanto a HNs (L1) o HNNs (L2)
- TC: atributos del hablante en cuanto a la existencia de TC (TC) o no TC (NTC)
- Estilo: en referencia a si la narrativa es oral (datos del estudio del 2012) o escrita (datos del presente trabajo).

Después de la eliminación de las formas de sujeto no analizables, se codificaron 1.106 registros. Se mantiene la forma GenNP con el propósito de comparar de manera consistente los datos escritos con los orales del 2012, sin embargo, esta categoría se

elimina cuando comparamos los datos escritos entre L1 y L2, con y sin TC. La tabla de frecuencia con GenNP es como sigue:

Tabla 4.5.1 – Frecuencia y distribución de las formas de sujeto (todos los grupos) con GenNP

	Null	Overt	GenN	Dem	DefNP	PrpN	IndNP	Total
Frecuencia	451	68	64	17	318	126	62	1.106
Porcentaje	40.78	6.15	5.79	1.54	28.75	11.39	5.61	100

A diferencia del estudio oral de ByLQ (2012), el cual tiene en cuenta siete formas referenciales, los datos escritos del presente estudio tienen en cuenta seis al eliminar GenNP (ver Tabla 4.5.2, más abajo). Se tomó esta decisión ya que, en mi opinión, las expresiones generales pueden fácilmente clasificarse como FNs definidas o indefinidas, según se pueda o no identificar al referente de manera específica. ByLQ consideraron que algunas expresiones referenciales son semánticamente demasiado generales como para poder identificar al referente (M. Lubbers Quesada, comunicación personal, 2018). Dentro de esta categoría se incluyen expresiones como *los dos*, *las dos mujeres*, *el otro*, *uno de los hombres*, *uno de ellos*, etc. De esta manera, las investigadoras crearon la categoría de GenNP para poder recoger estas expresiones. Sin embargo, en los datos obtenidos para este estudio, el ejemplo (10) más abajo es el único caso que se encontró de una expresión referencial que no permite la identificación del personaje. Aunque la expresión *la otra mujer* es definida, haciendo suponer un referente que es por lo menos familiar, en la escena hay dos mujeres y, por lo tanto, la sola expresión no permite la recuperación del referente, como tampoco lo hace el contexto:

- (10) El hombre y una de las mujeres están durmiendo y roncando en el banco de lo que parece un parque. *La otra mujer* está sentada en el banco pero no duerme [P14.L1]

Los ejemplos (11) y (12) que siguen a continuación son casos representativos de GenNP que han sido reclasificados como DefNP e IndNP, respectivamente:

- (11) ...*el otro hombre* aparece de nuevo... [P8.L1]

- (12) *Otra mujer* pasa por el banco y [P40.L2]

En (11), la expresión referencial definida *el otro hombre* aparece en la línea 16, pero más arriba, en la línea 9, el participante nos había introducido a este personaje como a Charlie Chaplin: “aparece otro hombre (Chaplin), también a conquistarla” (el paréntesis hace parte de la redacción de este participante), por lo cual *el otro hombre* es plenamente identificable y se clasifica como DefNP. Por otro lado, en (12) el referente *otra mujer* no es identificable puesto que el personaje es introducido en la narrativa en ese momento. De esta manera, la expresión puede clasificarse como IndNP.

En la base de datos se resaltaron 64 registros cuyas formas se reconocen como DefNP o IndNP en vez de GenNP, lo cual representa el 5.79% de los datos totales. De estos, 26 (40.62%) se reconocen como DefNP y 38 (59.37%) como IndNP.

Tabla 4.5.2 Frecuencia y distribución de las formas de sujeto (todos los grupos) sin GenNP

	Null	Overt	Dem	DefN	PrpN	IndNP	Total
Frecuencia	451	68	17	344	126	100	1.106
Porcentaje	40.78	6.15	1.54	31.10	11.39	9.04	100

El análisis estadístico se llevó a cabo con la colaboración de un equipo de estadísticos del Centro de Asesoría Estadística (*Statistical Consulting Center*) de la Universidad de Georgia (Athens, USA). Este equipo utilizó una metodología similar a la del estudio del 2012, con el objetivo de poder comparar de manera consistente los datos de ambos estudios. El propósito de trabajar con estadísticos a través de una colaboración interdisciplinaria es el de identificar y aplicar un modelo que permita examinar las formas de sujeto y su relación con otras variables lingüísticas: Estado Cognitivo, Grupo, Estilo y TC. Los objetivos de aplicar los distintos modelos estadísticos apropiados de este estudio son, a continuación:

- Demostrar el uso de las formas referenciales para cada EC utilizando tablas de frecuencia bidimensional/de doble entrada (*2-way frequency table*).
- Investigar la diferencia entre los grupos utilizando una prueba de chi-cuadrado (o una prueba exacta de Fisher) en las tablas de frecuencias.
- Aplicar una regresión logística con la penalización de Firth para aquellas categorías de formas que pueden ser potencialmente diferentes entre los dos grupos, con el objetivo de investigar la diferencia entre los grupos para cada forma referencial.

Las tablas de frecuencias bidimensionales presentan la distribución conjunta de dos variables categóricas. Las celdas de la tabla representan las instancias observadas de cada resultado posible. Ya que el análisis de este trabajo se enfoca en el uso de las expresiones referenciales de sujeto entre diferentes grupos/TC/estilos, consiguientemente las tablas de frecuencias se construyen con base en estas variables.

La prueba de chi-cuadrado se usa para probar el grado de dependencia entre dos variables categóricas. Se busca rechazar la hipótesis nula de que el uso de las formas de sujeto son las mismas entre los dos grupos. En esta prueba, entre menor sea el valor de p , mayor será la desviación de la hipótesis nula. Sin embargo, cuando las muestras son demasiado pequeñas, la prueba de chi-cuadrado resulta inadecuada, por lo tanto, y al igual que en el estudio del 2012, hay dos alternativas posibles. En primer lugar, se eliminan algunas celdas donde los valores son muy pequeños para que la prueba de chi-cuadrado pueda funcionar. Pero aún así, en algunos casos esto resulta ineficiente y se aplica la segunda alternativa, que es el uso de la prueba exacta de Fisher.

Si la prueba de chi-cuadrado arroja valores significativos, se concluye que hay diferencias en el uso de las formas referenciales de sujeto entre los grupos. En este caso, nos interesa investigar cuál es la forma que difiere entre los dos grupos. A raíz de que el resultado para cada forma es binario (se usa o no se usa), se aplica una regresión logística para determinar si los dos grupos tienen la misma probabilidad de seleccionar una forma sobre la otra. Como veremos en el análisis en el siguiente capítulo, en algunos casos los dos grupos demuestran una completa separación en el uso de las formas; por ejemplo, en el EC TYP, todos los usos de DefNP son hechos por hablantes L1 y ninguno por hablantes L2. Cuando esto ocurre, el modelo de regresión logística no funciona y, para corregir este problema, se utiliza un modelo de inferencia con una penalización de Firth (corrección utilizada para corregir cuando hay una separación completa o casi completa en los datos). Además, si el uso total de una forma tiene menos de cinco instancias, o si las formas tienen el mismo porcentaje de ocurrencia entre los dos grupos, la regresión logística es innecesaria.

En el Capítulo 5 se presentan los resultados cuantitativos obtenidos del estudio estadístico que se llevó a cabo. Estos resultados incluyen la comparación entre los HNs y HNNs en las composiciones escritas, la comparación entre los datos escritos con los datos orales del estudio de ByLQ (2012), y los resultados entre las narrativas escritas con y sin TC. En el Capítulo 6 se lleva a cabo la discusión de los resultados cuantitativos y el análisis cualitativo de los mismos.

CAPÍTULO 5

5 RESULTADOS CUANTITATIVOS

5.1 Introducción

En este capítulo se presentan los resultados cuantitativos obtenidos del análisis estadístico realizado. La primera sección se enfoca en una comparación entre los datos orales reportados en Blackwell y Lubbers Quesada (ByLQ) (2012) y los datos escritos recolectados para el presente estudio de participantes nativos de español (L1) y participantes no-nativos (L2). Es importante mencionar que del estudio de ByLQ solo se comparan los resultados de los grupos 3 y 4, ya que estos son los que corresponden con los niveles del presente estudio: el grupo 3 de su estudio son estudiantes de español de nivel avanzado que corresponden con los hablantes no nativos (HNNs) del presente estudio; mientras que el grupo 4 de aquel estudio son los hablantes nativos (HNs). Cuando señalamos los totales de ambos grupos del estudio de ByLQ, en todos los casos nos estamos únicamente refiriendo a los grupos 3 y 4.

La segunda sección se enfoca en los resultados obtenidos con respecto al factor Terreno Común (TC), y su potencial efecto en la selección de las expresiones referenciales de sujeto en las narrativas escritas. De esta manera, revisaremos si el TC puede ofrecer algunas explicaciones con respecto a la selección de las diferentes formas, según el estado cognitivo (EC). Cabe recordar que, para este análisis, solo se tienen en cuenta las narrativas escritas del presente estudio y se excluyen las orales del estudio de ByLQ. Esto se debe a que en su estudio no se aisló este factor como variable y por lo

tanto no se diseñaron instrucciones específicas con respecto al TC. Aunque en la mayoría de los casos el entrevistador indicó que no había visto la película, por lo tanto, asumiendo una falta de TC, esto no fue consistente con todos los participantes (M. Quesada, comunicación personal, 2018). Además, de las dos autoras que llevaron a cabo el estudio, una de ellas había visto la película y no la otra. Este hecho se tuvo en cuenta al analizar los resultados orales, ya que en algunas ocasiones el haber o no haber visto la película es lo que permite identificar correctamente un referente. En el caso de las narrativas escritas del presente estudio, la investigadora ha visto la película, y la diferencia en el TC radica en las instrucciones específicas dadas a los participantes (ver la metodología en el Capítulo 4 de este trabajo).

5.2 Resultados cuantitativos: estudio oral vs estudio escrito

En esta sección empezaremos mostrando los resultados generales de la distribución de las formas de sujeto (sujetos nulos — Null; pronombres personales explícitos — Overt; demostrativos — Dem; FNs generales no pronominales — GenNP; FNs definidas — DefNP; nombres propios — PrpN; FNs indefinidas — IndNP) según el EC, presentados en tablas de frecuencia, para luego pasar a analizar cada EC por separado en subsecciones, lo cual significa que tenemos siete subsecciones. A su vez, cada subsección incluye las siguientes comparaciones:

- Comparación entre L1 y L2 en las narrativas escritas (sin GenNP)
- Comparación entre L1 y L2 entre las narrativas orales y las escritas (con GenNP):
 - Comparación oral y escrita para hablantes L1

- Comparación oral y escrita para hablantes L2
- Comparación entre Grupo (L1/L2) y medio de Producción (Oral/Escrito)

Al comparar los resultados entre L1 y L2 aplicamos la prueba chi-cuadrado con un nivel de significación establecido igual o menor de 0.05 ($p \leq 0.05$) y la prueba exacta de Fisher con un nivel de significación igual o menor de 0.10 ($p \leq 0.10$) (ver Capítulo 4). En aquellos casos en los cuales los resultados arrojan valores con significación estadística o al límite de la significación, es decir, donde el valor de p no alcanza significación estadística, pero se nota una tendencia hacia la significación (p. ej. $p=0.0503$) (ambos casos resaltados en letra negrilla en este capítulo), podemos concluir que los hablantes L1 y L2 difieren en su uso de formas referenciales de sujeto. Para poder determinar cuáles son las formas en las cuales difieren los dos grupos de hablantes se pasa a aplicar un modelo de regresión logística, el cual se lleva a cabo únicamente sobre aquellas formas con instancias por encima de cinco. También se excluyen aquellas formas cuyos porcentajes de distribución son iguales o que difieren de un 1%, ya que un modelo de regresión no es necesario en estos casos. Las celdas resaltadas en gris en las tablas son aquellas que se eliminan.

La comparación entre los factores Grupo y Producción es una investigación general donde se realiza un ajuste logístico a través de la aplicación de un modelo de regresión logística múltiple. Los resultados revelan cómo Grupo y Producción afectan el uso de cada forma de sujeto. En este modelo, el valor de p refleja la significación, y el coeficiente mide el tamaño del efecto; es decir, indica hasta qué punto el factor bajo análisis afecta los resultados. En este análisis general solo se incluyen aquellos resultados que arrojaron significación estadística.

5.2.1 Resultados generales

En general, en el presente estudio se registró un número menor de instancias de expresiones referenciales de sujeto que en el estudio de ByLQ (ver Tabla 5.2.1, más abajo). Este resultado proviene de los HNs, ya que en el presente estudio produjeron un total de 517 expresiones de sujeto en sus narrativas escritas en comparación con 793 instancias en el estudio oral previo; por otro lado, los HNNs del presente estudio produjeron 589 expresiones de sujeto en comparación con 421 expresiones de HNNs en el estudio oral de ByLQ. También es relevante notar que el menor número de expresiones referenciales de sujeto por parte de los HNs en el presente estudio se registran en el EC en foco. En el estudio oral previo, los HNs produjeron 425 instancias en comparación con 213 en el presente estudio escrito. Se observa la tendencia opuesta en las expresiones de sujeto de los HNNs; es decir, los HNNs realizaron 185 sujetos en el EC en foco en el estudio oral previo en comparación con 252 en el presente estudio escrito. En las narrativas escritas, en comparación con las narrativas orales de ByLQ, los HNNs expresaron un número mayor de sujetos en todos los estados; mientras que los HNs produjeron un número mayor solo en los estados familiar (FAM) y de tipo identificable (TYP) y un número menor en los estados en foco (INF), activado recuperable (ACTR), activado (ACT) y identificable de modo único (UNI). Para el EC activado no recuperable (ACTNR), los HNs elaboraron el mismo número de sujetos (50 instancias) que ese grupo en el estudio previo.

En la Tabla 5.2.2, más abajo, se observa que la tendencia en el uso de las formas según EC en las narrativas escritas es similar a la que reportan ByLQ para las narrativas orales; esto es, tanto los HNs como los HNNs, seleccionan una forma referencial según

las predicciones de la Givenness Hierarchy; específicamente, los participantes optan por referirse a un referente en el estado en foco o activado y recuperable con un sujeto nulo y a un referente que es de tipo identificable con una FN indefinida. En los estados familiar e identificable de modo único, los participantes seleccionan FNs definidas. Las únicas excepciones son los estados cognitivos ACT y ACTNR: en las narrativas escritas los dos grupos (HNs y HNNs) seleccionan mayormente FNs definidas en el estado ACT, mientras que en las narrativas orales la forma más utilizada por parte de los HNs son los pronombres explícitos.

Tabla 5.2.1 - Comparación de instancias de expresiones referenciales para ECs entre datos orales (Blackwell y Lubbers Quesada) y datos escritos (presente estudio)

Estado Cognitivo	INF		ACTR		ACT		ACTNR	
Frecuencia Porcentaje								
Producción (oral/escrita)	*O	**E	O	E	O	E	O	E
L1	425 69.7%	213 45.8%	96 70.1%	44 39.6%	69 57.0%	53 48.2%	50 56.2%	50 48.1%
L2	185 30.3%	252 54.2%	41 29.9%	67 60.4%	52 43.0%	57 51.8%	39 43.8%	54 51.9%
TOTAL	610 100%	465 100%	137 100%	111 100%	121 100%	110 100%	89 100%	104 100%
Estado Cognitivo	FAM		UNI		TYP		TOTAL	
Frecuencia Porcentaje								
Producción	O	E	O	E	O	E	O	E
L1	82 55.8%	68 47.2%	29 63.0%	37 45.7%	42 65.6%	52 57.1%	793 65.3%	517 46.7%
L2	65 44.2%	76 52.8%	17 37.0%	44 54.3%	22 34.4%	39 42.9%	421 34.7%	589 53.3%
TOTAL	147 100%	144 100%	46 100%	81 100%	64 100%	91 100%	1214 100%	1106 100%

*O = datos orales de Blackwell y Lubbers Quesada (2012)

**E = datos escritos del presente estudio

En el estado ACTNR, la tendencia de los dos grupos también es hacia FNs definidas en las narrativas escritas, pero en las narrativas orales la tendencia es hacia expresiones generales (p. ej. *los dos, uno de ellos, etc.*) (HNs) y nombres propios (HNNs).

Tabla 5.2.2 - Frecuencia y distribución de formas de sujeto por estado cognitivo (los dos grupos) en narrativas escritas (con GenNP)

Frecuencia Porcentaje	INF	ACTR	ACT	ACTNR	FAM	UNI	TYP
Null	374 80.43	60 54.05	8 7.27	6 5.77	2 1.39	1 1.23	0 0.00
Overt	31 6.67	12 10.81	25 22.73	0 0.00	0 0.00	0 0.00	0 0.00
GenNP	1 0.22	4 3.60	1 0.91	21 20.19	10 6.94	1 1.23	26 28.57
Dem	2 0.43	2 1.80	6 5.45	1 0.96	2 1.39	3 3.70	1 1.10
DefNP	37 7.96	21 18.92	58 52.73	51 49.04	91 63.19	55 67.90	5 5.49
PrpN	20 4.30	12 10.81	12 10.91	23 22.12	38 26.39	21 25.93	0 0.00
IndNP	0 0.00	0 0.00	0 0	2 1.92	1 0.69	0 0.00	59 64.84
Total N	465 100.00	111 100.00	110 100.00	104 100.00	144 100.00	81 100.00	91 100.00

Las celdas resaltadas con un fondo gris indican la forma más utilizada en el presente estudio, los resultados en negrilla corresponden con la forma más utilizadas en el estudio oral de ByLQ.

Como mencionado anteriormente, al comparar las narrativas escritas entre HNs y HNNs, utilizaremos los datos sin la categoría GenNP, pero estos se incluirán cuando comparemos los datos escritos con los orales de ByLQ quienes consideraron algunas expresiones referenciales demasiado generales semánticamente como para poder identificar al referente. Se incluyeron expresiones como *los dos, las dos mujeres, el otro, uno de los hombres, uno de ellos, etc.* y crearon la categoría de GenNP para poder recoger estas expresiones. Sin embargo, en el presente estudio la mayoría de las

expresiones GenNP fueron reclasificadas como DefNP o IndNP. De hecho, en los datos obtenidos para este estudio, solo hubo un caso donde el referente no era recuperable con la forma producida. Los ejemplos (1) y (2) que siguen a continuación son casos representativos de GenNP que han sido vueltos a clasificarse como DefNP e IndNP, respectivamente, y la Tabla 5.2.3 muestra los resultados generales del presente estudio sin la categoría GenNP:

(1) L16. ...*el otro hombre* aparece de nuevo... [P8.L1.TC]²⁰

(2) L3. *Otra mujer* pasa por el banco y [P40.L2.NTC]

En (1) y (2) el contexto nos ayuda a identificar el personaje: en (1), la expresión referencial *el otro hombre* aparece en la línea 16, pero más arriba, en las líneas 8 y 9, el participante nos había introducido a este personaje como a Charlie Chaplin: “aparece otro hombre (Chaplin), también a conquistarla” (el nombre propio entre paréntesis hace parte de la redacción de este participante). Ya que, en general, una expresión definida se utiliza cuando el referente es identificable, en este caso *el otro hombre* es plenamente identificable por lo que ha sido previamente introducido en la narrativa como Chaplin. La expresión, así sea general, permite identificar el referente sin problemas de ambigüedad. Por otro lado, en (2) el referente no es identificable con la expresión *otra mujer*, el contexto tampoco ofrece posibilidades de identificación y, de hecho, el personaje es introducido en la narrativa en ese momento. De esta manera, la expresión puede clasificarse como IndNP.

El ejemplo (3) más abajo es el único caso que se encontró de una expresión referencial que no permite la identificación del personaje. Aunque la expresión es

²⁰ L: # línea; P: # participante; L1/L2: nativo/no nativo; TC/NTC: con Terreno Común/sin Terreno Común.

definida, haciendo suponer un referente que es por lo menos familiar, en la escena hay dos mujeres y, por lo tanto, la sola expresión no permite la recuperación del referente, como tampoco lo hace el contexto. Además, al no haber TC entre el lector y el escritor, tampoco se puede recurrir a este:

(3) L5. *La otra mujer* está sentada en el banco pero no duerme [P14.L1.NTC]

Tabla 5.2.3 - Frecuencia y distribución de formas de sujeto por estado cognitivo (los dos grupos) en narrativas escritas (sin GenNP)

Frecuencia Porcentaje	INF	ACTR	ACT	ACTNR	FAM	UNI	TYP
Null	374 80.43	60 54.05	8 7.27	6 5.77	2 1.39	1 1.23	0 0.00
Overt	31 6.67	12 10.81	25 22.73	0 0.00	0 0.00	0 0.00	0 0.00
Dem	2 0.43	2 1.80	6 5.45	1 0.96	2 1.39	3 3.70	1 1.10
DefNP	38 8.17	25 22.52	58 52.73	62 59.62	101 70.14	55 67.90	5 5.49
PrpN	20 4.30	12 10.81	12 10.91	23 22.12	38 26.39	21 25.93	0 0.00
IndNP	0 0.00	0 0.00	1 0.91	12 11.54	1 0.69	1 1.23	85 93.41
Total N	465 100.00	111 100.00	110 100.00	104 100.00	144 100.00	81 100.00	91 100.00

Las celdas resaltadas con un fondo gris indican la forma más utilizada en el presente estudio

En las siguientes secciones, para cada EC se sigue el mismo formato de presentación de los resultados: en un primer lugar se presenta una tabla de frecuencia que muestra las diferencias de selección de uso de las formas en las narrativas escritas del presente estudio, entre L1 y L2. Como mencionado anteriormente, las formas resaltadas con un fondo gris son aquellas que se eliminan para la prueba de chi-cuadrado y en la subsecuente aplicación del modelo de regresión logística. Los valores resaltados en

negrilla indican que hay significación estadística. Luego, se presentan las diferencias de selección de uso de las formas de sujeto entre las narrativas escritas del presente estudio y las narrativas orales de ByLQ, entre L1 y L2. También aquí solo se aplica el modelo de regresión logística sobre aquellas categorías con más de cinco instancias, y que no sean proporcionalmente iguales. Por último, se aplican ajustes logísticos que hacen parte de la investigación general donde se usa un modelo de regresión logística múltiple. Se analiza el efecto del Grupo (L1/L2) y el efecto de Producción (Oral/Escrito), y los resultados revelan si los factores Grupo y Producción afectan el uso de cada forma de sujeto.

5.2.2 Estado cognitivo INF

La Tabla 5.2.4 muestra las diferencias de selección de uso de las formas en las narrativas escritas, entre L1 y L2.

Tabla 5.2.4 – INF: Formas de sujeto entre L1 y L2 (instancias y porcentajes) en narrativas escritas

Frecuencia (Porcentaje)	L1	L2
Null	173 (81)	201 (80)
Overt	9 (4)	22 (9)
Dem	1 (0)	1 (0)
DefNP	23 (11)	15 (6)
PrpN	7 (3)	13 (5)
IndNP	0 (0)	0 (0)
Total N (%)	213 (100)	252 (100)

Comparación (L1-Escrito) versus (L2-Escrito)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=7.8023$, $p=0.0503$

Prueba exacta de Fisher $p=0.0722$

La prueba chi-cuadrado indica que las diferencias entre L1 y L2 en las narrativas escritas se encuentra en el borde de la significancia.

Tabla 5.2.5 – INF: Regresión logística para Null, Overt, DefNP y PrpN entre L1 y L2

Forma	P-valor
Null	0.6926
Overt	0.0481
DefNP	0.0576
PrpN	0.3167

La regresión logística hecha sobre las categorías de Null, Overt, DefNP y PrpN apoyan los resultados de que los hablantes L1 y L2 difieren en su uso de Overt y DefNP. Los HNs (L1) tienen una menor probabilidad estadística de utilizar Overt y una mayor probabilidad de utilizar DefNP que los HNNs (L2), cuando el referente está en foco.

La tabla 5.2.6 muestra las diferencias de selección de uso de las formas de sujeto entre las narrativas escritas del presente estudio y las orales de ByLQ, entre L1 y L2.

Tabla 5.2.6 – INF: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas

Frecuencia (Porcentaje)	L1 Oral ByLQ	L2 Oral ByLQ	L1 Escrito	L2 Escrito
Null	385 (91)	120(65)	173 (81)	201 (80)
Overt	23 (5)	49 (26)	9 (4)	22 (9)
GenNP	3 (1)	2 (1)	0 (0)	1 (0)
Dem	2 (0)	0 (0)	1 (0)	1 (0)
DefNP	8 (2)	6 (3)	23 (11)	14 (6)
PrpN	4 (1)	8 (4)	7 (3)	13 (5)
IndNP	0 (0)	0 (0)	0 (0)	0 (0)
Total N (%)	425 (100)	185 (99)	213 (99)	252 (100)

Comparación (L1-Oral) versus (L1-Escrito)

Chi-cuadrado sin (GenNP, Dem, IndNP) $df=3$, $X^2=29.484$, $p<0.0001$

Prueba exacta de Fisher $p < 0.0001$

Comparación (L2-Oral) versus (L2- Escrito)

Chi-cuadrado sin (GenNP, Dem, IndNP) $df=3$, $X^2=25.337$, $p < 0.0001$

Prueba exacta de Fisher $p < 0.0001$

Se encontraron diferencias estadísticas en el uso de expresiones referenciales de sujeto entre las narrativas orales de ByLQ y las escritas del presente estudio para L1 y para L2.

Tabla 5.2.7 -INF: Regresión logística sobre Null, Overt, DefNP y PrpN para L1 entre narrativas orales y escritas

Forma	<i>p</i> -valor
Null	0.0010
Overt	0.5115
DefNP	<0.0001
PrpN	0.0388

Tabla 5.2.8 - INF: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre narrativas orales y escritas

Forma	<i>p</i> -valor
Null	0.0005
Overt	<0.0001
DefNP	0.2445
PrpN	0.6855

La regresión logística en 5.2.7 muestra que los hablantes L1 tienen una menor probabilidad de utilizar Null y una mayor probabilidad de utilizar DefNP y PrpN en las narrativas escritas de este estudio que en las orales del estudio previo. Por otro lado, los hablantes L2 tienen una menor probabilidad de utilizar Overt y una mayor probabilidad de utilizar Null en las narrativas escritas que en las orales (Tabla 5.2.8).

Tabla 5.2.9 – INF: Ajuste logístico sobre Null

Término	Estimación	<i>p</i> -valor
Escrito	-0.8000	0.0009
L2	-1.6513	<0.0001
Escrito*L2	1.5583	<0.0001

Tabla 5.2.10 – INF: Ajuste logístico sobre Overt

Término	Estimación	<i>p</i> -valor
Escrito	-0.2599	0.5183
L2	1.8401	<0.0001
Escrito*L2	-1.0663	0.0293

Tabla 5.2.11 – INF: Ajuste logístico sobre DefNP

Término	Estimación	<i>p</i> -valor
Escrito	1.8421	<0.0001
L2	0.5580	0.3080
Escrito*L2	-1.2797	0.0494

Tabla 5.2.12 – INF: Ajuste logístico sobre PrpN

Término	Estimación	<i>p</i> -valor
L2	1.0569	0.0051

La investigación general a través de los ajustes logísticos, tomando en cuenta y comparando los resultados del presente estudio y los de ByLQ, indica que los factores Producción y Grupo afectan la selección de la referencia de sujeto de manera significativa de la siguiente manera:

- Hay una tendencia decreciente en el uso de Null en el grupo L1 y una tendencia creciente en el grupo L2 cuando el medio de producción de narración es la escritura (Tabla 5.2.9).

- Hay una tendencia decreciente en el uso de Overt tanto en L1 como en L2 en la escritura. Este efecto es mayor en L2 (Tabla 5.2.10).
- Hay una tendencia creciente en el uso de DefNP tanto en L1 como L2 en la escritura. El efecto es mayor en L1 (Tabla 5.2.11).
- Hay una tendencia creciente en el uso de PrpN tanto en L1 como L2 en la escritura, pero solo tiene significancia estadística en L2 (Tabla 5.2.12).

5.2.3 Estado cognitivo ACTR

La Tabla 5.2.13 muestra las diferencias de selección de uso de las formas de sujeto en las narrativas escritas, entre L1 y L2.

Tabla 5.2.13 – ACTR: Formas de sujeto entre L1 y L2

Frecuencia (Porcentaje)	L1	L2
Null	24 (55)	36 (54)
Overt	2 (5)	10 (15)
Dem	1 (2)	1 (1)
DefNP	10 (23)	15 (22)
PrpN	7 (16)	5 (7)
IndNP	0 (0)	0 (0)
Total N (%)	44 (101)	67 (99)

Comparación (L1-Escrito) versus (L2-Escrito)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=4.4098$, $p=0.2205$

Prueba exacta de Fisher $p=0.2282$

No se encontraron diferencias significativas entre L1 y L2, en las narrativas escritas.

Tabla 5.2.14 – ACTR: Regresión logística para Overt y PrpN entre L1 y L2

Forma	<i>p</i> -valor
Overt	0.0692
PrpN	0.1663

No se encontraron diferencias significativas entre los dos grupos para Overt y PrpN.

Tabla 5.2.15 – ACTR: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas

Frecuencia (Porcentaje)	L1 Oral ByLQ	L2 Oral ByLQ	L1 Escrito	L2 Escrito
Null	79 (82)	24 (59)	24 (55)	36 (54)
Overt	8 (8)	14 (34)	2 (5)	10 (15)
GenNP	4 (4)	0 (0)	1 (2)	3 (4)
Dem	2 (2)	0 (0)	1 (2)	1 (1)
DefNP	2 (2)	1 (2)	9 (20)	12 (18)
PrpN	1 (1)	2 (5)	7 (16)	5 (7)
IndNP	0 (0)	0 (0)	0 (0)	0 (0)
Total N (%)	96 (99)	41 (100)	44 (100)	67 (99)

Comparación (L1-Oral) versus (L1-Escrito)

Chi-cuadrado sin (Dem, IndNP) $df=4$, $X^2=28.718$, $P<0.0001$

Prueba exacta de Fisher $p<0.0001$

Comparación (L2-Oral) versus (L2-Escrito)

Chi-cuadrado sin (GenNP, Dem, IndNP) $df=3$, $X^2=9.4181$, $P=0.0241$

Prueba exacta de Fisher $p=0.0198$

Tanto L1 como L2 muestran diferencias significativas en el uso de sujetos entre las narrativas orales y escritas.

Tabla 5.2.16 - ACTR: Regresión logística sobre Null, Overt, DefNP y PrpN para L1 entre narrativas orales y escritas

Forma	<i>p</i> -valor
Null	0.0007
Overt	0.4012
DefNP	0.0003
PrpN	0.0006

Tabla 5.2.17 – ACTR: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre narrativas orales y escritas

Forma	<i>p</i> -valor
Null	0.6254
Overt	0.0213
DefNP	0.0080
PrpN	0.5896

Los resultados de la regresión logística indican que los hablantes L1 tienen una menor probabilidad de utilizar Null en las narrativas escritas de este estudio que en las orales de ByLQ, mientras que tienen una mayor probabilidad de utilizar DefNP y PrpN (Tabla 5.2.16). Los hablantes L2 tienden a utilizar menos Overt y más DefNP en las narrativas escritas que en las orales (Tabla 5.2.17).

Tabla 5.2.18 - ACTR: Ajuste logístico sobre Null

Término	Estimación	<i>p</i> -valor
Escrito	-1.3390	0.0008
L2	-1.1914	0.0041
Escrito*L2	1.1586	0.0417

Tabla 5.2.19 - ACTR: Ajuste logístico sobre Overt

Término	Estimación	<i>p</i> -valor
Escrito	-0.9802	0.0191
L2	1.6296	0.0001

Tabla 5.2.20 - ACTR: Ajuste logístico sobre DefNP

Término	Estimación	<i>p</i> -valor
Escrito	2.3439	0.0002

Tabla 5.2.21 - ACTR: Ajuste logístico sobre PrpN

Término	Estimación	<i>p</i> -valor
Escrito	1.6890	0.0104

La investigación general a través de los ajustes logísticos, tomando en cuenta y comparando los resultados del presente estudio y los de ByLQ, indica que los factores Producción y Estilo afectan la selección de la referencia de sujeto de manera significativa de la siguiente manera:

- Hay una tendencia decreciente en el uso de Null en los grupos L1 y L2, cuando el medio de producción de narración es la escritura. Este efecto es mayor en los hablantes L1 (Tabla 5.2.18).
- Hay una tendencia decreciente en el uso de Overt en las narrativas escritas. Este efecto tiene significancia solamente en los hablantes L1 (Tabla 5.2.19).
- Hay una tendencia creciente en el uso de DefNP en las narrativas escritas. Esta tendencia es significativa tanto en L1 como en L2 (Tabla 5.2.20).
- Hay una tendencia creciente en el uso de PrpN en las narrativas escritas. Esta tendencia es significativa tanto en L1 como en L2 (Tabla 5.2.21).

5.2.4 Estado cognitivo ACT

La Tabla 5.2.22 muestra las diferencias de selección de uso de las formas en las narrativas escritas, entre L1 y L2.

Tabla 5.2.22 - ACT: Formas de sujeto entre L1 y L2

Frecuencia (Porcentaje)	L1 Escrito	L2 Escrito
Null	3 (6)	5 (9)
Overt	12 (23)	13 (23)
Dem	4 (8)	2 (4)
DefNP	28 (53)	30 (53)
PrpN	5 (9)	7 (12)
IndNP	1 (2)	0 (0)
Total N (%)	53 (101)	57 (101)

Comparación (L1-Escrito) versus (L2-Escrito)

Chi-cuadrado sin IndNP $df=4$, $X^2=1.3825$, $p=0.8472$

Prueba exacta de Fisher $p=0.8487$

No se encontraron diferencias significativas entre L1 y L2, en las narrativas escritas.

Tabla 5.2.23 - ACT: Regresión logística sobre Null, Dem y PrpN entre L1 y L2

Forma	P-valor
Null	0.5276
Dem	0.3481
PrpN	0.6314

No se encontraron diferencias significativas en el uso de Null, Dem y PrpN entre L1 y L2, en las narrativas escritas, a través del análisis de regresión logística.

Tabla 5.2.24 – ACT: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas

Frecuencia (Porcentaje)	L1 Oral ByLQ	L2 Oral ByLQ	L1 Escrito	L2 Escrito
Null	8 (12)	1 (2)	3 (6)	5 (9)
Overt	28 (41)	17 (33)	12 (23)	13 (23)
GenNP	3 (4)	0 (0)	1 (2)	0 (0)
Dem	3 (4)	2 (4)	4 (8)	2 (4)
DefNP	18 (26)	23 (44)	28 (53)	30 (53)
PrpN	9 (13)	9 (17)	5 (9)	7 (12)
IndNP	0 (0)	0 (0)	0 (0)	0 (0)
Total N (%)	69 (100)	52 (100)	53 (101)	57 (101)

Comparación (L1-Oral) versus (L1-Escrito)

Chi-cuadrado sin IndNP $df=5$, $X^2=11.227$, $p=0.0471$

Prueba exacta de Fisher $p=0.0418$

Comparación (L2-Oral) versus (L2-Escrito)

Chi-cuadrado sin (GenNP, IndNP) $df=4$, $X^2=4.1539$, $p=0.3856$

Prueba exacta de Fisher $p=0.4028$

Los resultados de la Tabla 5.2.24 muestran que hay diferencias significativas entre las narrativas escritas del presente estudio y las orales de ByLQ en el uso de las formas de sujeto para los hablantes L1. En contraste, los hablantes L2 demuestran un uso similar en ambos medios de producción de las narrativas.

Tabla 5.2.25 - ACT: Regresión logística sobre Null, Overt, Dem, DefNP y PrpN para L1 entre narrativas orales y escritas

Forma	p -valor
Null	0.2460
Overt	0.0343
Dem	0.4538
DefNP	0.0025
PrpN	0.5320

Tabla 5.2.26 - ACT: Regresión logística sobre Null, Overt, Dem, DefNP y PrpN para L2 entre narrativas orales y escritas

Forma	<i>p</i> -valor
Null	0.1009
Overt	0.2483
Dem	0.9255
DefNP	0.3805
PrpN	0.4589

El análisis de regresión logística indica que, en las narrativas escritas, los hablantes L1 tienen más tendencia a utilizar DefNP y menos tendencia a utilizar Overt en las narrativas escritas que en las orales (Tabla 5.2.25). Los hablantes L2 no demuestran diferencias significativas en el uso de los sujetos en ACT, entre las narrativas orales y escritas (Tabla 5.2.26). No se llevó a cabo una regresión logística múltiple para Null, Dem ni PrpN, ya que estas formas no arrojaron resultados significativos para ningún grupo.

Tabla 5.2.27 - ACT: Ajuste logístico sobre Overt

Término	Estimación	<i>p</i> -valor
Escrito	-0.6997	0.0178

Tabla 5.2.28 - ACT: Ajuste logístico sobre DefNP

Término	Estimación	<i>p</i> -valor
Escrito	0.7777	0.0041

Los resultados del ajuste logístico a través del modelo de regresión múltiple validan las anteriores conclusiones con respecto al uso de Overt y DefNP en el estado activado: cuando el medio de producción de la narrativa es escrito, hay una tendencia

significativa hacia la disminución en el uso de Overt y hacia el aumento en el uso de DefNP para los dos grupos, L1 y L2.

5.2.5 Estado cognitivo ACTNR

La Tabla 5.2.29 muestra las diferencias de selección de uso de las formas en las narrativas escritas, entre L1 y L2.

Tabla 5.2.29 - ACTNR: Formas de sujeto entre L1 y L2

Frecuencia (Porcentaje)	L1 Escrito	L2 Escrito
Null	4 (8)	2 (4)
Overt	0 (0)	0 (0)
Dem	1 (2)	0 (0)
DefNP	28 (56)	34 (63)
PrpN	8 (16)	15 (28)
IndNP	9 (18)	3 (6)
Total N (%)	50 (100)	54 (101)

Comparación (L1-Escrito) versus (L2-Escrito)

Chi-cuadrado sin (Overt, Dem) $df=3$, $X^2=6.1495$, $p=0.1046$

Prueba exacta de Fisher $p=0.1060$

En general, no se encontraron diferencias significativas entre L1 y L2, en las narrativas escritas.

Tabla 5.2.30 - ACTNR: Regresión logística sobre Null, DefNP, PrpN y IndNP entre L1 y L2

Forma	p -valor
Null	0.3446
DefNP	0.4696
PrpN	0.1452
IndNP	0.0435

El análisis de regresión logística indica que, en las narrativas escritas, los hablantes L1 muestran significativamente una mayor probabilidad de utilizar IndNP que los hablantes L2 cuando el referente está en el estado ACTNR.

Tabla 5.2.31 – ACTNR: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas

Frecuencia (Porcentaje)	L1 Oral ByLQ	L2 Oral ByLQ	L1 Escrito ByLQ	L2 Escrito ByLQ
Null	0 (0)	1 (3)	4 (8)	2 (4)
Overt	1 (2)	0 (0)	0 (0)	0 (0)
GenNP	22 (44)	4 (10)	14 (28)	7 (13)
Dem	0 (0)	0 (0)	1 (2)	0 (0)
DefNP	15 (30)	10 (26)	22 (44)	29 (54)
PrpN	12 (24)	23 (59)	8 (16)	15 (28)
IndNP	0 (0)	1 (3)	1 (2)	1 (2)
Total N (%)	50 (100)	39 (101)	50 (100)	54 (101)

Comparación (L1-Oral) versus (L1-Escrito)

Chi-cuadrado sin (Overt, Dem, IndNP) $df=3$, $X^2=7.8926$, $p=0.0483$

Prueba exacta de Fisher $p=0.0459$

Comparación (L2-Oral) versus (L2-Escrito)

Chi-cuadrado sin (Overt, Dem, IndNP) $df=3$, $X^2=9.8883$, $p=0.0195$

Prueba exacta de Fisher $p=0.0176$

Tanto L1 como L2 demuestran diferencias significativas en el uso de las referencias de sujeto, entre las narrativas orales y escritas.

Tabla 5.2.32 - ACTNR: Regresión logística sobre Null, GenNP, DefNP y PrpN para L1 entre narrativas orales y escritas

Forma	<i>p</i> -valor
Null (Firth)	0.0499
GenNP	0.0945
DefNP	0.1462
PrpN	0.3160

Tabla 5.2.33 - ACTNR: Regresión logística sobre Null, GenNP, DefNP y PrpN para L2 entre narrativas orales y escritas

Forma	<i>p</i> -valor
Null	0.7561
GenNP	0.6881
DefNP	0.0061
PrpN	0.0024

El análisis de regresión logística indica que los hablantes L1 tienen una mayor tendencia a utilizar Null en las narrativas escritas del presente estudio vs las orales de ByLQ (Tabla 5.2.32). Los hablantes L2 tienden a utilizar menos PrpN y más DefNP en las narrativas escritas que en las orales (Tabla 5.2.33).

No se llevó a cabo una regresión logística múltiple para Null, ya que esta forma no arrojó resultados significativos para ningún grupo.

Tabla 5.2.34 - ACTNR: Ajuste logístico sobre GenNP

Término	Estimación	<i>p</i> -valor
L2	-1.4334	0.0002

Tabla 5.2.35 - ACTNR: Ajuste logístico sobre DefNP

Término	Estimación	<i>p</i> -valor
Escrito	0.9015	0.0033

Tabla 5.2.36 – ACTNR: Ajuste logístico sobre PrpN

Término	Estimación	p-valor
Escrito	-0.9781	0.0037
L2	1.1482	0.0007

La investigación general a través de los ajustes logísticos, tomando en cuenta y comparando los resultados del presente estudio y los de ByLQ, indica que los factores Producción y Estilo afectan la selección de la referencia de sujeto de manera significativa de la siguiente manera:

- Los hablantes L1 utilizan más GenNP que los hablantes L2 en ambos medios de producción, escrito y oral (Tabla 5.2.34)
- Hay una tendencia creciente en el uso de DefNP en L1 y L2 cuando el medio de producción de la narración es la escritura. (Tabla 5.2.35)
- Hay una tendencia decreciente en el uso de PrpN en L1 y L2 cuando el medio de producción es la escritura (Tabla 5.2.36)

5.2.6 Estado cognitivo FAM

La Tabla 5.2.37 muestra las diferencias de selección de uso de las formas en las narrativas escritas, entre L1 y L2.

Tabla 5.2.37 - FAM: Formas de sujeto entre L1 y L2

Frecuencia (Porcentaje)	L1 Escrito	L2 Escrito
Null	0 (0)	2 (3)
Overt	0 (0)	0 (0)
Dem	1 (1)	1 (1)
DefNP	44 (65)	57 (75)
PrpN	22 (32)	16 (21)
IndNP	1 (1)	0 (0)
Total N (%)	68 (99)	76 (100)

Comparación (L1-Escrito) versus (L2-Escrito)

Chi-cuadrado solamente con DefNP y PrpN $df=1$, $X^2=1.7355$, $p=0.1877$

Prueba exacta de Fisher $p=0.2099$

No se encontraron diferencias significativas entre L1 y L2, en las narrativas escritas.

Tabla 5.2.38 - FAM: Regresión logística sobre DefNP y PrpN entre L1 y L2

Forma	P-valor
DefNP	0.1778
PrpN	0.1244

El análisis de regresión logística indica que no se encontraron diferencias significativas en el uso de DefNP y PrpN entre L1 y L2, en las narrativas escritas.

Tabla 5.2.39 – FAM: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas

Frecuencia (Porcentaje)	L1 Oral ByLQ	L2 Oral ByLQ	L1 Escrito	L2 Escrito
Null	3 (4)	0 (0)	0 (0)	2 (3)
Overt	1 (1)	0 (0)	0 (0)	0 (0)
GenNP	10 (12)	2 (3)	7 (10)	3 (4)
Dem	0 (0)	0 (0)	1 (1)	1 (1)
DefNP	55 (67)	39 (60)	37 (54)	54 (71)
PrpN	13 (16)	24 (37)	22 (32)	16 (21)
IndNP	0 (0)	0 (0)	1 (1)	0 (0)
Total N (%)	82 (100)	65 (100)	68 (98)	76 (98)

Comparación (L1-Oral) versus (L1-Escrito)

Chi-cuadrado con GenNP, DefNP y PrpN $df=2$, $X^2=5.403$, $p=0.0671$

Prueba exacta de Fisher $p=0.0396$

Comparación (L2-Oral) versus (L2-Escrito)

Chi-cuadrado con GenNP, DefNP y PrpN $df=2$, $X^2=3.7683$, $p=0.1520$

Prueba exacta de Fisher $p=0.1253$

Los resultados demuestran que hay diferencias significativas entre las narrativas escritas del presente estudio y las orales de ByLQ en el uso de las formas de sujeto para los hablantes L1. En contraste, los hablantes L2 demuestran un uso similar en ambos medios de producción.

Tabla 5.2.40 - FAM: Regresión logística sobre GenNP, DefNP y PrpN para L1 entre narrativas orales y escritas

Forma	<i>p</i> -valor
GenNP	0.7139
DefNP	0.1130
PrpN	0.0173

Tabla 5.2.41 - FAM: Regresión logística sobre GenNP, DefNP y PrpN para L2 entre narrativas orales y escritas

Forma	<i>p</i> -valor
GenNP	0.7796
DefNP	0.1676
PrpN	0.0371

El análisis de regresión logística sobre GenNP, DefNP y PrpN indica que, cuando el referente está en el estado FAM, los hablantes L1 tienen una mayor tendencia a utilizar PrpN en las narrativas escritas que en las orales, mientras que los hablantes L2 tienen una tendencia opuesta.

Tabla 5.2.42 - FAM: Ajuste logístico sobre GenNP

Término	Estimación	<i>p</i> -valor
L2	-1.2461	0.0172

Tabla 5.2.43 - FAM: Ajuste logístico sobre DefNP

Término	Estimación	<i>p</i> -valor
Escrito	-0.5346	0.1142
L2	-0.3060	0.3757
Escrito*L2	1.0270	0.0371

Tabla 5.2.44 - FAM: Ajuste logístico sobre PrpN

Término	Estimación	<i>p</i> -valor
Escrito	0.9316	0.0193
L2	1.1336	0.0043
Escrito*L2	-1.7178	0.0018

La investigación general a través de los ajustes logísticos, tomando en cuenta y comparando los resultados del presente estudio y los de ByLQ, indica que los factores Producción y Estilo afectan la selección de la referencia de sujeto de manera significativa de la siguiente manera:

- Los hablantes L2 tienen una tendencia significativa a utilizar menos GenNP que los hablantes L1, tanto en las narrativas orales como en las escritas (Tabla 5.2.42).
- Hay una tendencia decreciente en el uso de DefNP por parte de los hablantes L1 cuando el medio de producción de las narrativas es escrito (Tabla 5.2.43).
- Hay una tendencia creciente de los hablantes L1 a utilizar PrpN cuando el medio de producción es la escritura; mientras que los hablantes L2 tienden a usar más PrpN en las narrativas orales (Tabla 5.2.44).

5.2.7 Estado cognitivo UNI

La Tabla 5.2.45 muestra las diferencias de selección de uso de las formas en las narrativas escritas, entre L1 y L2.

Tabla 5.2.45 - UNI: Formas de sujeto entre L1 y L2

Frecuencia (Porcentaje)	L1	L2
Null	0 (0)	1 (2)
Overt	0 (0)	0 (0)
Dem	0 (0)	3 (7)
DefNP	25 (68)	30 (68)
PrpN	11 (30)	10 (23)
IndNP	1 (3)	0 (0)
Total N (%)	37 (101)	44 (100)

Comparación (L1-Escrito) versus (L2-Escrito)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=0.0806$, $p=0.7765$

Prueba exacta de Fisher $p=0.6888$

No se encontraron diferencias significativas entre L1 y L2, en las narrativas escritas.

Tabla 5.2.46 - UNI: Regresión logística sobre PrpN entre L1 y L2

Forma	p -valor
PrpN	0.4744

El análisis de regresión logística indica que no se encontró ninguna diferencia significativa en el uso de PrpN entre L1 y L2, en las narrativas escritas.

Tabla 5.2.47 – UNI: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas

Frecuencia (Porcentaje)	L1	L2	L1	L2
	Oral ByLQ	Oral ByLQ	Escrito	Escrito
Null	0 (0)	0 (0)	0 (0)	1 (2)
Overt	0 (0)	0 (0)	0 (0)	0 (0)
GenNP	0 (0)	0 (0)	1 (3)	0 (0)
Dem	0 (0)	0 (0)	0 (0)	3 (7)
DefNP	24 (83)	9 (53)	25 (68)	30 (68)
PrpN	5 (17)	8 (47)	11 (30)	10 (23)
IndNP	0 (0)	0 (0)	0 (0)	0 (0)
Total N (%)	29 (100)	17 (100)	37 (101)	44 (100)

Comparación (L1-Oral) versus (L1-Escrito)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=0.9007$, $p=0.3426$

Prueba exacta de Fisher $p=0.3135$

Comparación (L2-Oral) versus (L2-Escrito)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=1.7627$, $p=0.1843$

Prueba exacta de Fisher $p=0.2160$

No se encontraron diferencias significativas en el uso de sujetos entre las narrativas escritas del presente estudio y las orales de ByLQ.

Tabla 5.2.48 - UNI: Regresión logística sobre DefNP y PrpN para L1 entre narrativas orales y escritas

Forma	p -valor
DefNP	0.1554
PrpN	0.2344

Tabla 5.2.49 - UNI: Regresión logística sobre DefNP y PrpN para L2 entre narrativas orales y escritas

Forma	p -valor
DefNP	0.2710
PrpN	0.0677

La regresión logística indica que no hay diferencias significativas según la forma de sujeto.

Tabla 5.2.50 - UNI: Ajuste logístico sobre DefNP

Término	Estimación	<i>p</i> -valor
Escrito	-0.8346	0.1671
L2	-1.4508	0.0358
Escrito*L2	1.4790	0.0783

Tabla 5.2.51 - UNI: Ajuste logístico sobre PrpN

Término	Estimación	<i>p</i> -valor
Escrito	0.7084	0.2448
L2	1.4508	0.0358
Escrito*L2	-1.8144	0.0345

La investigación general a través de los ajustes logísticos, tomando en cuenta y comparando los resultados del presente estudio y los de ByLQ, indica que los factores Producción y Estilo afectan la selección de la referencia de sujeto de manera significativa de la siguiente manera:

- Hay una menor tendencia de los hablantes L1 a utilizar DefNP en las narrativas escritas cuando el referente es UNI. Esta tendencia es opuesta para PrpN.
- Los hablantes L2 tienden a utilizar menos DefNP que los L1 en las narrativas orales, pero más PrpN. La tendencia es opuesta en las narrativas escritas.

5.2.8 Estado cognitivo TYP

La Tabla 5.2.52 muestra las diferencias de selección de uso de las formas en las narrativas escritas, entre L1 y L2.

Tabla 5.2.52 - TYP: Formas de sujeto entre L1 y L2

Frecuencia (Porcentaje)	L1	L2
Null	0 (0)	0 (0)
Overt	0 (0)	0 (0)
Dem	1 (2)	0 (0)
DefNP	5 (10)	0 (0)
PrpN	0 (0)	0 (0)
IndNP	46 (88)	39 (100)
Total N (%)	52 (100)	39 (100)

Comparación (L1-Escrito) versus (L2-Escrito)

Chi-cuadrado con DefNP y IndNP $df=1$, $X^2=2.3955$, $p=0.1217$

Prueba exacta de Fisher $p=0.0683$

En general, no se encontraron diferencias significativas entre L1 y L2, en las narrativas escritas.

Tabla 5.2.53 - TYP: Regresión logística sobre DefNP y IndNP entre L1 y L2

Forma	<i>p</i> -valor
DefNP (Firth)	0.0516
IndNP (Firth)	0.0277

El modelo de regresión logística indica que, en las narrativas escritas, los hablantes L1 tienden a utilizar más DefNP que los hablantes L2, mientras que los hablantes L2 tienden a utilizar más IndNP que los hablantes L1.

Tabla 5.2.54 – TYP: Formas de sujeto entre L1 y L2, en las narrativas orales y escritas

Frecuencia (Porcentaje)	L1 Oral ByLQ	L2 Oral ByLQ	L1 Escrito	L2 Escrito
Null	0 (0)	0 (0)	0 (0)	0 (0)
Overt	1 (2)	0 (0)	0 (0)	0 (0)
GenNP	0 (0)	0 (0)	14 (27)	12 (31)
Dem	0 (0)	0 (0)	1 (2)	0 (0)
DefNP	3(7)	5 (23)	5 (10)	0 (0)
PrpN	0 (0)	0 (0)	0 (0)	0 (0)
IndNP	38 (90)	17 (77)	32 (62)	27 (69)
Total N (%)	42 (99)	22 (100)	52 (101)	39 (100)

Comparación (L1-Oral) versus (L1-Escrito)

Chi-cuadrado con GenNP, DefNP y IndNP $df=2$, $X^2=14.094$, $p=0.0008$

Prueba exacta de Fisher $p=0.0001$

Comparación (L2-Oral) versus (L2-Escrito)

Chi-cuadrado con GenNP, DefNP y IndNP $df=2$, $X^2=15.759$, $p=0.0004$

Prueba exacta de Fisher $p<0.0001$

Tanto los hablantes L1 como L2 muestran diferencias significativas en el uso de sujetos referenciales entre las narrativas escritas del presente estudio y las orales de ByLQ, cuando el referente está en TYP.

Tabla 5.2.55 - TYP: Regresión logística sobre GenNP, DefNP y IndNP para L1 entre narrativas orales y escritas

Forma	p -valor
GenNP (Firth)	<0.0001
DefNP	0.6673
IndNP	0.0008

Tabla 5.2.56 - TYP: Regresión logística sobre GenNP, DefNP y IndNP para L2 entre narrativas orales y escritas

Forma	<i>p</i> -valor
GenNP (Firth)	0.0019
DefNP (Firth)	0.0028
IndNP	0.4969

El análisis de regresión logística indica que tanto los hablantes L1 como los L2 tienden a utilizar más GenNP cuando el medio de producción es escrito. Los hablantes L1 tienden a utilizar menos IndNP en las narrativas escritas que en las orales (Tabla 5.2.55). Los hablantes L2 tienden a utilizar menos DefNP en las narrativas escritas que las orales (Tabla 5.2.56).

Tabla 5.2.57 - TYP: Ajuste logístico sobre GenNP

Término	Estimación	<i>p</i> -valor
Escrito	3.9549	<0.0001

Tabla 5.2.58 - TYP: Ajuste logístico sobre DefNP

Término	Estimación	<i>p</i> -valor
Escrito	0.2676	0.0707
L2	1.2661	0.0829
Escrito*L2	-3.4796	0.0093

Tabla 5.2.59 - TYP: Ajuste logístico sobre IndNP

Término	Estimación	<i>p</i> -valor
Escrito	-1.1983	0.0045

La investigación general a través de los ajustes logísticos, tomando en cuenta y comparando los resultados del presente estudio y los de ByLQ, indica que los factores Producción y Estilo afectan la selección de la referencia de sujeto de manera significativa de la siguiente manera:

- Hay una tendencia creciente en el uso de GenNP en L1 y L2 cuando el medio de producción es la escritura (Tabla 5.2.57).
- Hay una tendencia decreciente en el uso de DefNP en el grupo L2 cuando el medio de producción es la escritura (Tabla 5.2.58).
- Hay una tendencia decreciente en el uso de IndNP en L1 y L2 cuando el medio de producción es la escritura (Tabla 5.2.59).

5.3 Resultados cuantitativos: Terreno Común (TC) vs No Terreno Común (NTC)

En esta sección se presentan los resultados cuantitativos aislando el factor TC. Para este análisis, solo se consideran los resultados de las narrativas escritas del presente estudio. Puesto que ByLQ no tomaron en cuenta este factor, no es posible una comparación con sus datos orales. Así como en la sección 5.2, tenemos siete subsecciones debido a que el análisis se lleva a cabo para cada uno de los siete EC. A su vez, cada subsección incluye las siguientes comparaciones:

- Comparación entre las narrativas TC y NTC
- Comparación de L1 y L2 entre las narrativas TC y NTC
 - Comparación L1 vs L2 en narrativas TC
 - Comparación L1 y L2 en narrativas NTC
 - Comparación entre TC y NTC para L1

- Comparación entre TC y NTC para L2
- Comparación general entre Grupo (L1/L2) y TC (TC/NTC)

Para el análisis, aplicamos la prueba chi-cuadrado y la prueba exacta de Fisher (ver Capítulo 4). En aquellos casos en los cuales los resultados arrojan valores con significación estadística o al límite de la significación (en ambos casos resaltados en letra negrilla), podemos concluir que hay diferencias en el uso de las formas referenciales entre las narrativas TC y NTC. Para poder determinar cuáles son las formas que difieren, se pasa a aplicar un modelo de regresión logística, el cual se lleva a cabo únicamente sobre aquellas formas con instancias por encima de cinco. También se excluyen aquellas formas cuyos porcentajes de distribución son iguales o que difieren de un 1%, ya que un modelo de regresión no es necesario en estos casos. Las celdas resaltadas en gris en las tablas son aquellas que se eliminan.

La comparación entre los factores Grupo y TC es una investigación general donde se realiza un ajuste logístico a través de la aplicación de un modelo de regresión logística múltiple. Los resultados revelan cómo Grupo y TC afectan el uso de cada forma de sujeto. En este modelo, el valor de p refleja la significación (con un nivel de significación establecido igual o menor de 0.05 para la prueba chi-cuadrado y de 0.10 para la prueba exacta de Fisher), y el coeficiente mide el tamaño del efecto; es decir, indica hasta qué punto el factor bajo análisis afecta los resultados. En este análisis general solo se incluyen aquellos resultados que arrojaron significación estadística.

5.3.1 Estado cognitivo INF

La Tabla 5.3.1 muestra la distribución de las formas de sujeto entre TC y NTC, para todos los participantes (grupos L1 y L2 combinados), en el estado INF.

Tabla 5.3.1 – INF: Formas de sujeto entre TC y NTC

Frecuencia (Porcentaje)	TC	NTC
Null	151 (78)	223 (82)
Overt	13 (7)	18 (7)
Dem	1 (1)	1 (0)
DefNP	17 (9)	21 (8)
PrpN	11 (6)	9 (3)
IndNP	0 (0)	0 (0)
Total N (%)	193 (101)	272 (100)

Comparación TC vs NTC

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=1.8632$, $p=0.6013$

Prueba exacta de Fisher $p=0.7169$

En general, no se encontraron diferencias significativas entre las narrativas TC y NTC, cuando el referente está INF.

Tabla 5.3.2 – INF: Regresión logística sobre Null, DefNP y PrpN entre TC y NTC

Forma	p -valor
Null	0.3173
DefNP	0.6714
PrpN	0.2146

No se encontraron resultados significativos en las formas de sujeto, a través del análisis de regresión logística.

Tabla 5.3.3 – INF: Formas de sujeto entre L1 y L2, TC y NTC

Frecuencia (Porcentaje)	L1 TC	L2 TC	L1 NTC	L2 NTC
Null	56 (77)	95 (79)	117 (84)	106 (80)
Overt	4 (5)	9 (8)	5 (4)	13 (10)
Dem	0 (0)	1 (1)	1 (1)	0 (0)
DefNP	12 (16)	5 (4)	11 (8)	10 (8)
PrpN	1 (1)	10 (8)	6 (4)	3 (2)
IndNP	0 (0)	0 (0)	0 (0)	0 (0)
Total N (%)	73 (99)	120 (100)	140 (101)	132 (100)

Comparación (L1-TC) vs (L1-NTC)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=5.1906$, $p=0.1584$

Prueba exacta de Fisher $p=0.2125$

Comparación (L2-TC) vs (L2-NTC)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=6.1082$, $p=0.1065$

Prueba exacta de Fisher $p=0.1053$

Comparación (L1-TC) vs (L2-TC)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=11.904$, $p=0.0077$

Prueba exacta de Fisher $p=0.0094$

Comparación (L1-NTC) vs (L2-NTC)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=4.9683$, $p=0.1741$

Prueba exacta de Fisher $p=0.1759$

Se encontraron diferencias significativas únicamente entre los participantes L1 y L2 en las narrativas con TC, entre todas las formas de sujeto, excluyendo Dem e IndNP.

Tabla 5.3.4 – INF: Regresión logística sobre Null, Overt, DefNP y PrpN para L1 entre TC y NTC

Forma	<i>p</i> -valor
Null	0.2294
Overt	0.5189
DefNP	0.0617
PrpN	0.2241

Tabla 5.3.5 – INF: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre TC y NTC

Forma	<i>p</i> -valor
Null	0.8226
Overt	0.5081
DefNP	0.2481
PrpN	0.0266

Tabla 5.3.6 – INF: Regresión logística sobre Null, Overt, DefNP y PrpN para TC entre L1 y L2

Forma	<i>p</i> -valor
Null	0.6895
Overt	0.5820
DefNP	0.0041
PrpN	0.0257

Tabla 5.3.7 – INF: Regresión logística sobre Null, Overt, DefNP y PrpN para NTC entre L1 y L2

Forma	<i>p</i> -valor
Null	0.4834
Overt	0.0348
DefNP	0.2481
PrpN	0.9307

Entre las narrativas TC y NTC, el análisis de regresión logística indica que los hablantes L2 tienden a utilizar más PrpN cuando se asume TC entre el escritor y el lector (Tabla 5.3.5). Entre los grupos L1 y L2, cuando hay TC, los hablantes L1 utilizan más

DefNP que L2 y los L2 utilizan más PrpN que L1 (Tabla 5.3.6). Cuando se asume que no hay TC, los hablantes L2 utilizan más Overt que L1 (Tabla 5.3.7).

Al realizar los ajustes logísticos para el análisis general de los factores Grupo y TC, solo se muestran aquellos con significación estadística, así como en la sección 5.1. El ajuste logístico sobre Null no muestra resultados significativos y por lo tanto no se presenta.

Tabla 5.3.8 – INF: Ajuste logístico sobre Overt

Término	Estimación	<i>p</i> -valor
L2	0.7739	0.0573

Tabla 5.3.9 – INF: Ajuste logístico sobre DefNP

Término	Estimación	<i>p</i> -valor
TC	0.8360	0.0605
L2	-0.0395	0.9308
TC*L2	-1.4700	0.0406

Tabla 5.3.10 – INF: Ajuste logístico sobre PrpN

Término	Estimación	<i>p</i> -valor
TC	-1.1706	0.2828
L2	-0.6551	0.3614
TC*L2	2.5339	0.0477

La investigación general a través de los ajustes logísticos sobre los factores Grupo y TC indica lo siguiente:

- Los hablantes L2 incrementan significativamente su uso de Overt cuando no hay TC (Tabla 5.3.8).
- El factor TC tiene un efecto sobre el uso de DefNP: cuando hay TC, los

hablantes L2 tienden a utilizar menos DefNP que los hablantes L1 (Tabla 5.3.9).

- El factor TC tiene un efecto sobre el uso de PrpN: cuando hay TC, los hablantes L2 tienden a utilizar más PrpN que los hablantes L1 (Tabla 5.3.10).

5.3.2 Estado cognitivo ACTR

La Tabla 5.3.11 muestra la distribución de las formas de sujeto entre TC y NTC, para todos los participantes (L1 y L2), en el estado ACTR.

Tabla 5.3.11 – ACTR: Formas de sujeto entre TC y NTC

Frecuencia (Porcentaje)	TC	NTC
Null	30 (59)	30 (50)
Overt	5 (10)	7 (12)
Dem	0 (0)	2 (3)
DefNP	9 (18)	16 (27)
PrpN	7 (14)	5 (8)
IndNP	0 (0)	0 (0)
Total N (%)	51 (101)	60 (100)

Comparación TC vs NTC

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=2.1861$, $p=0.5347$

Prueba exacta de Fisher $p=0.4659$

En general, no se encontraron diferencias significativas entre las narrativas TC y NTC, cuando el referente está ACTR.

Tabla 5.3.12 – ACTR: Regresión logística sobre Null, Overt, DefNP y PrpN entre TC y NTC

Forma	<i>p</i> -valor
Null	0.3501
Overt	0.7522
DefNP	0.2538
PrpN	0.3625

No se encontraron resultados significativos en las formas de sujeto, a través del análisis de regresión logística.

Tabla 5.3.13 – ACTR: Formas de sujeto entre L1 y L2, TC y NTC

Frecuencia (Porcentaje)	L1 TC	L2 TC	L1 NTC	L2 NTC
Null	10 (62)	20 (57)	14 (50)	16 (50)
Overt	0 (0)	5 (14)	2 (7)	5 (16)
Dem	0 (0)	0 (0)	1 (4)	1 (3)
DefNP	2 (12)	7 (20)	8 (29)	8 (25)
PrpN	4 (25)	3 (9)	3 (11)	2 (6)
IndNP	0 (0)	0 (0)	0 (0)	0 (0)
Total N (%)	16 (99)	35 (100)	28 (101)	32 (100)

Comparación (L1-TC) vs (L1-NTC)

Chi-cuadrado con (Null, DefNP, PrpN) $df=2$, $X^2=2.5571$, $p=0.2784$

Prueba exacta de Fisher $p=0.4136$

Comparación (L2-TC) vs (L2-NTC)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=0.4704$, $p=0.9253$

Prueba exacta de Fisher $p=0.9304$

Comparación (L1-TC) vs (L2-TC)

Chi-cuadrado con (Null, DefNP, PrpN) $df=2$, $X^2=2.1966$, $p=0.3334$

Prueba exacta de Fisher $p=0.2156$

Comparación (L2-TC) vs (L2-NTC)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=1.3496$, $p=0.7174$

Prueba exacta de Fisher $p=0.8731$

En general, no se encontraron diferencias significativas entre L1 y L2, ni entre TC y NTC.

Tabla 5.3.14 – ACTR: Regresión logística sobre Null, DefNP y PrpN para L1 entre TC y NTC

Forma	<i>p</i> -valor
Null	0.4213
DefNP	0.2052
PrpN	0.2214

Tabla 5.3.15 – ACTR: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre TC y NTC

Forma	<i>p</i> -valor
Null	0.5580
Overt	0.8779
DefNP	0.6240
PrpN	0.7169

Tabla 5.3.16 – ACTR: Regresión logística sobre Null, DefNP y PrpN para TC entre L1 y L2

Forma	<i>p</i> -valor
Null	0.7176
DefNP	0.5037
PrpN	0.1274

Tabla 5.3.17 – ACTR: Regresión logística sobre Overt, DefNP y PrpN para NTC entre L1 y L2

Forma	<i>p</i> -valor
Overt	0.2986
DefNP	0.7551
PrpN	0.5325

No se encontraron resultados significativos a través del análisis de regresión logística cuando el referente está ACTR, para ninguna de las formas de sujeto.

Los ajustes logísticos realizados sobre Null, Overt, DefNP y PrpN no arrojaron resultados significativos para los factores Grupo y TC; por lo tanto, en general, no se encontraron diferencias significativas entre los grupos, o entre las narrativas con y sin TC, cuando el referente está en el estado ACTR.

5.3.3 Estado cognitivo ACT

La Tabla 5.3.18 muestra la distribución de las formas de sujeto entre TC y NTC, para todos los participantes (L1 y L2), en el estado ACT.

Tabla 5.3.18 - ACT: Formas de sujeto entre TC y NTC

Frecuencia (Porcentaje)	TC	NTC
Null	3 (6)	5 (8)
Overt	10 (20)	15 (25)
Dem	0 (0)	6 (10)
DefNP	30 (60)	28 (47)
PrpN	6 (12)	6 (10)
IndNP	1 (2)	0 (0)
Total N (%)	50 (100)	60 (100)

Comparación TC vs NTC

Chi-cuadrado sin IndNP $df=4$, $X^2=6.5253$, $p=0.1632$

Prueba exacta de Fisher $p=0.1395$

En general, no se encontraron diferencias significativas entre las narrativas TC y NTC, cuando el referente está ACT.

Tabla 5.3.19 - ACT: Regresión logística sobre Null, Overt, Dem, DefNP y PrpN entre TC y NTC

Forma	<i>p</i> -valor
Null	0.6367
Overt	0.5319
Dem (Firth)	0.0206
DefNP	0.1623
PrpN	0.7381

Los resultados del análisis de regresión logística indican que los hablantes tienden a utilizar menos Dem en la narrativa, cuando se asume TC.

Tabla 5.3.20 - ACT: Formas de sujeto entre L1 y L2, TC y NTC

Frecuencia (Porcentaje)	L1 TC	L2 TC	L1 NTC	L2 NTC
Null	0 (0)	3 (11)	3 (10)	2 (7)
Overt	5 (22)	5 (19)	7 (23)	8 (27)
Dem	0 (0)	0 (0)	4 (13)	2 (7)
DefNP	14 (61)	16 (59)	14 (47)	14 (47)
PrpN	3 (13)	3 (11)	2 (7)	4 (13)
IndNP	1 (4)	0 (0)	0 (0)	0 (0)
Total N (%)	23 (100)	27 (100)	30 (100)	30 (101)

Comparación (L1-TC) vs (L1-NTC)

Chi-cuadrado con Overt, DefNP, PrpN $df=2$, $X^2=0.5114$, $p=0.7744$

Prueba exacta de Fisher $p=0.1726$

Comparación (L2-TC) vs (L2-NTC)

Chi-cuadrado sin (Dem, IndNP) $df=3$, $X^2=1.1507$, $p=0.7649$

Prueba exacta de Fisher $p=0.6717$

Comparación (L1-TC) vs (L2-TC)

Chi-cuadrado con Overt, DefNP, PrpN $df=2$, $X^2=0.0465$, $p=0.9770$

Prueba exacta de Fisher $p=0.5093$

Comparación (L1-NTC) vs (L2-NTC)

Chi-cuadrado sin IndNP $df=4$, $X^2=1.6000$, $p=0.8088$

Prueba exacta de Fisher $p=0.8279$

En general, no se encontraron resultados significativos entre L1 y L2, ni entre TC y NTC.

Tabla 5.3.21 – ACT: Regresión logística sobre Null, Overt, DefNP y PrpN para L1 entre TC y NTC

Forma	<i>p</i> -valor
Null	0.4272
Overt	0.8906
DefNP	0.3034
PrpN	0.4334

Tabla 5.3.22 - ACT: Regresión logística sobre Null, Overt, DefNP y PrpN para L2 entre TC y NTC

Forma	<i>p</i> -valor
Null	0.5533
Overt	0.4622
DefNP	0.3410
PrpN	0.7982

Tabla 5.3.23 - ACT: Regresión logística sobre Overt, DefNP y PrpN para TC entre L1 y L2

Forma	<i>p</i> -valor
Overt	0.7769
DefNP	0.9078
PrpN	0.8342

Tabla 5.3.24 - ACT: Regresión logística sobre Overt, Dem y PrpN para NTC entre L1 y L2

Forma	<i>p</i> -valor
Overt	0.7655
Dem	0.3853
PrpN	0.3853

No se encontraron diferencias significativas a través del análisis de regresión logística cuando el referente está ACT, para ninguna de las formas de sujeto.

Los ajustes logísticos realizados sobre Null, Overt, DefNP y PrpN no arrojaron resultados significativos para los factores Grupo y TC y por lo tanto, en general, no se encontraron diferencias significativas entre los grupos, o entre las narrativas con y sin TC, cuando el referente está en el estado ACT.

5.3.4 Estado cognitivo ACTNR

La Tabla 5.3.25 muestra la distribución de las formas de sujeto entre TC y NTC, para todos los participantes (L1 y L2), en el estado ACTNR.

Tabla 5.3.25 - ACTNR: Formas de sujeto entre TC y NTC

Frecuencia (Porcentaje)	TC	NTC
Null	3 (8)	3 (4)
Overt	0 (0)	0 (0)
Dem	1 (3)	0 (0)
DefNP	23 (62)	39 (58)
PrpN	6 (16)	17 (25)
IndNP	4 (12)	8 (12)
Total N (%)	37 (101)	67 (99)

Comparación TC vs NTC

Chi-cuadrado sin (Overt, Dem) $df=3$, $X^2=1.5319$, $p=0.6749$

Prueba exacta de Fisher $p=0.5122$

En general, no se encontraron diferencias significativas entre las narrativas TC y NTC, cuando el referente está ACTNR.

Tabla 5.3.26 - ACTNR: Regresión logística sobre Null, DefNP, PrpN y IndNP entre TC y NTC

Forma	P-valor
Null	0.4564
DefNP	0.6936
PrpN	0.2727
IndNP	0.8624

No se encontraron resultados significativos a través del análisis de regresión logística, para ninguna de las formas de sujeto, cuando el referente está ACTNR.

Tabla 5.3.27 - ACTNR: Formas de sujeto entre L1 y L2, TC y NTC

Frecuencia (Porcentaje)	L1 TC	L2 TC	L1 NTC	L2 NTC
Null	1 (8)	2 (8)	3(8)	0 (0)
Overt	0 (0)	0 (0)	0 (0)	0 (0)
Dem	1 (8)	0 (0)	0 (0)	0 (0)
DefNP	7 (58)	16 (64)	21 (55)	18 (62)
PrpN	0 (0)	6 (24)	8 (21)	9 (31)
IndNP	3 (25)	1 (4)	6 (16)	2 (7)
Total N (%)	12 (99)	25 (100)	38 (100)	29 (100)

Comparación (L1-TC) vs (L1-NTC)

Chi-cuadrado con DefNP, PrpN y IndNP $df=2$, $X^2=3.0536$, $p=0.2172$

Prueba exacta de Fisher $p=0.1931$

Comparación (L2-TC) vs (L2-NTC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=0.0214$, $p=0.8837$

Prueba exacta de Fisher $p=0.5373$

Comparación (L1-TC) vs (L2-TC)

Chi-cuadrado con DefNP, PrpN y IndNP $df=2$, $X^2=6.3926$, $p=0.0409$

Prueba exacta de Fisher $p=0.0535$

Comparación (L1-NTC) vs (L2-NTC)

Chi-cuadrado con DefNP, PrpN y IndNP $df=2$, $X^2=1.7424$, $p=0.4184$

Prueba exacta de Fisher $p=0.3295$

En general, no se encontraron resultados significativos entre L1 y L2, ni entre TC y NTC.

Tabla 5.3.28 - ACTNR: Regresión logística sobre DefNP, PrpN y IndNP para L1 entre TC y NTC

Forma	<i>p</i> -valor
DefNP	0.8516
PrpN (Firth)	0.0953
IndNP	0.4817

Tabla 5.3.29 - ACTNR: Regresión logística sobre DefNP, PrpN y IndNP para L2 entre TC y NTC

Forma	<i>p</i> -valor
DefNP	0.8835
PrpN	0.5638
IndNP	0.6392

Tabla 5.3.30 - ACTNR: Regresión logística sobre DefNP, PrpN y IndNP para TC entre L1 y L2

Forma	<i>p</i> -valor
DefNP	0.7401
PrpN (Firth)	0.0724
IndNP	0.0631

Tabla 5.3.31 - ACTNR: Regresión logística sobre DefNP, PrpN y IndNP para NTC entre L1 y L2

Forma	<i>p</i> -valor
DefNP	0.5752
PrpN	0.3538
IndNP	0.2534

No se encontraron resultados significativos a través del análisis de regresión logística cuando el referente está ACTNR, para ninguna de las formas de sujeto.

Los ajustes logísticos realizados sobre DefNP, PrpN e IndNP no arrojaron resultados significativos para los factores Grupo y TC; por lo tanto, en general, no se encontraron diferencias significativas entre los grupos, o entre las narrativas con y sin TC, cuando el referente está en el estado ACTNR.

5.3.5 Estado cognitivo FAM

La Tabla 5.3.32 muestra la distribución de las formas de sujeto entre TC y NTC, para todos los participantes (L1 y L2), en el estado FAM.

Tabla 5.3.32 - FAM: Formas de sujeto entre TC y NTC

Frecuencia (Porcentaje)	TC	NTC
Null	0 (0)	2 (3)
Overt	0 (0)	0 (0)
Dem	2 (3)	0 (0)
DefNP	49 (69)	52 (71)
PrpN	19 (27)	19 (26)
IndNP	1 (1)	0 (0)
Total N (%)	71 (100)	73 (100)

Comparación TC vs NTC

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2 < 0.0001$, $p=1.0000$

Prueba exacta de Fisher $p=0.3632$

En general, no se encontraron diferencias significativas entre las narrativas TC y NTC, cuando el referente está FAM.

Tabla 5.3.33 - FAM: Regresión logística sobre DefNP y PrpN entre TC y NTC

Forma	<i>p</i> -valor
DefNP	0.7712
PrpN	0.9205

No se encontraron resultados significativos a través del análisis de regresión logística cuando el referente está en FAM para las formas DefNP y PrpN.

Tabla 5.3.34 - FAM: Formas de sujeto entre L1 y L2, TC y NTC

Frecuencia (Porcentaje)	L1 TC	L2 TC	L1 NTC	L2 NTC
Null	0 (0)	0 (0)	0 (0)	2 (6)
Overt	0 (0)	0 (0)	0 (0)	0 (0)
Dem	1 (3)	1 (2)	0 (0)	0 (0)
DefNP	19 (61)	30 (75)	25 (68)	27 (75)
PrpN	10 (32)	9 (22)	12 (32)	7 (19)
IndNP	1 (3)	0 (0)	0 (0)	0 (0)
Total N (%)	31 (99)	40 (99)	37 (100)	36 (100)

Comparación (L1-TC) vs (L1-NTC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2 < 0.0001$, $p=1.0000$

Prueba exacta de Fisher $p=0.5501$

Comparación (L2-TC) vs (L2-NTC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2 < 0.0001$, $p=1.0000$

Prueba exacta de Fisher $p=0.5095$

Comparación (L1-TC) vs (L2-TC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=0.5828$, $p=0.4452$

Prueba exacta de Fisher $p=0.5114$

Comparación (L1-NTC) vs (L2-NTC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=0.7359$, $p=0.3910$

Prueba exacta de Fisher $p=0.1972$

En general, no se encontraron resultados significativos entre L1 y L2, ni entre TC y NTC.

Tabla 5.3.35 - FAM: Regresión logística sobre DefNP y PrpN para L1 entre TC y NTC

Forma	p -valor
DefNP	0.5898
PrpN	0.9878

Tabla 5.3.36 - FAM: Regresión logística sobre DefNP y PrpN para L2 entre TC y NTC

Forma	p -valor
DefNP	1.0000
PrpN	0.7439

Tabla 5.3.37 - FAM: Regresión logística sobre DefNP y PrpN para TC entre L1 y L2

Forma	p -valor
DefNP/PrpN	0.2163
PrpN	0.3583

Tabla 5.3.38 - FAM: Regresión logística sobre DefNP y PrpN para NTC entre L1 y L2

Forma	p -valor
DefNP	0.4825
PrpN	0.2039

No se encontraron resultados significativos a través del análisis de regresión logística, para ninguna de las formas de sujeto, cuando el referente está en FAM.

Los ajustes logísticos realizados sobre DefNP y PrpN no arrojaron resultados significativos para los factores Grupo y TC.

5.3.6 Estado cognitivo UNI

La Tabla 5.3.39 muestra la distribución de las formas de sujeto entre TC y NTC, para todos los participantes (L1 y L2), en el estado UNI.

Tabla 5.3.39 - UNI: Formas de sujeto entre TC y NTC

Frecuencia (Porcentaje)	TC	NTC
Null	1 (2)	0 (0)
Overt	0 (0)	0 (0)
Dem	2 (4)	1 (3)
DefNP	32 (71)	23 (64)
PrpN	10 (22)	11 (31)
IndNP	0 (0)	1 (3)
Total N (%)	45 (99)	36 (101)

Comparación TC vs NTC

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=0.3251$, $p=0.5685$

Prueba exacta de Fisher $p=0.6888$

En general, no se encontraron diferencias significativas entre las narrativas TC y NTC, cuando el referente está en UNI.

Tabla 5.3.40 - UNI: Regresión logística sobre DefNP y PrpN entre TC y NTC

Forma	<i>p</i> -valor
DefNP	0.4896
PrpN	0.3961

No se encontraron resultados significativos a través del análisis de regresión logística para DefNP y PrpN, cuando el referente está en UNI.

Tabla 5.3.41 - UNI: Formas de sujeto entre L1 y L2, TC y NTC

Frecuencia (Porcentaje)	L1 TC	L2 TC	L1 NTC	L2 NTC
Null	0 (0)	1 (4)	0 (0)	0 (0)
Overt	0 (0)	0 (0)	0 (0)	0 (0)
Dem	0 (0)	2 (8)	0 (0)	1 (5)
DefNP	16 (76)	16 (67)	9 (56)	14 (70)
PrpN	5 (24)	5 (21)	6 (38)	5 (25)
IndNP	0 (0)	0 (0)	1 (6)	0 (0)
Total N (%)	21 (100)	24 (100)	16 (100)	20 (100)

Comparación (L1-TC) vs (L1-NTC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=0.4526$, $p=0.5011$

Prueba exacta de Fisher $p=0.3580$

Comparación (L2-TC) vs (L2-NTC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2<0.0001$, $p=1.0000$

Prueba exacta de Fisher $p=1.0000$

Comparación (L1-TC) vs (L2-TC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2<0.0001$, $p=1.0000$

Prueba exacta de Fisher $p=0.6791$

Comparación (L1-NTC) vs (L2-NTC)

Chi-cuadrado con DefNP y PrpN $df=1$, $X^2=0.2282$, $p=0.6328$

Prueba exacta de Fisher $p=0.5144$

En general, no se encontraron resultados significativos entre L1 y L2, ni entre TC y NTC.

Tabla 5.3.42 - UNI: Regresión logística sobre DefNP y PrpN para L1 entre TC y NTC

Forma	<i>p</i> -valor
DefNP	0.1998
PrpN	0.3680

Tabla 5.3.43 - UNI: Regresión logística sobre DefNP y PrpN para L2 entre TC y NTC

Forma	<i>p</i> -valor
DefNP	0.8130
PrpN	0.7430

Tabla 5.3.44 - UNI: Regresión logística sobre DefNP y PrpN para TC entre L1 y L2

Forma	<i>p</i> -valor
DefNP	0.4802
PrpN	0.8108

Tabla 5.3.45 - UNI: Regresión logística sobre DefNP y PrpN para NTC entre L1 y L2

Forma	<i>p</i> -valor
DefNP	0.3938
PrpN	0.4192

No se encontraron resultados significativos a través del análisis de regresión logística cuando el referente está en UNI, para ninguna de las formas de sujeto.

Los ajustes logísticos realizados sobre DefNP y PrpN no arrojaron resultados significativos para los factores Grupo y TC.

5.3.7 Estado cognitivo TYP

La Tabla 5.3.46 muestra la distribución de las formas de sujeto entre TC y NTC, para todos los participantes (L1 y L2), en el estado TYP.

Tabla 5.3.46 - TYP: Formas de sujeto entre TC y NTC

Frecuencia (Porcentaje)	TC	NTC
Null	0 (0)	0 (0)
Overt	0 (0)	0 (0)
Dem	0 (0)	1 (2)
DefNP	3 (7)	2 (4)
PrpN	0 (0)	0 (0)
IndNP	40 (93)	45 (94)
Total N (%)	43 (100)	48 (100)

Comparación TC vs NTC

Chi-cuadrado con DefNP y IndNP $df=1$, $X^2=0.0105$, $p=0.9185$

Prueba exacta de Fisher $p=0.8243$

En general, no se encontraron diferencias significativas entre las narrativas TC y NTC, cuando el referente está en TYP.

Tabla 5.3.47 - TYP: Regresión logística sobre DefNP y IndNP entre TC y NTC

Forma	P-valor
DefNP	0.5567
IndNP	0.8892

No se encontraron resultados significativos a través del análisis de regresión logística para DefNP y PrpN, cuando el referente está en TYP.

Tabla 5.3.48 - TYP: Formas de sujeto entre L1 y L2, TC y NTC

Frecuencia (Porcentaje)	L1 TC	L2 TC	L1 NTC	L2 NTC
Null	0 (0)	0 (0)	0 (0)	0 (0)
Overt	0 (0)	0 (0)	0 (0)	0 (0)
Dem	0 (0)	0 (0)	1 (3)	0 (0)
DefNP	3 (13)	0 (0)	2 (7)	0 (0)
PrpN	0 (0)	0 (0)	0 (0)	0 (0)
IndNP	20 (87)	20 (100)	26 (90)	19 (100)
Total N (%)	23 (100)	20 (100)	29 (100)	19 (100)

Comparación (L1-TC) vs (L1-NTC)

Chi-cuadrado con DefNP y IndNP $df=1$, $X^2=0.0538$, $p=0.8166$

Prueba exacta de Fisher $p=0.8032$

Comparación (L2-TC) vs (L2-NTC)

Separación completa

Comparación (L1-TC) vs (L2-TC)

Chi-cuadrado con DefNP y IndNP $df=1$, $X^2=1.1547$, $p=0.2826$

Prueba exacta de Fisher $p=0.2359$

Comparación (L1-NTC) vs (L2-NTC)

Chi-cuadrado con DefNP y IndNP $df=1$, $X^2=0.2064$, $p=0.6496$

Prueba exacta de Fisher $p=0.7027$

En general, no se encontraron resultados significativos entre L1 y L2, ni entre TC y NTC.

Tabla 5.3.49 - TYP: Regresión logística sobre DefNP y IndNP para L1 entre TC y NTC

Forma	<i>p</i> -valor
DefNP	0.4567
IndNP	0.7629

Tabla 5.3.50 - TYP: Regresión logística sobre DefNP y IndNP para TC entre L1 y L2

Forma	<i>p</i> -valor
DefNP (Firth)	0.1268
IndNP (Firth)	0.9462

Tabla 5.3.51 - TYP: Regresión logística sobre DefNP y IndNP para NTC entre L1 y L2

Forma	<i>p</i> -valor
DefNP (Firth)	0.3680
IndNP (Firth)	0.8382

No se encontraron resultados significativos a través del análisis de regresión logística, para ninguna de las formas de sujeto, cuando el referente está en TYP.

Tabla 5.3.52 - TYP: Ajuste logístico sobre IndNP

Término	Estimación	<i>p</i> -valor
L2	2.4018	0.0277

El ajuste logístico sobre IndNP indica que el factor Grupo tiene un efecto sobre esta forma de sujeto cuando el referente está en TYP: el grupo L2 tiene una mayor tendencia a utilizar IndNP tanto en las narrativas con y sin TC.

CAPÍTULO 6

6 DISCUSIÓN Y ANÁLISIS CUALITATIVO

6.1 Introducción

A través de los resultados cuantitativos, hemos podido establecer que tanto el factor Grupo como el factor (medio de) Producción tienen un efecto en la selección de las expresiones referenciales de sujeto en los estados cognitivos (ECs) de la jerarquía cognitiva. A través de los diferentes modelos estadísticos utilizados, encontramos evidencia significativa acerca de las diferencias de uso entre los datos de las narrativas orales de Blackwell y Lubbers Quesada (ByLQ) (2012) y los datos de las narrativas escritas del presente estudio, así como entre los hablantes nativos (HNs) y los hablantes no-nativos (HNNs) del presente estudio. Por otro lado, los resultados concernientes a la variable terreno común (TC) de la actual investigación no demuestran un efecto significativo y son pocos los resultados estadísticos con evidencia de que este factor haya sido aislado exitosamente o de que se haya tenido en cuenta en las composiciones escritas de los participantes. Sin embargo, a través del análisis cualitativo llevado a cabo, hay evidencia de que el factor TC hace parte integral de la GH, pero no de la manera esperada en esta investigación. Por ejemplo, los HNs utilizan otras estrategias para establecer el TC en sus narrativas, como lo es una mención explícita del hecho de que el amigo ficticio ha visto la película, en una especie de pre-introducción en las composiciones.

En las siguientes siete subsecciones, que corresponden con cada uno de los EC de la jerarquía cognitiva, se van a discutir en más detalle los resultados obtenidos en el

estudio estadístico, integrando un análisis cualitativo de los mismos y observando ejemplos extraídos de los datos del estudio escrito. Para cada uno de los estados comenzamos describiendo las diferencias entre los HNs y HNNs en sus composiciones escritas; en segundo lugar, revisamos las hipótesis de ByLQ en su estudio oral y pasamos a establecer si estas se cumplen de igual manera en el presente estudio escrito; luego, analizamos cómo se comparan las diferencias del actual estudio escrito con aquellas del estudio oral de ByLQ entre los participantes de ambos estudios; en la tercera sección de cada estado se analizan los resultados concernientes al factor TC, entre las narrativas escritas de los HNs y los HNNs.

El análisis cualitativo se presenta teniendo en cuenta las hipótesis planteadas en este estudio y que fueron introducidas en el Capítulo 4. A continuación, se repiten dichas hipótesis para facilitar su revisión:

- Hipótesis 1: Debido a que la GH es universal, tanto en las narrativas orales como en las escritas, los participantes seguirán la GH, con mayor adherencia a los estados extremos de INF y TYP; es decir, los participantes seleccionarán un sujeto nulo para un referente con EC INF y una FN indefinida para un sujeto con EC TYP.
- Hipótesis 2: Se verá un acercamiento mayor al HN por parte de los HNNs, ya que en el discurso escrito los hablantes tienen más oportunidad de pensar en sus respuestas y corregirlas; más específicamente, los HNNs seleccionarán las formas de sujeto de una manera más semejante a las selecciones de los HNs en las narrativas escritas que en las narrativas orales de ByLQ.

- Hipótesis 3: Al ser el discurso escrito más cuidado que el discurso oral, y al no poder verificar el nivel de atención y memoria, los participantes, tanto HNs como HNNs, usarán expresiones referenciales más elaboradas en las narrativas escritas que en las orales.
- Hipótesis 4: Es posible aislar el TC como factor en la selección de expresiones referenciales y las suposiciones de los participantes con respecto al nivel de atención y memoria se verán afectados. Más específicamente, al introducir los personajes en el discurso que suponen previa familiaridad con el referente, el EC TYP corresponderá mayormente con expresiones definidas en las narrativas con un TC establecido que en las narrativas sin un TC determinado.
- Hipótesis 5: El factor TC desempeñará un papel más importante en las narrativas escritas de los HNs que de los HNNs. Se observará una codificación más notable del TC por parte de los HNs a través de un mayor uso de expresiones definidas que los HNNs en las narrativas con TC, y en un mayor uso de cláusulas con información de fondo en narrativas sin TC (p. ej. “El/un hombre con bastón”), según el personaje esté siendo introducido o reintroducido en la narrativa.

6.2 Análisis y discusión INF (GH pronostica NULL)

6.2.1 Datos escritos HNs vs HNNs

La prueba de chi-cuadrado sin demostrativos ni expresiones nominales indefinidas revela que hay diferencias al borde de la significancia entre las composiciones

escritas de los HNs y los HNNs ($p=0.0503$). La prueba exacta de Fisher muestra diferencias significativas ($p=0.0722$). A través de los modelos de regresión logística, observamos que estas diferencias provienen de los pronombres explícitos (Overt, $p=0.0481$) y de las FN definidas (DefNP, $p=0.0576$). En las narrativas escritas del presente estudio, los HNs utilizan menos formas mínimas y un mayor número de expresiones elaboradas que los HNNs. En los ejemplos (1) y (2) vemos que, en la misma escena, el HN selecciona la FN definida *el hombre* en L4 para contrastar con la joven del enunciado anterior; mientras que el HNN utiliza un sujeto nulo que, aunque es adecuado al ser el tópico de la oración, vemos también que el HNN recurre a un uso erróneo del objeto directo *la* en *comprarla* para hacer el mismo contraste:

(1) Ejemplo [P2.L1.TC]

L3. El padre_i se marcha a ligar con una joven que está flirteando con él.
L4. *El hombre*_i va a comprar unos refrescos para ambos y

(2) Ejemplo [P31.L2.NTC]

L5. El esposo_i espía una mujer soltera y
L6. [\emptyset]_i la siguió para coquetear con ella.
L7. [\emptyset]_i Sale para comprarla una bebida,

A lo largo de esta sección veremos que el foco contrastivo, como aquel del ejemplo (1), más arriba, es un factor pragmático que es usado con frecuencia en las narrativas escritas en los dos grupos de hablantes.

6.2.2 Comparación entre datos escritos y datos orales

ByLQ, en su estudio de 2012, habían propuesto la hipótesis para el estado INF que los hablantes usarían sujetos nulos para referirse a referentes que se encontraran en el inmediato centro de atención y memoria, por lo tanto, la entidad más prominente y el tópico de la oración (p. 154). Los resultados de su estudio arrojan datos que muestran que

esta hipótesis se cumple. Los datos escritos del presente estudio también apoyan esta hipótesis y, al igual que ByLQ, la distribución es muy clara en este estado, aunque la adherencia a la GH es menos fuerte en las narrativas escritas ya que también vemos que la escritura tiende a aumentar el uso de las formas de sujeto más elaboradas: expresiones nominales definidas y nombres propios.

Al comparar los medios de producción, oral y escrito, los sujetos nulos siguen siendo la forma más utilizada en las narrativas escritas, lo cual es esperado cuando el referente está en el centro de atención. Esto confirma la hipótesis uno de este estudio, la cual predice que los participantes seguirán la jerarquía cognitiva, con mayor adherencia a los estados INF y TYP (para TYP, ver Sección 6.8). Sin embargo, en las narrativas escritas se hace un menor uso de formas más mínimas, como lo es el uso de sujetos nulos (en los HNs) y de pronombres explícitos (tanto en HNs como en HNNs). Contrariamente, el uso de sujetos nulos es mayor en los HNNs y esto hace que el efecto del factor Grupo, en este caso el ser no-nativo, sea positivo; es decir, hay un acercamiento al uso nativo en las narrativas escritas, como predicho por la hipótesis dos:

- Uso de Null en HNs: 91% oral vs 81% escrito ($p < 0.0010$)
- Uso de Null en HNNs: 65% oral vs 80% escrito ($p < 0.0005$)

En el siguiente ejemplo (3) de un participante HN, Chaplin (a quien también se le conoce por el nombre de Charlot) es el tópico de la oración en la línea 31 (L31). La continuidad de tópico se mantiene a través de un sujeto nulo en L32 y L33:

(3) Ejemplo [P4.L1.TC]

L31 Charlot conduce al hombre cerca del lago,
L32. [\emptyset] mide el nivel del agua

L33. y [\emptyset] acaba²¹ echándolo.

Tanto los HNs como los HNNs hacen un mayor uso de expresiones nominales definidas en las narrativas escritas, aunque este cambio es solo estadísticamente significativo en los HNs:

- Uso de DefNP en HNs: 2% oral vs 11% escrito ($p < 0.0001$)
- Uso de DefNP en HNNs: 3% oral vs 6% escrito ($p = 0.2445$)

Este es un resultado esperado en la hipótesis tres del vigente estudio, donde se predice el uso de expresiones referenciales más elaboradas que en el discurso oral.

El uso de nombres propios incluye pocas instancias en el estado INF, pero estas también ocurren en mayor proporción en las narrativas escritas. La diferencia es significativa para los HNs:

- Uso de PrpN en HNs: 1% oral vs 3% escrito ($p = 0.0388$)
- Uso de PrpN en HNNs: 4% oral vs 5% escrito ($p = 0.6855$)

El uso de una expresión más elaborada no es en sí un fenómeno que vaya en contra de la GH. Uno de los principios de esta jerarquía es lo que Gundel (2010:155) denomina *one-to-many mapping*. Según este principio, al ser esta una jerarquía unidireccional e implicativa, es igualmente apropiado utilizar una forma relacionada con uno de los ECs más altos en la jerarquía, ya que estos estados están contenidos en las formas más bajas. En otras palabras, un referente DefNP que se encuentra en el foco de atención al ser el tópico del enunciado, también está activado, es familiar, es identificable de modo único y tiene identificación de tipo, por lo tanto, cualquier forma que el hablante

²¹ En todos los ejemplos de este trabajo se mantiene fielmente la producción de los participantes, ya sea agramatical o con errores de tipografía, ya que en muchos casos demuestran que los participantes escribieron de manera espontánea, sin planear o revisar el texto.

utilice va a cumplir con los requerimientos mínimos del EC en cuestión. Por ejemplo, la frase definida *el hombre* puede usarse adecuadamente para un referente en foco. Sin embargo, y al igual que en el estudio oral (p. 154), cuando los hablantes optan por una forma más elaborada para referentes en foco (una FN definida o un nombre propio) es generalmente debido a factores pragmáticos. Estos factores incluyen el foco contrastivo o la reiteración del tópico debido a otro tipo de énfasis, por ejemplo, al comenzar una nueva escena o acción. Las principales diferencias con el estudio oral de ByLQ son el mayor número de estos casos en las narrativas escritas, y el hecho de que se encontró evidencia de estos factores pragmáticos también en los HNNs y no solamente en los HNs, como fue el caso en el estudio oral. En los ejemplos (4) y (5), más abajo, vemos que los sujetos explícitos se utilizan cuando un personaje en foco comienza una nueva escena. En (4), el HN utiliza la FN definida *el padre* en L3 para llevar al foco de atención al referente, el cual es identificable de modo único a través de la previa mención de la familia en L2. En L4 este referente está en foco pero el HN utiliza la expresión definida *el hombre* cuando un sujeto nulo hubiera sido suficiente para su identificación; al utilizar una expresión más elaborada, el escritor indica el inicio de una nueva escena. De hecho, en esta escena el personaje ya no se encuentra con la joven sino que está caminando para buscar las bebidas; por lo tanto, el uso de la expresión definida *el hombre* puede indicar bien sea el cambio de escena o un cambio en la acción para el referente en foco. Cinco de los 10 participantes nativos del presente estudio narraron este mismo cambio en la escena del corto con una FN definida o con un pronombre explícito.

(4) Ejemplo [P2.L1.TC]

L2. mientras la madre y la hija echan la siesta

L3. el padre se marcha a ligar con una joven que está flirteando con él.

L4. *El hombre* va a comprar unos refrescos para ambos y

En el ejemplo (5), más abajo, el HNN repite el nombre propio (*Charlie*) al indicar la nueva escena donde Chaplin ve al hombre y a la mujer en L14. Puede ser que este hablante haya adquirido el sujeto nulo como lo vemos a través de su uso anterior en L12 y L13, como también es posible que el uso del sujeto nulo se deba a la influencia de la L1, puesto que en el inglés es viable utilizar el sujeto nulo en construcciones correferenciales coordinadas (mismo sujeto, mismo tiempo verbal) con la conjunción *and*, como lo demostraron Torres Cacoullós y Travis (2014), por ejemplo, “*I* just put that mayonnaise on top and \emptyset_i put in on some boiled eggs” (p. 20):

(5) Ejemplo [P24.L2.TC]

L11. Ahora vemos que *Charlie Chaplin* viene caminando por el parque y
L12. sin darse cuenta [\emptyset] camina arriba de un aspersor y
L13. [\emptyset] se moja todo.
L14. *Charlie* ve al hombre con la mujer misteriosa sentados en otro banco del parque y

Es de notar que, al repetir el nombre de Chaplin en L14, el HNN utiliza solamente el primer nombre y no el apellido (utilizado en L11), lo cual le brinda más cohesión al texto. Esto tiene sentido también desde el punto de vista de la accesibilidad: Ariel (1999:40) nos dice que los nombres completos (nombre y apellido), se encuentran en el extremo más bajo de la escala de los marcadores de baja accesibilidad, seguidos por el solo apellido y luego por el primer nombre, que, según la autora, señalan una accesibilidad relativamente alta.

Hay varios ejemplos de uso pragmático de foco contrastivo, tanto en HNs como en HNNs. Generalmente, en estos casos el sujeto explícito es obligatorio (Matos Amaral y Schwenter, 2005) ya que el hablante busca hacer más prominente al referente en foco

en una situación de posible ambigüedad, ya sea porque hay más de un referente en escena, por el contexto o por morfología verbal ambigua (por ejemplo, una forma verbal que no identifica al referente a través del número o el género).

En el ejemplo (6) que sigue vemos un caso pragmático de foco contrastivo donde el HN activa en L8 al personaje que ha sido previamente mencionado en la narrativa, *el hombre*, y este referente se mantiene en foco en L9 y L10:

(6) Ejemplo [P1.L1.TC]

- L8. Luego que *el hombre*_i se encuentra con la dama,
- L9. *el*_i le ofrece una bebida.
- L10. Mientras que *el hombre*_i va en busca de la bebida,
- L11. otro hombre pasa

Sin embargo, en L9 el escritor utiliza un pronombre explícito para este referente y una FN definida en L10. En L9 el contraste es innecesario porque con un pronombre nulo se puede acceder al referente *el hombre* en L8 sin ningún esfuerzo cognitivo, como lo predice la hipótesis de la posición del antecedente de Carminati (2002) (discutida en el Capítulo 2), al encontrarse el referente en posición de sujeto en la oración anterior. Sin embargo, debido a que *la dama* ha sido activada en posición de objeto en L8, el uso del pronombre explícito puede servir para aclarar que es el hombre quien ofrece la bebida a la dama y no viceversa, resaltando de esta manera el contraste entre los dos referentes. En L10, el escritor utiliza la FN definida *el hombre* para indicar un foco contrastivo específico entre las acciones del referente en foco y aquellas de un nuevo personaje que es introducido en L11. Por lo tanto, aunque un pronombre nulo en L10 puede acceder de manera adecuada al referente en foco, *el hombre*, el uso de la FN definida establece un contraste específico:

En el ejemplo (7) vemos un caso de foco contrastivo de un HNN. Charlie es el tópico del discurso desde L28 hasta L31, pero, al haber dos hombres más en la escena, el HNN opta por el nombre propio en L30. Gramaticalmente el uso de un sujeto explícito no es necesario en este caso porque el pronombre de objeto directo *los* ayuda a seleccionar correctamente al referente:

(7) Ejemplo [P24.L2.TC]

- L26. *Charlie* empieza a molestar a otra gente, dos hombres,
- L27. Y cuando uno de ellos se enoja
- L28. *Charlie* le saca la botella de gaseosa que estaba tomando y
- L29. [\emptyset] se la revienta en la cabeza.
- L30. Cuando *Charlie* los deja,
- L31. [\emptyset] se encuentra otra vez con el primer hombre,

También vemos que el uso de ciertos adverbios sirve para enfatizar el contraste pragmático entre dos escenas o acciones. En el ejemplo (8), el HN utiliza un nombre propio en L53 para el referente en foco, después de haber utilizado un sujeto nulo en la línea anterior; esto se debe a que el adverbio *mientras* resalta el contraste con la acción anterior:

(8) Ejemplo [P19.L1.NTC]

- L51. y en ese momento *chaplin*_i le da un golpe y
- L52. [\emptyset]_i empuja al hombre quien cae en el lago,
- L53. mientras *chaplin*_i le mira caer
- L54. [\emptyset]_i no se percata que detrás de él hay un policía

Esto también ocurre con los HNNs, como lo vemos con el uso del pronombre personal *él* en L15 del ejemplo (9), seguido del mismo adverbio, *mientras*:

(9) Ejemplo [P22.L2.TC]

- L13. Después de un rato juntos, *el hombre*_i le dijo a la coqueta
- L14. que [\emptyset]_i le va a comprar una bebida.
- L15. Mientras *él*_i le recoge la botella fría en el quiosco, vemos a Charlie Chaplin, el héroe.

Hay algunos casos en que la selección de una forma más elaborada no parece tener ninguna razón pragmática ni sintáctica. Sin embargo, pueden seguirse considerando adecuados, tal vez porque el estilo escrito permite una mayor elaboración de las formas sin que estas resulten redundantes. No se encontraron este tipo de ejemplos en los textos de los HNNs. En el ejemplo (10) se utiliza una frase léxica en L4, *el caballero*, después de haberse referido a este como *el hombre* en L3. La descripción más semánticamente elaborada podría atribuirse al efecto estilístico de la escritura, al ser este tipo de expresión poco común en el discurso oral; no obstante, una discusión de motivos estilísticos no forma parte del alcance del presente estudio. Así que de aquí en adelante nos limitaremos a enfocar solamente las causas pragmáticas que ayudan a construir la cohesión discursiva y en los ECs que motivan la selección de formas de sujeto:

(10) Ejemplo [P3.L1.TC]

- L2. Aprovechando que las mujeres se quedan dormidas,
- L3. *el hombre* las deja para hablar con otra mujer que pasaba por ahí.
- L4. *El caballero* quiere invitarla.

En las narrativas escritas del presente estudio hay casos de reiteración de tópico relacionado con el referente inanimado *la película*, donde un sujeto nulo hubiera sido adecuado. En los datos se encontraron tres de estos ejemplos entre los HN, ninguno con los HNNs. También hay tres de estos casos en los datos orales con los HNs, ninguno con los HNNs. A continuación, se presenta una instancia representativa extraída de las narrativas escritas del presente estudio, donde el referente es llevado al foco de atención en L1, pero sus dos subsecuentes menciones en L2 y L3 son con una FN definida en vez de un sujeto nulo, como hubiera sido lo esperado:

(11) Ejemplo [P7.L1.TC]

L1. Acabo de ver *una película* de de Charles Chaplin de los años 20 o 30 en blanco y negro.

L2. Además de ser en blanco y negro, *la película* también era en silencio con un fondo de música de piano.

L3. *La película* estuvo compuesta de varias escenas cortas con diferentes títulos.

Debemos tener en cuenta las siguientes dos consideraciones: en primer lugar, para una entidad inanimada las únicas formas referenciales posibles son el sujeto nulo, una frase léxica (*la/una película*) o un artículo o pronombre demostrativo (*esta película/ésta*); en español no podemos utilizar un pronombre personal para hacer referencia a un sujeto inanimado (p. ej. *#él/ella*). En segundo lugar, la repetición de la expresión referencial *la/una película* puede deberse al efecto de la perseveración (*priming*), definido como la repetición de la forma anterior por parte del hablante (Torres Cacoullos y Travis, 2014:24), el cual es un fenómeno psicolingüístico que ocurre en la producción de varias lenguas (Abreu, 2012:1) y, aunque ocurre también en las narrativas escritas, se ha encontrado mayor evidencia en las narrativas orales (Tannen, 1987:226). En este estudio no analizamos los efectos de la perseveración, pero se tiene en cuenta como posible explicación para algunas de las ocurrencias como aquellas representadas por el ejemplo (11) anterior.

Se encontraron algunas instancias entre los HNNs donde el uso de expresiones explícitas en el estado INF puede deberse a una adquisición parcial del sujeto nulo, causando una ruptura en la cohesión de la narrativa. Como lo indica Lubbers Quesada (2015:36), la noción cognitivo-discursiva de la cohesión se ve afectada por la selección de la forma referencial: el uso de un sujeto nulo en español contribuye a la cohesión de la narrativa, mientras que el uso repetitivo de expresiones explícitas causa una ruptura en la

cohesión debido a las dificultades de procesamiento que le implican al interlocutor. En el ejemplo (12) vemos un sobreuso de la FN definida *el hombre*, y en el ejemplo (13) un sobreuso del pronombre *ella*. Aunque los textos son coherentes, esta repetición de las expresiones explícitas en L23 del ejemplo (12) y en L21 del ejemplo (13) es lo que causa la ruptura en la cohesión narrativa:

(12) Ejemplo [P26.L2.NTC]

L22. *el hombre* se va con la mujer.

L23. *El hombre* le dice que quiere jugar al escondite

(13) Ejemplo [P32.L2.NTC]

L20. *Ella* le vendó los ojos al hombre y

L21. *ella* se fue.

6.2.3 Comparación narrativas escritas TC y NTC

El análisis del factor TC no revela resultados significativos sobre la codificación de esta variable a través del sujeto nulo. De hecho, en las narrativas escritas hay un uso proporcional casi idéntico (78% TC vs 82% NTC, Tabla 5.3.1). Aunque la diferencia no es significativa ($p=0.3173$), vemos que, sorpresivamente, hay un mayor uso de sujetos nulos en las narrativas NTC y este fenómeno sucede en ambos grupos de hablantes.

Además, tampoco se encontraron diferencias entre los HNs y HNNs sobre esta forma.

Desde el punto de vista del factor TC, el anterior ejemplo (1) (Sección 6.2.1) es un caso destacable ya que se trata de un HN en cuya narrativa de 208 palabras no utiliza en ningún momento sujetos nulos, a pesar de asumirse TC entre escritor y lector. Esto puede deberse a un posible conflicto entre los factores TC y Producción: al ser una narrativa donde se asume existencia de TC, esperaríamos menos expresiones elaboradas y más expresiones mínimas. Pero, hemos visto que el factor Producción es significativo en

el aumento de expresiones definidas en los HNs; por lo tanto, debemos concluir en este caso que el factor Producción es más fuerte que el factor TC. También debemos considerar una posible dificultad de aislar el TC en las narrativas escritas, lo cual contradice las hipótesis cuatro y cinco del presente estudio, donde se predice la posibilidad de aislar el TC y la evidencia de codificación del mismo por parte de los HNs en las narrativas escritas.

No podemos atribuirle la explicación de un mayor número de formas elaboradas en este estado al factor TC debido a que en los resultados cuantitativos no hay evidencia de este entre las narrativas TC y NTC ($p=0.6013$). Sin embargo, es de notar que el ejemplo (7) en la sección anterior (6.2.2) es una narrativa donde se asume TC, por lo tanto, de haberse usado un sujeto nulo en L30, se hubiera podido recurrir a este para la identificación del referente. El hecho de que el hablante no lo haya hecho, de nuevo es evidencia en contra de las hipótesis cuatro y cinco de este estudio con respecto a la codificación del TC en la narrativa y a favor de la escritura como factor que motiva un mayor uso de expresiones definidas, como planteado en la hipótesis tres.

6.2.4 Conclusiones estado INF

En conclusión, los datos muestran que tanto los HNs como los HNNs se adhieren a la GH al utilizar el sujeto nulo en el estado INF, donde el referente se encuentra en el centro de atención y es el tópico de la oración. Sin embargo, al comparar con las narrativas orales, esta adherencia es menos fuerte en las narrativas escritas, ya que hay evidencia de un significativo uso más elevado de expresiones más elaboradas en estas últimas, como lo son las expresiones nominales definidas y los nombres propios. Esta tendencia es mayor en los HNs. Este es un resultado esperado, como planteado por las

hipótesis uno y tres de este estudio, donde se predice una clara adherencia a la jerarquía para el EC INF, al mismo tiempo que un uso de expresiones referenciales más elaboradas al no poder verificar los niveles de atención y memoria del lector.

También vemos un posible efecto de perseveración de las formas referenciales, concretamente en la repetición del referente inanimado *la película*, en tres de las narrativas de los HNs. Dicha repetición puede deberse a que a través de este fenómeno lingüístico, el cual es automático e inconsciente, se facilita el procesamiento cognitivo, así como al hecho de que el referente *la película* es inanimado y son pocas las opciones posibles en español para hacer referencia a este tipo de entidades.

Por otra parte, no hay evidencia de que el factor TC afecte la selección de los sujetos en las narrativas escritas, lo cual vemos a través de la casi igual proporción de formas entre las narrativas TC y NTC y por la distribución contraria a lo esperado de las formas de sujeto nulo y de las formas más elaboradas. El aumento de sujetos nulos proviene de ambos grupos de hablantes, mientras que las expresiones léxicas muestran un comportamiento contrario entre los dos grupos: los HNs disminuyen el uso de expresiones definidas y aumentan el uso de nombres propios, mientras que los HNNs aumentan las expresiones definidas y disminuyen el uso de nombres propios. Hay que añadir que el único resultado significativo para el factor TC en el estado INF es entre HNs y HNNs en narrativas con TC, donde vemos que los HNNs tienden a utilizar más la forma nula (1% L1 vs 8% L2). Sin embargo, los participantes son diferentes en ambos grupos y es posible que el conocimiento del personaje de Chaplin no se encuentre al mismo nivel entre ellos y que esto afecte el uso de nombres propios. De todas maneras, estos resultados son sorprendentes, puesto que esperaríamos que se incrementaran todas

las formas de expresiones más elaboradas en las narrativas sin TC.

6.3 Análisis y discusión ACTR (GH pronostica NULL)

6.3.1 Datos escritos HNs vs HNNs

A través de la prueba de chi-cuadrado sin demostrativos ni expresiones nominales indefinidas no se encontraron diferencias significativas entre los HNs y HNNs en las narrativas escritas ($p=0.2205$). Debido a que, proporcionalmente, los resultados son casi idénticos entre los dos grupos para los sujetos nulos y las expresiones definidas, solo se realizó una regresión logística para los pronombres explícitos y el nombre propio.

Tampoco a través de este modelo se encontraron diferencias significativas (Tabla 5.2.14).

6.3.2 Comparación datos escritos vs datos orales

Aunque el referente no es el tópico de la oración ni es el foco de la oración, en el estudio oral de ByLQ se predice el uso de un sujeto nulo en el estado activado y recuperable, ya que este es completamente recuperable a través del contexto y de la semántica de los verbos (p. 155). La hipótesis se cumple en ambos estudios, sin embargo, en línea con los resultados de las narrativas escritas en el EC INF, los resultados revelan una menor frecuencia de formas más mínimas y una mayor frecuencia de aquellas más elaboradas, en comparación con las narrativas orales.

En el estudio oral de ByLQ, los HNs hacen un uso significativamente mayor de sujetos nulos que los HNNs (82% HNs vs 59% HNNs), y las autoras concluyen que esto se debe a que los HNs, en el estado activado y recuperable, no necesitan especificar un referente activado de manera explícita cuando este es identificable por el contexto, y que esto es algo que los estudiantes deben aprender (p. 156). Sin embargo, en el estudio

escrito esta división ya no es tan clara: tanto los HNs como los HNNs hacen un menor uso de sujetos nulos en las narrativas escritas (de manera significativa para los HNs), y esto resulta en un acercamiento al uso nativo en las narrativas escritas, prácticamente eliminando la diferencia entre los dos grupos de hablantes, lo cual confirma la hipótesis dos de este estudio, en cuanto al acercamiento al habla nativa por parte de los HNNs:

- Uso de Null HNs: 82% oral vs 55% escrito ($p=0.0007$)
- Uso de Null HNNs: 59% oral vs 54% escrito ($p=0.6254$)

El uso del pronombre explícito también ocurre con menor frecuencia en las narrativas escritas y esta diferencia es significativa en los HNNs. En el estudio oral, esta forma de expresión referencial ocupa el segundo lugar para ambos grupos en este estado, mientras que en nuestros datos escritos ocupa el cuarto lugar en los HNs y el tercero en los HNNs, confirmando que el EC de activado y recuperable desfavorece el uso de formas más mínimas en la escritura:

- Uso de Overt HNs: 8% oral vs 5% escrito ($p=0.4012$)
- Uso de Overt HNNs: 34% oral vs 15% escrito ($p=0.0213$)

En las narrativas escritas se observa un mayor uso de las formas más elaboradas en los dos grupos de hablantes, como predicho por la hipótesis tres del presente estudio. Las FNs definidas pasan a ocupar el segundo lugar en ambos grupos y la diferencia con respecto a las narrativas orales es significativa. En ninguno de los dos estudios se encontraron diferencias significativas entre los dos grupos de hablantes en el uso de FNs definidas, ya que proporcionalmente su uso es idéntico en el estudio oral y casi idéntico en el escrito. En el estudio de ByLQ no se aplicó la regresión logística debido a un número muy bajo de instancias para FNs definidas en el estado activado y recuperable:

- Uso de DefNP en HNs: 2% oral vs 20% escrito ($p=0.0003$)
- Uso de DefNP en HNNs: 2% oral vs 18% escrito ($p=0.0080$)

La diferencia en el uso de nombres propios entre los participantes de ByLQ y los del actual estudio es significativa en los HNs. Entre HNs y HNNs, en ninguno de los dos estudios se encontró una diferencia significativa en el uso de nombres propios para este EC:

- Uso de PrpN en HNs: 1% oral vs 16% escrito ($p=0.0006$)
- Uso de PrpN en HNNs: 5% oral vs 7% escrito ($p=0.5896$)

Los nombres propios se usan exclusivamente para el personaje de Charlie Chaplin, ya que es el único con nombre en el corto. Aunque nos encontramos aquí con un resultado significativo en el uso de PrpN, cabe mencionar un posible problema metodológico que puede afectar la correcta interpretación de los resultados para esta forma: al haberse llevado a cabo ambos estudios con participantes diferentes, y aunque el personaje de Chaplin es mundialmente reconocido (lo cual es una razón para la selección de este corto), no podemos afirmar con absoluta certeza que ambos grupos de participantes tengan el mismo nivel de conocimiento acerca de él. Es por este motivo que, aunque en algunas instancias los resultados arrojen valores significativos a lo largo de este estudio, no debemos considerarlos conclusivos.

En la Sección 5.2.3 del Capítulo 5 vimos que la investigación general a través de los ajustes logísticos confirma que la escritura incrementa la tendencia a utilizar FNs definidas (Tabla 5.2.20) y nombres propios (Tabla 5.2.21) de manera significativa tanto en HNs como en HNNs, y también en este estado hay varios casos de factores pragmáticos que motivan esta selección, tanto en los HNs como en los HNNs. El

siguiente es un ejemplo de un HN donde vemos que el participante, en un principio, describe y comenta sobre la escena que activa a Chaplin como referente (pero es la escena de la película que es el foco de atención) en L25; luego, al volver a la descripción de las acciones, repite el nombre de Chaplin en L26 aunque, como el único referente en la escena en foco, es recuperable con un pronombre nulo:

(14) Ejemplo [P4.L1.TC]

L25. La escena de la postura de Charlot en el suelo con los pies en el banco y su balanceo es muy buena.

L26. *Charlot* se repone y

En el ejemplo (15), el HNN enumera las acciones en vez de enlazar las cláusulas con oraciones subordinadas o relativas, lo cual conlleva a la repetición explícita de los sujetos. Esto puede deberse a una falta de adquisición lingüística o al estilo espontáneo de las narrativas del estudio. En este ejemplo, ambos referentes, *el hombre* en L20 y *Charlie* en L21, están en el estado ACTR:

(15) Ejemplo [P23.L2.TC]

L20. *El hombre* usa una botella para pegar a Charlie.

L21. *Charlie* se desmaya y

Los siguientes ejemplos demuestran cómo el factor Producción puede estar ligado bien sea a factores cognitivos de atención y memoria, así como a factores pragmáticos. En el ejemplo (16) de un HN, Charlie es el tópico en L29 y en L30, pero en L31 interviene otro personaje, *un policía*, el cual lleva a Charlie fuera del foco de atención. Sin embargo, cuando Charlie es vuelto a introducir en L32, este sigue siendo recuperable con un pronombre nulo a través del contexto; aún así, el escritor utiliza el nombre propio *Charlot* para crear contraste con *un policía*:

(16) Ejemplo [P5.L1.TC]

L29. Entonces con la ayuda de su bastón [\emptyset]_i le acerca hasta un pequeño lago y
L30. [\emptyset]_i busca el lugar ideal para tirarle.
L31. Un policía que pasaba por allí ve toda la situación
L32. y entonces *Charlot*_i intenta escapar de esta situación tan comprometida.

En el siguiente ejemplo, el HN selecciona un nombre propio, *Charlot*, en vez de un sujeto nulo en L25, la cual sería la forma esperada ya que el referente es recuperable a través del contexto (se encuentra activado debido a su mención previa en L23 y a través del determinante *su* en L24). El uso del nombre propio señala el comienzo de una nueva acción y esta forma más elaborada ayuda a mantener cohesión textual:

(17) Ejemplo [P5.L1.TC]

L23. Por otro lado, *Charlot* continua en el banco
L24. y dos hombres se sientan a su lado.
L25. *Charlot* en un descuido aprovecha para beberse la bebida de uno de ellos con lo cual se genera un enfrentamiento entre ellos.

Este fenómeno de un mayor uso de expresiones referenciales más elaboradas en las narrativas escritas se ve en algunas ocasiones también con los HNNs. En (18) el HNN crea un contraste en L12 al utilizar el nombre propio *Charlie* debido a que en esta escena hay tres participantes activados (*Charlie, el hombre y la mujer*); sin embargo, la recuperación del referente es posible a través del uso de un sujeto nulo gracias al contexto. Al seleccionar una forma más elaborada para crear este contraste, se disminuye el esfuerzo cognitivo lo cual conlleva a una mayor cohesión textual, así como vimos en el ejemplo (17):

(18) Ejemplo [P40.L2.NTC]

L6. el hombre_i ofrece traerla algo de tomar.
L7. Mientras [\emptyset]_i está comprando la bebida

- L8. *Charlie Chaplin*_j viene y
- L9. [∅]_j se sienta con la mujer.
- L10. Cuando el hombre_i regresa
- L11. [∅]_i se enoja porque piensa que
- L12. *Charlie*_j quiere conquistarla a la mujer.

En el ejemplo (19), el HNN reintroduce el referente activado con la FN definida *el hombre* en L13, cuando un sujeto nulo hubiera sido suficiente ya que el contexto permite su recuperabilidad. El uso del sujeto explícito puede deberse a razones pragmáticas ya que en este caso nos encontramos con el comienzo de una nueva escena:

(19) Ejemplo [P22.L2.TC]

- L9. *Él*_i ve una oportunidad
- L10. porque su esposa y su hija están completamente dormidas.
- L11. [∅]_i Se fue para coquetear con la otra,
- L12. y ella está dispuesta de charlar con *él*_i también.
- L13. Después de un rato juntos, *el hombre*_i le dijo a la coqueta

Hay otros casos en los cuales los HNNs utilizan una FN definida aun cuando la morfología permite un sujeto nulo, esto lo vemos en el ejemplo (20) en donde la FN definida *los dos* es innecesaria ya que la morfología del verbo (*se sientan*) selecciona al referente plural. Sin embargo, este puede ser un caso motivado por razones pragmáticas para enfatizar la acción de los dos participantes que están en foco, o por razones cognitivas de procesamiento: al referirse a los participantes en foco con una frase explícita, *los dos*, el escritor guía al lector (o a él mismo) a hacer la transformación de dos sujetos singulares en L10 y L11 a uno plural en L12:

(20) Ejemplo [P37.L2.NTC]

- L10. *Él* se levanta y
- L11. [∅] la sigue y
- L12. *los dos* se sientan o en otra banca o en la misma banca pero al otro extremo.

6.3.3 Comparación narrativas escritas TC y NTC

En cuanto a los resultados concernientes al factor TC, no hay evidencia de que este haya afectado la selección de ninguna de las formas de sujeto en el estado ACTR ($p=0.4659$). Sin embargo, hay un uso más elevado de expresiones nominales definidas en las narrativas escritas NTC (18% TC vs 27% NTC), lo cual es un resultado esperado, pero no es significativo ($p=0.2538$, Tabla 5.3.12). Al mismo tiempo vemos un menor uso de otras expresiones elaboradas, como el nombre propio (14% TC vs 8% NTC), aunque, de nuevo, esto puede deberse a las características metodológicas de este estudio. Juntas, ambas formas muestran un uso mayor de formas explícitas en las narrativas NTC: 19 (35%) vs 16 (32%) TC. No obstante, las instancias son muy pocas y no podemos llegar a una conclusión definitiva. Encontramos alguna posible evidencia del TC en los datos con respecto al uso de cláusulas con información de fondo. Según la hipótesis cinco, los HNs codificarán el TC en sus narrativas aportando más información de fondo. Tenemos dos ejemplos de esto en el estado ACTR, ambos con HNs, ninguno con HNNs. El primero de ellos lo vemos en el ejemplo (21), cuando el hablante menciona a la mujer coqueta con la cláusula *la joven del principio*. Esta información adicional no es necesaria para identificar el referente, entre otras razones porque a este personaje se le describe en varios enunciados, y no solamente al principio, como mencionado por el participante:

(21) Ejemplo [P12.L1NTC]

L30. POr otro lado, *la joven del principio* y el hombre siguen en los matorrales

El segundo ejemplo es el siguiente, en el cual se utiliza la cláusula relativa *el hombre que estaba dando el masaje*, para referirse a Charlie:

(22) Ejemplo [P14.L1NTC]

L30. *El hombre que estaba dando el masaje, se desmaya*

Aunque este tipo de cláusulas no ocurren en el estado ACTR en las narrativas con TC, es difícil atribuirle exclusivamente al factor TC su uso: el participante #14 no parece tener otra forma de identificar al referente ya que en ningún momento de su narrativa utiliza nombres propios, y se refiere tanto a Chaplin como al hombre coqueto siempre como *el hombre*. Por lo tanto, para poder identificar a Chaplin en el enunciado del ejemplo (22) la información de fondo es obligatoria. El uso de este tipo de información contextual ocurre también en otros enunciados del mismo participante en otros EC, por ejemplo: “el hombre que apareció segundo” (FAM), “el hombre que se había ido con la mujer” (FAM), “el señor del bigote” (ACTNR), etc.; por lo tanto, puede atribuírsele a una característica individual de este participante para identificar a personajes sin nombre.

6.3.4 Conclusiones estado ACTR

En conclusión, en el estado ACTR de las narrativas escritas vemos la misma tendencia de las narrativas orales de ByLQ a utilizar sujetos nulos, lo cual es evidencia de la adherencia a la GH a través del uso de esta forma referencial mínima; pero, también en este estado, la tendencia es menos fuerte que en las narrativas orales al encontrarnos con un uso significativo mayor de las formas referenciales más elaboradas en las narrativas escritas. Vemos también cómo los factores pragmáticos y cognitivos afectan la selección de los sujetos, aunque el factor TC no parece tener ningún efecto evidente en este estado.

Los datos muestran un acercamiento al uso nativo por parte de los HNNs, lo cual va en línea con las predicciones de las hipótesis dos y tres de este estudio, donde se espera un acercamiento al uso nativo, así como un discurso más cuidado y un aumento de

las formas más elaboradas por parte de ambos grupos al no poder confirmar los niveles de atención y memoria.

Con los ejemplos presentados en esta sección, específicamente (17) y (18), estamos viendo algunas limitaciones de la GH ya que no explica cómo la cohesión textual ayuda a la cognición. Aunque el referente es recuperable con un sujeto nulo gracias al contexto o a la morfología verbal, es posible que el costo cognitivo que se requiere para enlazar de manera rápida y eficiente un pronombre nulo (o un pronombre explícito) con el referente intencionado sea demasiado alto; de esta manera, el uso de formas más elaboradas como FN definidas y nombres propios, aseguran de manera inequívoca la rápida recuperación del referente y, por consiguiente, una mayor cohesión textual.

6.4 Análisis y discusión ACT (GH pronostica OVERT)

6.4.1 Datos escritos HNs vs HNNs

La prueba de chi-cuadrado no revela diferencias significativas entre las composiciones escritas de los HNs y HNNs ($p=0.8487$), de hecho, el uso de todas las formas es casi idéntico para los dos grupos, como lo revelan los resultados de la Tabla 5.2.22, en el Capítulo 5. En nuestros datos escritos no se cumplen las predicciones de la GH para ninguno de los dos grupos, ya que los resultados van en contra de la GH que predice el uso de pronombres explícitos. No obstante, los datos coinciden con la hipótesis tres, puesto que vemos que las expresiones nominales definidas son la forma más utilizada:

- Uso de Overt, narrativas escritas: 23% HNs vs 23% HNNs
- Uso de DefNP, narrativas escritas: 53% HNs vs 53% HNNs

Los ejemplos (23) y (24) de la siguiente sección son casos representativos del uso de expresiones nominales definidas en HNs y HNNs, respectivamente.

6.4.2 Comparación entre datos escritos y datos orales

Los datos demuestran que, aunque ninguno de los dos grupos de hablantes de las narrativas escritas utilizó pronombres explícitos en primer lugar, las tendencias entre las dos modalidades de producción (oral y escrita) no son del todo diferentes, siendo los pronombres explícitos y las FN definidas las dos formas referenciales más utilizadas en los dos estudios (oral de ByLQ y el presente estudio de narrativas escritas). La predicción del estudio oral de ByLQ es el uso de pronombres explícitos cuando el referente está activado en la memoria, ya que el solo uso de un sujeto nulo no permitiría recuperar el referente por el contexto. Las autoras encontraron resultados sorprendentes en este EC, ya que solo en un 41% de los casos los HNs seleccionaron un pronombre explícito, cuando se hubiera esperado un porcentaje de uso más alto. Las FN definidas ocuparon un segundo lugar con el 26% de las instancias. En cuanto a los estudiantes avanzados, el pronombre explícito ocupó el segundo lugar con el 33% de las respuestas, y el primer lugar lo ocuparon FNs definidas con el 44% (p. 156). En el estudio escrito, el uso de pronombres explícitos es menor para los dos grupos de hablantes en comparación con las narrativas orales. La diferencia es significativa entre los HNs ($p=0.0343$). También se observa un acercamiento al uso nativo en las narrativas escritas para esta expresión referencial, como predicho por la hipótesis dos de este estudio. Como vemos, los HNNs seleccionan esta forma en proporción idéntica a los HNs (23%):

- Uso de Overt en HNs: 41% oral vs 23% escrito ($p=0.0343$)
- Uso de Overt en HNNs: 33% oral vs 23% escrito ($p=0.4589$)

La tendencia hacia el uso de DefNP tanto en HNs como en HNNs en las narrativas escritas es confirmada, como planteado por la hipótesis tres. La diferencia entre los datos escritos y los datos orales de ByLQ es significativa únicamente para los HNs ($p=0.0025$). Además, vemos de nuevo que en las narrativas escritas se presenta un acercamiento al habla nativa por parte de los HNNs. Evidencia de esto es la selección de DefNP en igual proporción entre HNs y HNNs en las narrativas escritas (53%), lo cual no sucede en las narrativas orales:

- Uso de DefNP en HNs: 26% oral vs 53% escrito ($p=0.0025$)
- Uso de DefNP en HNNs: 44% oral vs 53% escrito ($p=0.3805$)

En este EC, vemos una mayor necesidad de crear cohesión a través del uso de formas referenciales más elaboradas, especialmente cuando en la escena actual hay tres o más posibles referentes. En el siguiente ejemplo, el HN introduce al hombre y a las dos damas en L2 y L3, respectivamente, activando así los referentes. En L5 y L6, tanto el contexto como la morfología verbal permiten la recuperación de los referentes a través de los pronombres explícitos *él* y *ella*. Sin embargo, en L4 se introduce un nuevo tópico, *la escena*, y el referente *las tres personas* activa de manera colectiva al *hombre* y *las dos damas*, por lo que el lector, al llegar a las líneas L5 y L6, puede identificar a los referentes a través del uso de las FN definidas *el hombre* y *las damas* sin ningún esfuerzo cognitivo:

(23) Ejemplo [P1.L1.TC]

L2. ...Primeramente se encuentra un hombre sentado proveyendo sosten físico a

L3. dos damas que duermen.

L4. La escena da la impresion que las tres personas de la escena inicial estan juntos.

L5. Pero curiosamente *el hombre* se asegura que

L6. *las damas* siguen dormindas para ir detras de la dama

La repetición de las FNs definidas en L5 y en L6 del anterior ejemplo también pueden ser el resultado de perseveración sintáctica, también conocida como paralelismo estructural, la cual es definida como la tendencia de los hablantes de volver a utilizar estructuras sintácticas previamente procesadas (Cleland and Pickering 2006: 187).

Aunque la perseveración no fue un factor que se analizó de manera independiente en este trabajo, se observan algunos ejemplos como (23) que demuestran que hay otros factores que restringen la selección de formas referenciales, los cuales invalidan las predicciones de la GH. Encontramos una posible instancia de paralelismo estructural también con los HNNs. En el ejemplo (24), el pronombre *ella* sería suficiente para recuperar al referente, el cual se encuentra activado gracias a su mención previa en L16, sin embargo, el hablante opta por la frase nominal *la mujer*, posiblemente a causa de una perseveración estructural motivada por la mención anterior:

(24) Ejemplo [P40.L2.NTC]

L16. El hombre y la mujer se van y

L17. *la mujer* sugiere que jueguen a las escondidas.

El uso de nombres propios, cuando el referente está activado, es menor en las narrativas escritas que en las orales para los dos grupos del presente estudio, pero las diferencias no son significativas, manteniendo proporciones similares a los resultados del estudio oral de ByLQ:

- Uso de PrpN en los HN: 13% oral vs 9% escrito ($p=0.5320$)
- Uso de PrpN en los HNN: 17% oral vs 12% escrito ($p=0.4589$)

Aunque las instancias son pocas, hay algunos usos no esperados de sujetos nulos en el estado ACT. Esto también ocurrió en las narrativas orales causando rupturas en la comunicación. Para los grupos de HNs, en el estudio oral (p. 157) se encontraron ocho instancias (12%) en comparación con tres instancias (6%) en el presente estudio escrito; y para los HNNs, en el estudio oral hubo una instancia (2%) en comparación con cinco (9%) en el estudio escrito. Ninguna de estas diferencias entre los datos orales y los escritos es significativa ($p=0.2460$ HNs, $p=0.1009$ HNNs). Los ejemplos (25) y (26) que siguen, extraídos del presente estudio, son casos representativos para los HNs y HNNs, respectivamente. En (25), el sujeto nulo en L40 no nos permite identificar al referente (no sabemos quién se aleja). Sintácticamente suponemos que es la mujer quien lo hace, al ser el sujeto preverbal de la oración anterior, pero hay razones de cohesión discursiva que requieren un sujeto explícito, puesto que el hombre también está activado en la memoria:

(25) Ejemplo [P19.L1.NTC]

L.38. Mientras la pareja que se habían alejado comienzan a jugar a las escondidas,

L39. ella_i le tapa los ojos al señor_j con su pañuelo para que la busca,

L40. en esas [\emptyset]_i? se aleja y

L41. quien llegar cerca de nuevo es *Chaplín* quien decide jugarle una broma al hombre

En el ejemplo (26), aunque podríamos suponer por el contexto que es Charlie quien se sienta al lado de la chica en L22 y no ella quien se sienta al lado de Charlie, el uso del sujeto nulo requiere algo de esfuerzo cognitivo para volver a la acción en L19, donde la mujer está sentada. Un sujeto explícito eliminaría este esfuerzo, contribuyendo a la cohesión discursiva del texto:

(26) Ejemplo [P22.L2.TC]

L19. Un poco después, $[\phi]_i$ pasa por el banco donde se sienta la mujer esperando su bebida.

L20. *Charlie*_i quiere hablar con ella,

L21. pero ella no está muy dispuesta de platicar con él.

L22. $[\phi]_i$ Se sentó a su lado,

6.4.3 Comparación narrativas escritas TC y NTC

Hay un menor uso de FNs definidas en narrativas NTC que en aquellas TC (60% TC vs 47% NTC), lo cual es un resultado sorprendente debido a que, al no haber TC, se esperaba un mayor uso de expresiones referenciales más elaboradas. Sin embargo, la diferencia proporcional corresponde únicamente con dos instancias y por lo tanto los resultados no son indicativos de una tendencia, así como tampoco son estadísticamente significativos ($p=0.1623$, Tabla 5.3.19).

A pesar de que el ejemplo (24) de la sección anterior haga parte de una narrativa NTC, no hay evidencia de que la selección de la expresión más elaborada en L17 haya sido motivada por esta falta de TC; además, al haber un número mayor de instancias en las narrativas con TC, hay suficientes contraejemplos que demuestran la falta de codificación del TC en las narrativas escritas, en contra de la hipótesis cuatro del presente estudio.

Los dos ejemplos anteriores (25) y (26) también son evidencia de la dificultad de aislar el factor TC en las composiciones escritas, ya que ambos actúan como ejemplos contradictorios. En primer lugar, la narrativa del ejemplo (25) asume falta de TC, por lo tanto, no se puede recurrir a este para identificar el referente de L40. El ejemplo (26) es una narrativa donde se asume TC, por lo que podría argüirse que se recurre a este para la identificación del referente en L22. El problema radica en que este participante, en su

texto de 315 palabras y 43 instancias de expresiones referenciales de sujeto, no muestra otra evidencia de que el factor TC haya sido tenido en cuenta en su narrativa. Además, en los datos totales de este estudio, solo se encontraron tres instancias de uso de sujetos nulos en narrativas con TC en el estado ACT (Tabla 5.3.18) y ninguna de ellas sugiere que la selección del sujeto nulo se deba al TC.

El único resultado significativo con respecto al factor TC en el estado ACT es que los hablantes tienden a utilizar más pronombres demostrativos en las narrativas cuando no hay TC ($p=0.0206$, Tabla 5.3.19); este uso de los demostrativos también es pronosticado por la GH. Tenemos dos casos entre los HNNs y cuatro casos entre los HNs. A continuación, presentamos un ejemplo de cada grupo, pero en todas las instancias se observa un uso similar:

(27) Ejemplo [P11.L1.NTC]

L1. Hola: Pues le cuento que vi uno de los varios cortos que hizo Chaplin durante su carrera,

L2. *aquellos* que

L3. representan el cine mundo. Vi “A Woman” de 1915.

(28) Ejemplo [P32.L2.NTC]

L15. [ø] le dio una patada a Charlie

L16. (o [ø] lo empujó--no me acuerdo bien)

L17. y *este* se cayó hacia atrás.

En el ejemplo (27), el uso del demostrativo distal *aquellos* representa una selección elaborada, posiblemente por razones pragmáticas, para contrastar los cortos comunes de los cortos que ya forman parte del cine histórico conocido mundialmente. Se observa que este tipo de demostrativo es infrecuente en el habla: de las 17 instancias registradas en los datos escritos, solo hay dos casos, mientras que en el estudio oral de ByLQ no se dio ningún caso entre las nueve instancias de demostrativos que se generaron. En el ejemplo

(28), un hablante del presente estudio selecciona un sujeto explícito en L17, tal vez por razones pragmáticas ya que hay una interrupción en L16, indicada entre paréntesis por el escritor, y el comienzo de una nueva acción en L17; que se haya seleccionado un demostrativo en vez de un pronombre personal o un sujeto nulo puede ser por razones de cohesión en este caso.

6.4.4 Conclusiones estado ACT

En conclusión, vemos de nuevo un acercamiento al habla nativa en la selección de expresiones referenciales para el estado cognitivo ACT, en particular para las formas Overt y DefNP, y un uso mayor en las narrativas escritas de expresiones más elaboradas en el caso de FNs definidas para ambos grupos.

En el estudio oral de ByLQ, la hipótesis predecía el uso de pronombres personales, cumpliéndose únicamente con los HNs, siguiéndole de cerca el uso de FNs definidas, mientras que los HNNs de su estudio seleccionaron mayormente FNs definidas. En el presente estudio escrito, las FNs definidas son la forma más utilizada por ambos grupos de hablantes. Esta tendencia podría sugerir que, para el estado ACT, la GH en español predice también el uso de las frases definidas y no solamente los pronombres explícitos. Esta tendencia se ve reforzada en las narrativas escritas.

Asimismo, vemos que el factor TC no tiene un efecto significativo en la selección de sujetos en las narrativas escritas y el único resultado con significación estadística (los demostrativos) puede explicarse cualitativamente con razones pragmáticas y, tal vez, estilísticas.

6.5 Análisis y discusión ACTNR (GH pronostica GenNP)

6.5.1 Datos escritos HNs vs HNNs

En las narrativas escritas del presente estudio, todas las expresiones nominales generales (GenNP) fueron reclasificadas bien sea como DefNP o como IndNP. En el Apéndice B se presentan todas las instancias de GenNP utilizadas en el estado ACTNR en las narrativas escritas, entre ellas tenemos como ejemplo las expresiones *uno*, *uno de los hombres*, *la otra mujer*, etc. Cuando eliminamos la categoría GenNP, vemos que la tendencia a usar FNs definidas es muy clara en este EC. En general, no se encontraron diferencias significativas entre HNs y HNNs a través de la prueba de chi-cuadrado ($p=0.1046$, Tabla 5.2.29), demostrando un uso similar entre los dos grupos de hablantes:

- Uso de DefNP en HNs en las narrativas escritas: 56%
- Uso de DefNP en HNNs en las narrativas escritas: 63%

Al realizar un análisis de regresión logístico, encontramos que los HNs muestran una mayor probabilidad de utilizar FNs indefinidas que los HNNs ($p=0.0435$, Tabla 5.2.30). La diferencia entre los dos grupos de participantes no es significativa ($p=0.4696$):

- Uso de IndNP en HNs en las narrativas escritas: 18%
- Uso de IndNP en HNNs en las narrativas escritas: 6%

Las expresiones indefinidas que se realizan en este EC, aunque se clasifican como indefinidas en el presente estudio, son semánticamente generales y no-pronominales. Es por este motivo que en la siguiente sección (6.4.2) estas expresiones se consideran GenNP para poder hacer una comparación directa con el estudio de las narrativas orales de ByLQ. A continuación, se presentan de nuestros datos dos ejemplos representativos donde vemos que el uso de estas cláusulas generales indefinidas es igual entre los HNs y

los HNNs:

(29) Ejemplo [P12.L1.NTC]

L27. en el que *uno de los hombres* le empuja con el bastón

(30) Ejemplo [P24.L2.TC]

L27. y cuando *uno de ellos* se enoja

6.5.2 Comparación entre datos escritos y datos orales

Junto con el estado ACT, el estado activado y no recuperable (ACTNR) muestra resultados que se desvían ligeramente de los resultados orales del estudio de ByLQ. En su estudio se predice el uso de expresiones nominales generales para referirse a referentes que se encuentran en la memoria operativa, pero que no sean recuperables con el solo uso de un pronombre explícito (p. 158). Esta predicción se cumple únicamente con los HNs (44%) en los datos orales, mientras que los estudiantes avanzados utilizan dichas expresiones solamente el 10% de las veces, siendo el nombre propio la forma más utilizada (59%), seguida por las FNs definidas (26%). En el presente estudio escrito los datos confirman la tendencia hacia el uso de expresiones nominales definidas para ambos grupos.

Al introducir la categoría GenNP para comparar los resultados de las narrativas orales de ByLQ con las narrativas escritas del presente estudio, 11 instancias de DefNP y 10 de IndNP de los datos escritos se vuelven a clasificar como GenNP. Aún reclassificando estas instancias como GenNP, el uso de las FNs definidas sigue ocupando el primer lugar para los dos grupos de hablantes en las narrativas escritas. El mayor uso de DefNP en las narrativas escritas comparado con las narrativas orales es significativo en los HNNs:

- Uso de DefNP en HNs: 30% oral vs 44% escrito ($p=0.1462$)
- Uso de DefNP en HNNs: 26% oral vs 54% escrito ($p=0.0061$)

La frecuencia en el uso de GenNP es menor en los HNs en las narrativas escritas que en las orales. Aunque proporcionalmente la diferencia es considerable, estadísticamente el resultado no es significativo ($p=0.0945$) y proviene únicamente de ocho instancias; por lo tanto, el resultado solo es indicativo de esta tendencia. En los HNNs, el uso de GenNP es ligeramente mayor en las narrativas escritas (tres instancias) y el resultado no es significativo:

- Uso de GenNP en HNs: 44% oral vs 28% escrito ($p=0.0945$)
- Uso de GenNP en HNNs: 10% oral vs 13% escrito ($p=0.6881$)

En cuanto al tipo de expresiones generales, en el estudio oral de ByLQ se notó que los hablantes avanzados usaron una menor variedad de expresiones que los HNs (p. 158). Esta misma tendencia se observa en las narrativas escritas: siete HNs utilizaron expresiones generales en el estado ACTNR versus seis HNNs; entre los HNs se utilizaron nueve tipos de expresiones versus seis por parte de los HNNs²². Además de una mayor variedad, los HNs de los dos estudios muestran la habilidad de usar expresiones más complejas tipo *una de las escenas, una de las mujeres, los dos enamorados*. El ejemplo (31) ilustra el uso de una de estas expresiones en las narrativas escritas por parte de un HN y el ejemplo (32) por parte de un participante HNN:

(31) Ejemplo [P19.L1.NTC]

L4. en donde se encuentra un hombre sentado en una banca con dos mujeres.

L5. *una de ellas* duerme

²² Las expresiones que solo cambian el género se cuentan como la misma expresión: uno/a, uno de los hombres/una de las mujeres, etc.

- (32) Ejemplo [P36.L2.NTC]
 L3. Las mujeres parecen aburridas
 L4. porque *una* está dormida

En el estudio oral de ByLQ, las autoras atribuyen posiblemente al mayor número de sujetos producido por los HNs el hecho de que se dieran más expresiones generales: los HNs produjeron 50 expresiones referenciales de sujeto en el estado ACTNR versus 39 entre los HNNs (p. 158). Sin embargo, en las narrativas escritas la proporción no solamente es casi igual entre los dos grupos de hablantes, con solo cuatro instancias de diferencia, sino que fueron los HNNs quienes produjeron un número mayor de expresiones referenciales en este estado. Por lo tanto, podría argüirse que la razón se debe a la capacidad lingüística de los HNs de producir expresiones referenciales más complejas. La Tabla 6.5.1, más abajo, muestra la producción total de sujetos en el EC ACTNR, en donde vemos una frecuencia casi idéntica en las composiciones escritas entre los dos grupos de hablantes:

Tabla 6.5.1 - Total sujetos ACTNR entre las narrativas orales y escritas, L1 vs L2

Estado Cognitivo	ACTNR	
	Oral	Escrito
L1	50 (56.2%)	50 (48.1%)
L2	39 (43.8%)	54 (51.9%)
TOTAL	89 (100%)	104 (100%)

En general, en las narrativas escritas observamos que el uso de expresiones definidas en vez de expresiones más generales es necesario cuando el referente está ACTNR, puesto que las expresiones más elaboradas contribuyen a la identificación de la entidad en referencia, mientras que con una expresión general se corre el riesgo de seleccionar los sujetos incorrectos. Esto puede deberse a la necesidad de los hablantes de

elaborar con más detalle en el discurso escrito, como predicho por la hipótesis tres del presente estudio, debido a la imposibilidad de confirmar los niveles de atención y memoria del lector. En el ejemplo (33) que sigue, el HN selecciona la expresión definida *el viejo y la chica* en vez de una expresión más mínima como *los dos*, y en (34), en la misma escena el HNN selecciona *el hombre y la mujer*:

(33) Ejemplo [P9.L1.TC]

L11. El viejo se pone furioso
L12. porque [ø] piensa que
L13. el chaplin est'a ligando con ella.
L14. [ø] Lo echa del lugar.
L15. Y entonces *el viejo y la chica* se ponen a jugar a la gallinita ciega

(34) Ejemplo [P23.L2.TC]

L20. El hombre usa una botella para pegar a Charlie.
L21. Charlie se desmaya y
L22. *el hombre y la mujer* salen juntos.

En estos ejemplos vemos que una expresión más general como *los dos* en L15 y en L22 seleccionarían erróneamente al hombre y a Charlie al ser los tópicos de las oraciones precedentes, en vez de al hombre y a la mujer, aunque el contexto puede ayudar en el proceso de desambiguación.

Solo encontramos un ejemplo, perteneciente a un HNN, en donde una expresión general puede fácilmente reemplazar el uso de una FN definida sin causar ambigüedad, pero esto es gracias al contexto lingüístico, ya que el HNN inmediatamente menciona a Charlie en posición de objeto en L18:

(35) Ejemplo [P28.L2.TC]

L16. Charlie Chaplin está tocando a la coqueta en su pie y
L17. se enoja el primer hombre.
L18. *La coqueta y el hombre* se van y dejan Chaplin en el banco.

En el estado ACTNR observamos que en varias instancias las FNs definidas incluyen información de fondo. Esto ocurre en los dos grupos de hablantes y, mayormente, en las narrativas donde no se asume TC; sin embargo, es más probable que la información de fondo se deba a la necesidad de asegurar la identificación del referente debido a la restricción derivada del estado ACTNR, que a la falta de TC; esto es porque se requiere de un mayor esfuerzo cognitivo para la recuperación del referente en este EC. Algunos ejemplos de información de fondo que se encontraron en los presentes datos son los siguientes: *el hombre casado*, *el hombre que estaba sentado en el banco*, *el señor del bigote*, *el muchacho del bastón*, entre otros. A continuación, presentamos ejemplos representativos por parte de un HN (36) y de un HNN (37):

(36) Ejemplo [P14.L1.NTC]

L6. Cuando [ø] están sentados,
L7. una tercera mujer pasa por delante de los tres personajes anteriores
L8. y el hombre que estaba sentado en el banco la mira con interés.

(37) Ejemplo [P34.L2.NTC]

L48. Cuando el papá se quita la venda de los ojos
L49. [ø] descubre que es el muchacho con el bastón que estaba con la coqueta.
L50. El muchacho del bastón mide que tan profundo es el lago y
L51. [ø] termina tirando al papá al lago.

En cuanto al uso de nombres propios, se observa una tendencia contraria al comportamiento de las FNs definidas para los dos grupos de hablantes, ya que hay un menor número de ocurrencias de esta forma referencial en las narrativas escritas que en las orales de ByLQ. En las narrativas orales, los nombres propios son la forma más utilizada para los HNNs (59%), y pasan a ocupar un segundo lugar en las narrativas escritas (28%), esta diferencia entre las dos modalidades de producción es significativa

($p=0.0024$). Por otra parte, los HNs mantienen un uso similar (24% oral; 16% escrito) y esta diferencia no es significativa ($p=0.3160$). De nuevo, cabe recordar que el nivel de significación para esta forma se ve atenuado por posibles interferencias metodológicas del estudio.

Aunque en el estado ACTNR vemos un mayor número de ocurrencias de DefNP en las narrativas escritas con respecto a las orales de ByLQ, también hay un mayor uso de sujetos nulos, aunque el número de instancias es muy limitado (Tabla 5.2.31). En las narrativas orales solo se presenta un caso de un sujeto nulo y es por parte de un HNN; mientras que en las narrativas escritas tenemos seis casos (4 HN y 2 HNN). La diferencia entre los dos modos de producción es significativa en los HNs ($p=0.0499$) y no es significativa en los HNNs ($p=0.7561$). Vamos a ver algunos ejemplos de sujetos nulos de las narrativas escritas en ACTNR en la siguiente sección, ya que se analizan desde la perspectiva del TC.

6.5.3 Comparación narrativas escritas TC y NTC

No hay evidencia en los datos escritos de que el factor TC influya de manera significativa en la selección de sujetos en el estado ACTNR. Sin embargo, vemos que hay un mayor número de instancias de expresiones más descriptivas (DefNP+PrpN) provenientes de las narrativas sin TC. Este es un comportamiento esperado ya que la hipótesis cinco de este estudio predice que la falta de TC motiva a los hablantes a intentar asegurar una correcta identificación del referente a través de cláusulas léxicas más elaboradas. Combinadas estas dos formas, hay 29 (78%) instancias en narrativas TC versus 56 (83%) instancias en narrativas NTC. Esta diferencia proviene de un mayor uso de nombres propios en las narrativas NTC (16% TC vs 25% NTC), ya que las

expresiones definidas muestran un comportamiento contrario (62% TC vs 58% NTC). El análisis de regresión logística aplicado a estas formas no arrojó resultados significativos de estas diferencias: DefNP, $p=0.6936$ y PrpN, $p=0.2727$ (Tabla 5.3.26). Asimismo, en el análisis cualitativo de los datos no se encuentran ejemplos de estas formas donde sea evidente la codificación del TC; en cambio, su selección parece depender más del nivel de recuperabilidad del referente, puesto que una FN definida o un nombre propio seleccionan a la entidad intencionada de manera menos ambigua que una expresión general.

Con respecto al uso de sujetos nulos mencionados en la sección anterior, se esperaría que se hiciera un mayor uso de esta forma referencial en narrativas con TC; sin embargo, se observa un número idéntico de instancias en composiciones TC y NTC. Proporcionalmente, el uso es mayor en las composiciones TC, aunque esta diferencia no es significativa ($p=0.4564$, Tabla 5.3.26):

- Uso de Null en narrativas TC: 3 (8%)
- Uso de Null en narrativas NTC: 3 (4%)

Las instancias son demasiado pocas para llegar a conclusiones definitivas, pero hay algunos ejemplos individuales que podemos resaltar. En L10 del ejemplo (38), más abajo, el sujeto nulo se refiere a los dos referentes (*Charlie y el hombre*) que han sido sujetos en las cláusulas anteriores, sin embargo, el personaje de la mujer coqueta también se encuentra activado. Para evitar ambigüedad, el hablante debió utilizar una expresión general definida como *los dos hombres*, ya que ni siquiera se ha mencionado que el hombre ha regresado con las bebidas y por lo tanto no se puede recurrir al contexto. Al haber TC en la narrativa se elimina esta posible ambigüedad causada por el uso del sujeto

nulo:

- (38) Ejemplo [P8.L1.TC]
- L5. El hombre_i se va tras de ella
 - L6. Y [\emptyset]_i le ofrece una bebida,
 - L7. Mientras que el_i le trae la bebida
 - L8. Aparece otro hombre_j
 - L9. (Chaplin)_j, también a conquistarla,
 - L10. [\emptyset]_{ij} tienen una discusión

Aunque el anterior es un ejemplo donde se puede argüir por el efecto del TC, también es posible que el uso del sujeto nulo en L10 permita la recuperabilidad de los referentes simplemente porque selecciona a los dos referentes que han sido sujetos en las cláusulas anteriores (*Chaplin* en L9 y *otro hombre* en L8). Además, tenemos datos que sirven de contraejemplo con un HN en una narrativa donde no se asume TC. En el ejemplo (39), al no poderse recurrir al TC, se da una posible ruptura en la comunicación al utilizar sujetos nulos en L51 y L56 ya que, sintácticamente, seleccionan al referente incorrecto. En (39a), Charlie (*el hombre que apareció segundo*) es el tópico de las oraciones desde L48 hasta L50, pero no en L51, ya que quien lleva a cabo la acción de tocar el bigote es el hombre coqueto, el cual lleva los ojos vendados. En (39b), el tópico pasa a ser el hombre con los ojos vendados desde L53 a L55, pero en L56 quien ha sido reconocido es Charlie. En ambos casos se requiere de una frase léxica para la recuperación del referente ya que no se encuentra en el centro de atención y no está en una posición sintáctica prominente (posición de tópico). Además, al no haber TC en esta narrativa, tampoco se puede depender de este para la identificación del referente:

- (39) a. Ejemplo [P14.L1.NTC]
- L48. El hombre que apareció segundo_i, lo encuentra con los ojos tapados
 - L49. y [\emptyset]_i lo guía hacia el lago cogiéndolo del cuello con el

bastón.

L50. [∅]_i Lo pasea por el lago

L51. y al final [∅]^{*i/j} le toca el bigote

L52. y [∅]_j se da cuenta que no es la mujer.

b. Ejemplo [P14.L1.NTC]

L53. [∅]_j Se quita la venda de los ojos,

L54. [∅]_j lo reconoce,

L55. porque [∅]_j lo ha visto antes dándole un masaje en el pie a la señora

L56. y como [∅]^{*j/i} ha sido reconocido,

L57. el otro hombre_i lo tira al lago.

Los ejemplos (38) y (39) nos llevan a pensar que, aunque en algunas instancias el TC es un factor que no está lingüísticamente codificado por el hablante/escritor, sí es utilizado (o puede ser utilizado), por el oyente/lector en su proceso de asignación referencial. Este proceso será más o menos exitoso dependiendo del EC en el cual se encuentre el referente en la mente del oyente/lector, debido a los diferentes niveles de atención y memoria de cada estado.

Los dos casos de uso de Null entre los HNNs, mencionados al final de la Sección 6.5.2, son en composiciones donde se asume TC. Sin embargo, es difícil asegurar que su uso se deba a este factor y no a factores lingüísticos de adquisición. En el ejemplo (40) que sigue, hay dos hombres en la escena (L25) y por lo tanto el uso de un sujeto nulo en L26 no es adecuado puesto que el sujeto en L26 es singular, como evidenciado por la morfología verbal de *pisar*. Necesitamos una frase general como *uno/uno de ellos/uno de los hombres*. Esta parece ser una situación de falta de adquisición por parte del HNN, como lo vemos también por el uso incorrecto del objeto directo *le* en L26:

(40) Ejemplo [P28.L2.TC]

L25. Los hombres_i se van y ultimo,

L26. [∅]_j? le pisa a Charlie.

6.5.4 Conclusiones estado ACTNR

En conclusión, en el estado ACTNR la diferencia entre HNs y HNNs en las narrativas escritas no es significativa. Al comparar estos datos con las narrativas orales, vemos un desvío de los datos orales con una mayor tendencia hacia el uso de expresiones definidas en los datos escritos. El acercamiento al uso nativo se ve reflejado en el hecho de que, para los dos grupos de hablantes, las expresiones nominales definidas son la forma más seleccionada en las narrativas escritas, lo cual no sucede en las narrativas orales de ByLQ. El menor uso de nombres propios por parte de los HNNs en las narrativas escritas, combinado con el mayor uso por parte de los HNs, reduce la diferencia entre HNs y HNNs, en línea con la hipótesis dos del presente estudio.

Las expresiones generales eran la forma esperada en este estado y observamos un menor uso en las narrativas escritas en los dos grupos de hablantes. Sin embargo, vemos los mismos patrones en el discurso escrito de nuestro estudio que en el discurso oral del estudio previo: los HNs tienden a utilizar estas expresiones generales con mayor frecuencia y con un mayor nivel de complejidad que los HNNs.

No hay evidencia significativa de que el factor TC tenga un efecto en las composiciones escritas, y hemos visto ejemplos donde los mismos usos de una forma se presentan tanto en narrativas TC como en narrativas NTC. Aunque se observa un ligero aumento en el uso de las formas más elaboradas en las narrativas NTC, no hay evidencia de este a través del análisis cualitativo.

6.6 Análisis y discusión FAM (GH pronostica DefNP y PrpN)

6.6.1 Datos escritos HNs vs HNNs

En las narrativas escritas del presente estudio los resultados de la prueba chi-cuadrado muestran un uso similar en la selección de las formas en el estado familiar (FAM) entre los HNs y HNNs ($p=0.2099$, Tabla 5.2.37), siendo DefNP y PrpN las dos formas más utilizadas, como pronosticado por la jerarquía. A través del análisis de regresión logística no se encontraron diferencias significativas entre los hablantes para estas dos formas:

- Uso de DefNP, narrativas escritas: 65% HNs vs 75% HNNs ($P=0.1778$)
- Uso de PrpN, narrativas escritas: 32% HNs vs 21% HNNs ($P=0.1244$)

6.6.2 Comparación entre datos escritos y datos orales

La predicción del estudio oral de ByLQ era con respecto al uso de expresiones nominales definidas y nombres propios cuando el hablante puede identificar el referente debido a una mención anterior en el discurso (p. 159). Esta predicción se confirmó en su estudio y también en el presente estudio; o sea, se observa en nuestros datos escritos que las expresiones definidas son la forma más seleccionada para los dos grupos de hablantes y los nombres propios ocupan el segundo lugar.

Al sumar ambas formas (DefNP+PrpN), observamos un número casi idéntico de ocurrencias entre las dos modalidades de producción: 131 instancias en las narrativas orales de ByLQ vs 129 instancias en las narrativas escritas. En las dos modalidades estas instancias corresponden con el 89% del total en el estado FAM (Tabla 5.2.39).

El único resultado con significación estadística es con respecto al uso de los nombres propios, donde vemos que la escritura tiene un efecto significativo sobre esta

forma: los HNs tienden a utilizar más nombres propios en las narrativas escritas, mientras que los HNNs tienen una tendencia opuesta (Tablas 5.2.40 y 5.2.41, respectivamente):

- Uso de PrpN en HNs: 16% oral vs 32% escrito ($p=0.0173$)
- Uso de PrpN en HNNs: 37% oral vs 21% escrito ($p=0.0371$)

Aunque no es significativa, hay una tendencia contraria en el uso de expresiones nominales definidas, ya que los HNs hacen un menor uso de esta forma en las narrativas escritas y los HNNs hacen un mayor uso:

- Uso de DefNP HNs: 67% oral vs 54% escrito ($p=0.1130$)
- Uso de DefNP HNNs: 60% oral vs 71% escrito ($p=0.1676$)

Al ser las FNs definidas y los nombres propios las mismas dos formas más utilizadas en las narrativas orales, y teniendo en cuenta la salvedad metodológica de los nombres propios mencionada en la Sección 6.3.2 con respecto a la imposibilidad de verificar el conocimiento que los diferentes participantes pueden tener del personaje Charlie Chaplin, no podemos decir de manera concluyente que los cambios entre DefNP y PrpN sean a consecuencia de la producción escrita ya que pueden deberse a las características metodológicas del estudio. Sin embargo, esto no elimina el hecho de que los resultados hayan identificado las mismas tendencias de adherencia a la GH en cuanto al uso de las formas referenciales en el estado FAM.

En el EC FAM del estudio oral, ByLQ comentan cómo el uso de sujetos nulos y pronombres explícitos en este estado causa problemas de comunicación al no poder identificar correctamente al referente. En su estudio, se registraron tres instancias de sujetos nulos y una de un pronombre explícito, todas las instancias con HNs (p. 159). En el estudio escrito se presentan únicamente dos casos de sujetos nulos, pero solamente con

un HNN, y ningún caso de pronombres explícitos. Esto puede deberse a un lenguaje más cuidado en las composiciones escritas, como predicho por la hipótesis tres. Sin embargo, en el presente estudio también notamos problemas en la comunicación a través del uso de sujetos nulos. En el ejemplo (41), vemos que el sujeto nulo no permite recuperar al referente y, al haber otros personajes en escena (*las otras dos señoras*), el hablante obliga a un alto esfuerzo cognitivo con el uso de esta forma referencial mínima. En L42 el sujeto nulo hace referencia al hombre coqueto, el cual no hace parte de la escena actual, y por lo tanto no está activado en la memoria operativa (este personaje aparece por última vez en posición de tópico en L33). Además, al ser esta una narrativa sin TC, tampoco se puede recurrir a este para la identificación del referente:

(41) Ejemplo [P37.L2.NTC]

L41. el que la chica le roba su cartera al señor. Me pareció raro que no viéramos más a las otras dos señoras

L42. del principio, una de ellas siendo probablemente la esposa...la que [\emptyset] había dejado para seguir a la

L43. coqueta, pero igual me pareció una buena venganza por su comportamiento que le robaran su dinero!

En general, el uso de demostrativos no es común en ninguna de las narrativas, ya sea oral o escrita y, de hecho, ByLQ no reportan ninguna instancia en el estado FAM en sus narrativas orales, pero sí ocurre en las narrativas escritas del presente estudio, con una instancia en un HN (42) y otra en un HNN (43):

(42) Ejemplo [P4.L1.TC]

L23. al mismo tiempo que *aquél hombre* que fue a por refrescos se percata que le están robando la chica

(43) Ejemplo [P26.L2.TC]

34. *Ese Charlie* siempre sale con

En el ejemplo (42), el demostrativo nominal *aquél hombre* es una expresión deíctica con una función anafórica; es decir, es una expresión referencial de sujeto que toma su referencia del texto previo (García Salido, 2011:66). En este caso el referente es el hombre coqueto, el cual fue introducido previamente en la narrativa. Por otra parte, la función del demostrativo nominal *ese Charlie* del ejemplo (43) puede ser simplemente un énfasis pragmático de las características personales de ese personaje. Vemos que este participante no tuvo tiempo de terminar la narrativa y su oración está truncada, al ser la última de su texto. Sin embargo, parece que el hablante pretendía utilizar la conocida expresión en español “salirse con la suya” para describir el comportamiento de Charlie y el uso del demostrativo enfatizaría el enunciado.

6.6.3 Comparación narrativas escritas TC y NTC

En el estado cognitivo FAM no hay evidencia de que el factor TC tenga una influencia en la selección del sujeto en las narrativas escritas. Proporcionalmente tampoco se ven diferencias que puedan resaltarse. Sin embargo, a nivel individual, el caso del participante #9 sobresale por los siguientes dos motivos: en primer lugar, este es el único caso donde un HN utiliza una FN indefinida para un referente que es FAM y puede explicarse a raíz del efecto del TC. En L1 del ejemplo (44), el HN introduce al personaje en la narrativa a modo de conversación con el lector ficticio de la encuesta, para luego pasar a describir los eventos del videoclip, de nuevo introduciendo al mismo personaje en L2 con la frase indefinida *un hombre*. Por lo tanto, este sujeto se clasifica como FAM puesto que el personaje ya ha sido introducido en L1 como *el hombre que quería ligar con una chica joven*, llevando al referente a la memoria a través de su mención:

(44) Ejemplo [P9.L1.TC]

L1. Te acuerdas de la película que vimos? Si, la del hombre que quer'ia ligar con un chica joven, en el

L2. parque.... Que no te acuerdas? quieres que la recuerde? Pues el resumen es que hay *un hombre aburrido*

En segundo lugar, más adelante en L8 del ejemplo (45), el mismo hablante introduce al personaje de Chaplin, pero no lo hace con un nombre propio sino con la cláusula indefinida *otro tipo*, en el estado de tipo identificable (TYP). Luego, en L16, vuelve a mencionar a Chaplin, esta vez por nombre. Como el personaje ya ha sido presentado en la narrativa, se clasifica como FAM, sin embargo, de no haber TC en este texto, no se hubiera podido hacer una correferencia exitosa entre *otro tipo* en L8 y *Chaplin* en L16, y el uso del nombre propio hubiera causado un efecto de ruptura en la cohesión del texto. No se dan este tipo de ejemplos con HNNs:

(45) Ejemplo [P9.L1.TC]

L8. llega otro tipo_i que por descuido tira a la chica del banco.

.
. .

L16. En eso *chaplin*_i aprovecha que

6.6.4 Conclusiones estado FAM

En conclusión, en el EC FAM, tanto los HNs como los HNNs se adhieren a la GH y las pocas diferencias que se registran entre los dos grupos no afectan la distribución de las formas: las FN definidas y los nombres propios siguen siendo las dos formas referenciales más utilizadas en las dos modalidades de producción del discurso. La hipótesis uno de este estudio predice que los estados con mayor adherencia serán los de INF y TYP, los resultados obtenidos para FAM revelan que esta adherencia es igualmente clara para este EC.

Aunque estadísticamente el TC no muestra un efecto significativo, a través del análisis cualitativo se observa el caso de un HN en el ejemplo (44), quien explícitamente introduce el hecho de que su interlocutor ha visto la película. Esto no afecta la selección de las formas referenciales de sujeto ya que el hablante elige una FN indefinida para introducir al personaje en la narrativa, como es lo esperado, pero sí afecta el EC de la forma utilizada debido a que el referente ya se encuentra en la memoria, por lo tanto, se clasifica como FAM y no como TYP.

6.7 Análisis y discusión UNI (GH pronostica DefNP)

6.7.1 Datos escritos HNNs vs HNs

La prueba de chi-cuadrado en las composiciones escritas revela que no hay diferencias significativas en el uso de las formas para el EC identificable de modo único (UNI) entre los HNs y HNNs ($p=0.6888$, Tabla 5.2.45). La distribución de las FN definidas es idéntica para los dos grupos de hablantes, lo cual es evidencia de un uso casi nativo por parte de los HNNs:

- Uso de DefNP en HNs en las narrativas escritas: 68%
- Uso de DefNP en HNNs en las narrativas escritas: 68%

6.7.2 Comparación entre datos escritos y datos orales

En el estudio oral de ByLQ, la predicción para el EC UNI es que los hablantes se referirán a las entidades a través del uso de expresiones nominales definidas, cuyo simple uso permita la identificación del referente, bien sea por la expresión en sí al ya existir en la memoria a largo plazo del hablante, o por asociación con otro referente previamente activado en el discurso reciente (p.160). Tanto los resultados orales de ByLQ como

nuestros datos escritos apoyan esta predicción: las expresiones nominales definidas son la forma más usada para los HNs y los HNNs, seguidas por los nombres propios.

Como predicho por la hipótesis dos del presente estudio, en este estado vemos de nuevo un acercamiento al uso nativo en las narrativas escritas. Esto se debe a que los HNs hacen un menor uso de FN definidas en las narrativas escritas versus las orales; mientras que los HNNs hacen un mayor uso de estas. Sin embargo, las diferencias entre los dos estudios no son estadísticamente significativas:

- Uso de DefNP en HNs: 83% oral vs 68% escrito ($p=0.1554$)
- Uso de DefNP en HNNs: 53% oral vs 68% escrito ($p=0.2710$)

El uso de PrpN tiene una tendencia contraria; es decir, los HNs hacen un mayor uso de nombres propios en las narrativas escritas y los HNNs hacen un menor uso. De nuevo, ninguna de estas diferencias es estadísticamente significativa:

- Uso de PrpN en HNs: 17% oral vs 30% escrito ($p=0.2344$)
- Uso de PrpN en HNNs: 47% oral vs 23% escrito ($p=0.0677$)

Ejemplos de expresiones definidas cuando el referente es UNI son cuando los hablantes se refieren a la película en sí (*la película, el cortometraje*), o expresiones relacionadas con películas como marcos cognitivos (*el hilo conductor, el personaje principal, los subtítulos*). También se usan expresiones definidas cuando se ha introducido a la familia con expresiones como *una familia*, para subsecuentemente pasar a mencionar a sus miembros de manera individual (*la mujer y la hija, el padre, las mujeres*). Algunos ejemplos ilustrativos de estos casos son los siguientes y, como vemos, este uso de FNs definidas sucede tanto con los HNs, ejemplo (46) como con los HNNs, ejemplos (47) y (48):

- (46) Ejemplo [P2.L1.TC]
L1. *La película* trata sobre una familia que va al parque y
- (47) Ejemplo [P33.L2.NTC]
L20. Me encantan los cortometrajes mudos porque *los personajes*, mediante su lenguaje corporal, tienen que
- (48) Ejemplo [P22.L2TC]
L1. Este videoclip comienza mostrando una familia descansando el parque.
L2. *El padre* está de muy mal humor,

El uso de nombres propios en este estado se debe a que el personaje de Charlie Chaplin no necesita ningún otro tipo de introducción al ser universalmente conocido. Tanto los HNs, ejemplo (49), como los HNNs, ejemplo (50), utilizan un nombre propio para introducir a este personaje por primera vez en la narrativa:

- (49) Ejemplo [P10.L1TC]
L12. En ese momento aparece *Charlot*, que pasea por el parque,
- (50) Ejemplo [P24.L2TC]
L11. Ahora vemos que *Charlie Chaplin* viene caminando por el parque y

Un caso que difiere entre las narrativas orales de ByLQ y las narrativas escritas de este estudio para el EC UNI es el de los demostrativos. Como mencionado anteriormente en la Sección 6.6, en las dos modalidades de producción se encontraron pocos casos de demostrativos y estos son usados por los HNs con mayor frecuencia. La Tabla 6.7.1, a continuación, muestra el total de las ocurrencias de esta forma referencial en todos los estados de la GH, comparando las narrativas escritas del presente estudio con las orales de ByLQ:

Tabla 6.7.1 – Uso total de demostrativos en las narrativas escritas y orales, L1 vs L2, todos los ECs

Forma	Dem	
	Oral	Escrito
L1	7 (77.9%)	9 (52.9%)
L2	2 (22.2%)	8 (47.1%)
TOTAL	9 (100%)	17 (100%)

Como vemos en la anterior tabla, los demostrativos representan un porcentaje mínimo dentro del uso total de las expresiones referenciales: 0.74% en las narrativas orales y 1.54% en las narrativas escritas. Gundel et ál. (1993:292) dicen que, en español, los pronombres demostrativos *aquel/ese/este/este N* solo requieren activación y para los demostrativos *ese N/aquel N* el estado FAM es suficiente. En su estudio, en la distribución total de los sujetos referenciales notamos que no se encontró ningún demostrativo que requiriera los estados UNI o TYP. Igualmente, en el estudio oral de ByLQ (p. 153) no se encontró ningún demostrativo a partir del estado ACTNR; en otras palabras, todos los demostrativos requieren que sean, por lo menos, activados.

En las narrativas escritas del presente estudio se encontraron casos de demostrativos en UNI y en TYP. En UNI solo hay tres de estas instancias y es por parte de HNNs, representando el 7% del total en el estado UNI. Esto es inusual porque, como mencionado más arriba, los demostrativos suelen ser más utilizados por HNs. En todas las ocasiones las expresiones demostrativas son utilizadas para comenzar la descripción de la película, como vemos en los siguientes ejemplos:

(51) Ejemplo [P22.L2.TC]

L1. *Este videoclip* comienza mostrando una familia descansando el parque.

(52) Ejemplo [P23.L2.TC]

L1. *Este cortometraje* empieza con una familia sentada en un banco en el parque.

(53) Ejemplo [P37.L2.NTC]

L1. *Este clip* de una película silente de Chaplin comienza con 3 personas sentadas en una banca en un

En todos estos ejemplos el demostrativo nominal *este N* es utilizado en la primera línea de las composiciones para comenzar a describir la película. Se clasifica dentro del EC UNI porque es la primera mención del referente (*la película*) y por lo tanto no está activado en la memoria operativa; sin embargo, con base en la descripción de Gundel (2010:154), en el estado UNI se puede construir una representación única del referente con la sola información conceptualmente codificada en la oración. En otras palabras, aunque el referente es introducido por primera vez en la narrativa, el referente existe en la memoria a largo plazo y la sola expresión ayuda en el proceso de identificación; por lo tanto, el demostrativo actúa como una expresión definida capaz de seleccionar una entidad única. Como vimos en el Capítulo 3 del presente trabajo, este tipo de referencia es exofórica; es decir, lleva al lector fuera del texto y por lo tanto no se considera cohesiva (no une dos elementos en un texto). Al no haber ocurrido en las narrativas orales podría considerarse exclusiva de las narrativas escritas, aunque se necesitaría de mayor evidencia para llegar a una conclusión definitiva a este respecto.

6.7.3 Comparación narrativas escritas TC y NTC

En el análisis cuantitativo no se encontró evidencia de que el factor TC influyera en la selección del sujeto en las narrativas escritas. Proporcionalmente tampoco se ven diferencias que puedan resaltarse (Tabla 5.3.39): las dos formas más utilizadas, DefNP y

PrpN representan el 93% en las narrativas TC y el 95% en las narrativas NTC. Ninguna de estas formas arrojó resultados significativos entre los hablantes ($p=0.6888$). Este resultado es esperado en el estado UNI, ya que, como predicho por la GH para el español, se espera el uso de expresiones definidas que asuman familiaridad con el referente. En el ejemplo (54), más abajo, el hablante reconoce el TC que tiene con su interlocutor ficticio en una especie de pre-introducción en su composición. En L2, L3 y L4 el hablante utiliza frases definidas para los miembros individuales que se encuentran en el EC UNI, activados a través de la mención de la familia en L1:

(54) Ejemplo [P10.L1.TC]

- L1. ¿No te acuerdas? Sí, hombre, está la familia en un banco de un parque,
- L2. *el marido*,
- L3. *la mujer* -descolorida y ojerosa- duerme roncando y
- L4. *la hija* está sumida en el espanto y aburrimiento.

Igualmente, en el ejemplo (55) de una narrativa NTC, vemos cómo el marco cognitivo activado por la mención de la familia en L2 permite el uso de frases definidas para la misma escena en L3 y L4. Por lo tanto, no es el TC lo que determina el uso de expresiones definidas en estos casos sino su previa activación por parte del escritor:

(55) Ejemplo [P17.L1.NTC]

- L2. sólo música. Está supuestamente que disque una familia feliz en un parque,
- L3. y *la esposa*, que me imagino que es está profunda con otra señora, no sé quien sea, hija, tía, amiga en una banca.
- L4. Está *el esposo* también despierto y

Por último, es interesante que aun sin TC, el participante del ejemplo (53), más arriba, utiliza el demostrativo como si el lector supiera de qué video está hablando, lo cual puede deberse a las dificultades de aislar el TC como variable.

6.7.4 Conclusiones para el estado UNI

En conclusión, en el EC UNI las narrativas escritas siguen los mismos patrones que las narrativas orales, aunque vemos de nuevo un acercamiento al habla nativa por parte de los HNNs en las narrativas escritas, tanto en el uso de FN definidas, donde se observa un uso proporcional idéntico, como en el uso de nombres propios. Al igual que en el estado FAM, en el estado UNI la adherencia a la jerarquía se cumple de manera muy clara.

Los demostrativos presentan un caso destacable en este EC, puesto que únicamente en los datos escritos del presente estudio hay ocurrencias de esta forma de expresión referencial, lo cual sugiere que para los pronombres demostrativos, en particular *esteN*, el estado UNI es el mínimo requerido. Esta conclusión difiere de la propuesta de Gundel et ál. (1993), quienes predicen que solo requiere activación.

No hay evidencia de codificación del factor TC, lo cual es un resultado esperado en este estado, donde se predice el uso de expresiones definidas para referentes que sean familiares en el discurso, ya sea porque son identificables a través de la expresión misma al existir en la memoria a largo plazo, o estén activados por marcos cognitivos.

6.8 Análisis y discusión TYP (GH pronostica IndNP)

6.8.1 Datos escritos HNs vs HNNs

La prueba de chi-cuadrado con FN definidas e indefinidas revela que, en general, no hay diferencias significativas entre los hablantes en las composiciones escritas ($p=0.0683$, Tabla 5.2.52) en la selección de formas de sujeto para el EC de tipo identificable (TYP). Sin embargo, al realizar un análisis de regresión logística, los

resultados demuestran que los HNs tienden a utilizar más expresiones nominales definidas que los HNNs ($p=0.0516$), mientras que los HNNs tienden a utilizar más expresiones indefinidas ($p=0.0277$) (Tabla 5.2.53). Los casos de DefNP utilizados por los HNs representan el 10% de las ocurrencias, lo cual corresponde con cinco instancias: P#1 *los tres caballeros*, P#4 *el policía*, P#9 *la policía* y P#16 *la mujer/los hombres*. Estudiamos estos casos en la sección correspondiente con el TC (Sección 6.7.3).

Se observa que, cuando eliminamos GenNP al comparar únicamente las narrativas escritas entre los HNs y los HNNs, los dos grupos se adhieren a la jerarquía de una manera extremadamente clara, en línea con la hipótesis uno de este estudio:

- Uso de IndNP en HNs, narrativas escritas: 88%
- Uso de IndNP en HNNs, narrativas escritas: 100%

La diferencia entre los dos grupos de hablantes es significativa ($p=0.0277$ Firth). El 12% restante en los HNs está distribuido entre expresiones demostrativas (2%) y expresiones nominales definidas (10%). Describiremos estos casos a lo largo de las siguientes secciones, con ejemplos ilustrativos.

Las expresiones indefinidas son utilizadas para introducir a una entidad en la narrativa por primera vez. Algunos ejemplos son los siguientes:

(56) Ejemplo [P3.L1.TC]

L1. *Una familia* aparentemente respetable, formada por padre, madre e hija, descansan en un parque.

(57) Ejemplo [P15.L2.NTC]

L1. *Un señor* esta con su mujer y suegra en un parque en un día de verano.

6.8.2 Comparación entre datos escritos y datos orales

Según la predicción del estudio de ByLQ, los hablantes utilizarán una FN

indefinida (con o sin determinante) para referirse a un referente cuando se puede suponer que el oyente podrá entender el tipo de entidad que la expresión describe (p. 160). Tanto en los datos orales del estudio de ByLQ como en los escritos de la presente investigación se cumple esta predicción. Además, el EC TYP muestra los resultados más claros en concordancia con la GH, como predicho por la hipótesis uno del actual estudio.

Al introducir la categoría GenNP para la comparación entre los datos escritos del presente estudio con los datos orales de ByLQ, el uso de expresiones nominales indefinidas se ve reducido en ambos grupos de hablantes, aunque se mantiene como la forma más utilizada en este estado. La distribución queda de la siguiente manera:

- HNs narrativa escrita con GenNP: 62% IndNP, 27% GenNP
- HNNs narrativa escrita con GenNP: 69% IndNP, 31% GenNP

Estos resultados con GenNP son muy diferentes a los resultados orales, puesto que en el estudio de ByLQ ningún participante utilizó expresiones generales cuando el referente está en el EC TYP. Como consecuencia, el mayor número de instancias de GenNP en las narrativas escritas es altamente significativo, tanto por parte de los HNs ($p < 0.001$ Firth) como de los HNNs ($p = 0.0019$ Firth). En este caso utilizamos la prueba exacta de Fisher puesto que los resultados tienen una completa separación (ver Capítulo 4, sección 4.4). El ajuste logístico realizado confirma que la escritura disminuye la tendencia de los participantes de utilizar IndNP de manera significativa ($p = 0.0045$, Tabla 5.2.59):

- Uso de IndNP en HNs: 90% oral vs 62% escrito ($p = 0.0001$)
- Uso de IndNP en HNNs: 77% oral vs 69% escrito ($p < 0.0001$)

Cuando analizamos los ejemplos en detalle, vemos que todas las instancias de

GenNP en el estado TYP son utilizadas para introducir a personajes nuevos en la narrativa, además, todas las expresiones generales fueron reclasificadas como indefinidas en este estado. La introducción de algunos personajes con expresiones generales sucede tanto con los HNs, ejemplo (58), como con los HNNs, ejemplo (59):

(58) Ejemplo [P17.L1.NTC]

L32. En esas llega *otro señor* a reclamarle pero

(59) Ejemplo [P37.L2.NTC]

L17. Mientras tanto, *otro hombre* llega y

A parte de *otro señor* y *otro hombre* de los ejemplos anteriores, en este EC los participantes usaron expresiones como *otros dos hombres*, *otra mujer*, *dos hombres nuevos* y *dos muchachos*, entre otras.

En las dos modalidades de producción se dieron casos de expresiones referenciales definidas para el estado TYP, lo cual no es pronosticado por la jerarquía cognitiva, pero algunas de estas instancias codifican el TC en el estudio escrito, como veremos en la siguiente sección. En el estudio oral hay un total de ocho instancias de frases definidas en este EC: tres en los HNs y cinco en los HNNs. ByLQ (p. 161) dicen que cuando un HN utiliza una expresión definida en el estado TYP, lo hace para referirse a entidades genéricas tipo *la mujer ideal*; mientras que los HNNs utilizan frases definidas para introducir al policía en la narrativa, lo cual no se da entre los HNs. En las narrativas escritas hay cinco casos de expresiones definidas, todas pertenecientes a los HNs. El mayor uso de estas expresiones definidas en las narrativas escritas en los HNs no es significativo. Sí lo es el menor uso en los HNNs, ya que es completamente eliminado en las narrativas escritas:

- Uso de DefNP en HNs: 7% oral vs 10% escrito ($p=0.6673$)
- Uso de DefNP en HNNs: 5% oral vs 0% escrito ($p=0.0028$ Firth)

El uso de las expresiones definidas en el estudio escrito muestra varias diferencias con el estudio oral. En primer lugar, contrario al estudio oral, en las narrativas escritas los HNs sí utilizan expresiones definidas para introducir al policía por primera vez, y no únicamente los HNNs. Esto ocurre dos veces, y ambos casos son en narrativas con TC, lo cual podría ser evidencia de codificación lingüística del TC, según la hipótesis cuatro; sin embargo, es más probable que se deba al hecho de que este personaje representa al cuerpo policial de manera genérica, conclusión a la que también llegan ByLQ en su estudio (p. 151):

(60) Ejemplo [P4.L1TC]

L34. *El policía* que está detrás va también al agua con una estupenda pirueta de Charlot

(61) Ejemplo [P9.L1TC]

L20. pero *la polic'ia* lo vio

Las otras tres expresiones definidas por parte de los HNs provienen de dos participantes y, al igual que en el caso del policía, son utilizadas para introducir a personajes nuevos: *los tres caballeros*, *la mujer* y *los hombres*. Estos casos son analizados en la siguiente sección, dentro del marco del TC.

6.8.3 Comparación narrativas escritas TC y NTC

La codificación del TC en el estado TYP fue planteada por la hipótesis cuatro, la cual predice un mayor número de expresiones definidas en narrativas con TC debido a la familiaridad del referente. En el análisis cuantitativo del factor TC excluimos las expresiones generales (por lo que se comparan únicamente los textos escritos); es por esta

razón que no tenemos resultados estadísticos para esta forma referencial. Aun así, podemos realizar un análisis cualitativo para determinar posibles tendencias. En primer lugar, notamos que la mayoría de las expresiones generales en el estado TYP provienen de narrativas con TC, tanto en los textos de los HNs como de los HNNs:

- Uso de GenNP en HNs: 9 (64%) TC vs 5 (36%) NTC
- Uso de GenNP en HNNs: 8 (67%) TC vs 4 (33%) NTC

En este EC, todas las expresiones generales fueron reclasificadas como expresiones indefinidas, lo que conlleva a que el anterior sea un resultado inesperado puesto que se espera un menor uso de expresiones indefinidas en las narrativas con TC al ser los referentes más familiares. De todas maneras, el número de instancias es demasiado reducido como para llegar a resultados conclusivos.

Las tres expresiones definidas por parte de los HNs, mencionadas en la anterior sección, provienen de dos participantes y, al igual que en el caso del policía, son utilizadas para introducir a personajes nuevos: *los tres caballeros*, *la mujer* y *los hombres*. Sin embargo, solo uno de estos casos hace parte de las narrativas con TC (*los tres caballeros*), como vemos en el ejemplo (62b), más abajo. En la composición escrita de este participante se observan dos instancias donde hay una posible codificación lingüística del TC: en primer lugar, en L6 del ejemplo (62a), aunque el sujeto no es analizable en este estudio al estar en posición de objeto, notamos que el hablante introduce a la mujer coqueta por primera vez en el discurso con la expresión definida *la dama*. Más abajo, en L18 de (62b), el hablante describe la escena en la cual Chaplin pelea con los dos hombres del bar a través de la expresión definida *los tres caballeros*. Solo al existir TC podemos saber a cuáles caballeros hace referencia esta expresión. El referente

los tres caballeros en L18 se encuentra en el EC de tipo identificable al ser la primera mención en la narrativa:

(62) a. Ejemplo [P1.L1.TC]

L6. las damas siguen dormindas para ir detras de *la dama*

b. Ejemplo [P1.L1.TC]

L18. Sin embargo la escena antes de la ultima en que *los tres caballeros* estan peleando de forma animada

Los otros dos casos de expresiones definidas en este estado pertenecen a un mismo participante. Aunque se trata de una composición NTC, los personajes de la mujer en (63a) y de los hombres en (63b) son introducidos en la narrativa asumiendo familiaridad, aunque sea la primera vez que aparecen en el texto. Esto puede atribuirse a la dificultad de aislar el TC como factor, en contra de la hipótesis cuatro:

(63) a. Ejemplo [P16.L1.NTC]

L5. Y luego aparece *la mujer* que no puede dormir por los ronquidos de las otras personas.

b. Ejemplo [P16.L1.NTC]

L16. Luego llegando *los hombres* más con botellas bebiendo y

Otro ejemplo por destacar con respecto al TC es el siguiente, en el cual el HN hace explícita la existencia de este factor en su introducción, y por lo tanto utiliza expresiones definidas para los personajes. En L1 el sujeto no es analizable en este estudio por ser el argumento de un verbo copulativo y por lo tanto obligatorio; sin embargo, claramente el escritor asume que el referente *la familia* hace parte del TC entre él y el lector en su pre-introducción. Las subsiguientes menciones de los miembros de la familia también son expresiones definidas, pero están en el EC UNI, al pertenecer al marco

cognitivo de la familia (ver Sección 6.7):

(64) Ejemplo [P10.L1.TC]

L1. ¿No te acuerdas? Sí, hombre, está *la familia* en un banco de un parque,

L2. el marido,

L3. la mujer -descolorida y ojerosa- duerme roncando y

L4. la hija está sumida en el espanto y aburrimiento.

Por último, hay un solo caso de demostrativos en este estado y se utiliza para introducir al personaje de la mujer coqueta. El uso del demostrativo medial *esa N* en L6 se refiere a la mujer coqueta, a quien el hablante introduce en ese momento. El hecho de que la narrativa no tenga TC requiere que la expresión demostrativa haga referencia a una entidad que sea por lo menos familiar. Al no ser este el caso, el texto presenta una ruptura cohesiva ya que el demostrativo no tiene ningún nodo al cual hacer referencia:

(65) Ejemplo [P16.L1.NTC]

L6. A *esa otra mujer* que pasa A lado del hombre para coquetearle y sugiriendo pedirle algo votarle algo.

Es relevante notar que el TC, en varias de las composiciones escritas, es establecido no a través de la codificación lingüística de las formas referenciales de sujeto, sino a través de una pre-introducción en las narrativas. A través de esta estrategia, el hablante reconoce que su interlocutor ficticio ha visto la película, confirmando de esta manera el TC entre ellos. Sin embargo, al comenzar la narrativa, y tal vez debido a la imposibilidad de recibir retroalimentación en la modalidad escrita de la tarea, procede a introducir a los personajes con la forma esperada, esto es, con expresiones indefinidas. Cuando los participantes establecen el TC en la pre-introducción, vemos que ocurren una o más de las siguientes situaciones: en algunos casos, la pre-introducción es el único indicio de que el hablante reconoce el TC con su interlocutor ficticio ya que, después de

la introducción, no hay ninguna otra evidencia lingüística de codificación del TC. En el siguiente ejemplo la expresión *como sabes* es la única evidencia de que se ha reconocido TC entre escritor y lector:

(66) Ejemplo [P6.L1.TC]

L1. Es una película del gran Charles Chapling, y como sabes es una película de cine mudo. Y no hace sino

En segundo lugar, algunos participantes van a introducir a los personajes con expresiones definidas, pero únicamente en la pre-introducción, como lo vimos en el ejemplo (64), más arriba, parcialmente reproducido aquí como (67):

(67) Ejemplo [P10.L1.TC]

L1. ¿No te acuerdas? Sí, hombre, está la familia en un banco de un parque,

También sucede que, al haberse establecido el TC en la introducción, este factor no afecta la selección de la forma sino el EC. En el ejemplo (44), reproducido aquí en (68), observamos que el participante correctamente utiliza una frase indefinida para introducir al hombre coqueto en la narrativa en L2. Sin embargo, el EC de este referente ya no es TYP sino FAM debido a su mención anterior en L1; esto es, el referente ya ha sido llevado a la memoria por el escritor durante la pre-introducción de la composición:

(68) Ejemplo [P9.L1.TC]

L1. Te acuerdas de la película que vimos? Si, la del hombre que quer'ia ligar con un chica joven, en el

L2. parque.... Que no te acuerdas? quieres que la recuerde? Pues el resumen es que hay *un hombre aburrido*

Estas pre-introducciones donde se establece el TC ocurren en tres ocasiones entre los HNs. Aunque también ocurren en igual proporción entre los HNNs, el TC es más evidente y directo en las composiciones nativas. En los ejemplos (69) y (70) vemos que,

después de la introducción inicial estableciendo el hecho de que el amigo ficticio ha visto la película, en ningún otro momento de la narrativa hay evidencia de codificación lingüística del TC. En el ejemplo (69) el HNN pasa inmediatamente a introducir a los personajes con expresiones indefinidas, clasificadas en el estado TYP al ser la primera mención, y lo mismo sucede en el ejemplo (70):

(69) Ejemplo [P25.L2.TC]

L1. Fue chistoso ver a la película silenciosa de Chaplin, ¿no? Lo que sucedió fue que *una familia* estuvo en

(70) Ejemplo [P28.L2.TC]

L1. Oye, ¿viste ese video loco del cortometraje sin sonido y color? No sé si me gustó. Bueno, [ø] empieza
L2. en un parque y hay lo que llaman *una familia* sentada en el banco.

En el último ejemplo, más abajo, vemos un comportamiento similar al uso nativo ya que el HNN utiliza la introducción para establecer el TC e introduce a los personajes en L2 con expresiones definidas. Sin embargo, al no estar en posición de sujeto, estas expresiones no son analizables en este estudio:

(71) Ejemplo [P26.L2.TC]

L1. Acabo de mirar el videoclip de la coqueta y el travieso Charlie que me recomendaste. ¡Qué chistoso!
L2. Cuando [ø] comenzó con la señora durmiéndose al lado de su esposo, yo no esperaba lo que iba a

6.8.4 Conclusiones para el estado TYP

En conclusión, el estado TYP muestra los resultados más claros de adherencia a la jerarquía, así como sucede con el estudio oral, y en línea con la hipótesis uno del actual estudio. Al comparar las dos modalidades de producción vemos que la adherencia es menos fuerte en la escritura debido al incremento en el uso de expresiones generales, las

cuales no fueron utilizadas en las narrativas orales. Sin embargo, como en el presente estudio todas las expresiones generales en este EC se consideran indefinidas, podemos concluir que no hay diferencias en la adherencia a la jerarquía entre los dos estudios.

En cuanto al TC, en la hipótesis cuatro se plantea un mayor uso de expresiones definidas en el estado TYP, y aunque los resultados cuantitativos no muestran evidencia significativa de que los hablantes codifiquen lingüísticamente este factor, hay varios ejemplos cualitativos que revelan su influencia a través de otras estrategias, como lo es el uso de pre-introducciones en las composiciones con TC.

6.9 Conclusiones análisis cualitativo

Los resultados obtenidos en la presente investigación con datos de narrativas escritas muestran que, en general, tanto los HNs como los HNNs siguen las mismas tendencias que en las narrativas orales, cuyos datos provienen del estudio llevado a cabo por ByLQ (2012). Principalmente, se observa que las elecciones de los hablantes se adhieren a la GH, aunque esta adherencia es menos fuerte en la modalidad escrita. Al ser la escritura una modalidad más cuidada, hay un efecto significativo sobre el uso y distribución de las formas referenciales de sujeto al observarse un mayor número de instancias de expresiones más elaboradas, así como un acercamiento al habla nativa por parte de los HNNs.

Con respecto al TC, cuantitativamente no se obtuvieron resultados significativos que demuestren la influencia directa de esta variable, es solo a través de un análisis cualitativo de los datos generales, así como algunas instancias individuales, donde

observamos algunos posibles efectos de este factor sobre las narrativas escritas de los hablantes.

En el siguiente capítulo (Capítulo 7) se presenta un resumen junto con un análisis general de los resultados con respecto a las preguntas de investigación e hipótesis planteadas en el Capítulo 4.

CAPÍTULO 7

7 CONCLUSIONES

7.1 Resumen

En esta investigación se ha analizado la selección y uso que hacen los hablantes nativos y no-nativos de español de las expresiones referenciales de sujeto en narrativas escritas y se han comparado los datos con los datos orales del estudio de Blackwell y Lubbers Quesada (ByLQ) (2012). Se tomaron como base principal las propiedades cognitivas relacionadas con el nivel de atención y memoria con respecto al referente. La Givenness Hierarchy (GH) propuesta por Gundel, Hedberg y Zacharski (1993) presenta una jerarquía implicativa de estados cognitivos (ECs) que restringen las formas de sujeto y su distribución en el discurso. Estos ECs se basan en el nivel de recuperabilidad del referente según el estado de atención y memoria del oyente/lector, donde las formas más mínimas corresponden con referentes más prominentes en el discurso. El presente trabajo utiliza como marco teórico de referencia esta jerarquía, adaptada al español por ByLQ (2012).

A lo largo de esta investigación, y en línea con estudios previos acerca de las expresiones referenciales de sujeto en el español, se ha visto que el uso y selección de las expresiones referenciales están restringidas no solo por estos factores cognitivos, sino también por factores sintácticos relacionados con la posición del sujeto, tal como codificado por la hipótesis de posición del antecedente de Carminati (2002), así como por la función pragmática del sujeto, donde el sujeto nulo se relaciona con el

tópico/continuidad de tópico y tiene como función discursiva presentar información vieja o compartida, mientras que el sujeto explícito se relaciona con el foco/cambio de tópico y tiene como función introducir nueva información en el discurso o marcar un contraste (Lubbers Quesada, 2015).

La presente investigación tiene un doble objetivo: en primer lugar, se comparan las elecciones de los hablantes según la modalidad de producción (oral vs escrito) y, en segundo lugar, se analiza el terreno común (TC) como variable independiente para determinar su posible efecto en las elecciones de los escritores. La motivación principal provino de una falta de estudios de adquisición de segunda lengua que investigaran los factores cognitivos según la modalidad del discurso. En su disertación doctoral, Saunders reconoce este vacío al decir: “No claim has been made in this dissertation regarding written narratives. It is likely that the references used would differ somewhat...It would be interesting to note, nonetheless, the similarities and differences between the oral and written modalities” (1999:183-184).

En este trabajo se parte de la base de que hay un mayor costo de procesamiento en la producción escrita con respecto a la producción oral; aunque, como observado por McCutchen (1996) y Bourdin y Fayol (1994, 2002), las diferencias son notables en adultos únicamente cuando aparecen factores que afectan la carga cognitiva, por ejemplo, que se imponga un límite de tiempo para la redacción (Bourdin y Fayol, 2002), o en la producción de una segunda lengua (Schoonen, van Gelderen, de Glopper, Hulstijn, Simis, Snellings y Stevenson, 2003). Ambos factores están presentes en este estudio.

Respecto al TC, al ser esta una variable que hace parte de los factores pragmático-cognitivos que tanto Gundel et ál. (1993) como ByLQ (2012) discuten en sus

investigaciones (aunque no se aíslan), en este trabajo se pretendía determinar hasta qué punto es posible manipular el TC en las narrativas escritas y observar su posible influencia en la selección de expresiones por parte de los escritores.

En las siguientes secciones se resumen los principales resultados obtenidos, los cuales fueron descritos en detalle en los Capítulos 5 (análisis cuantitativo) y 6 (análisis cualitativo). Se comienza revisando los resultados de las narrativas escritas con respecto a las orales, para luego pasar a las conclusiones de los resultados del TC en las narrativas escritas. Por último, se hacen unas observaciones generales con respecto a la adquisición del español como L2 y la cognición.

7.2 El discurso oral vs el discurso escrito

Los datos escritos fueron recolectados por la que escribe estas líneas para el presente estudio; mientras que los datos orales provienen del trabajo de ByLQ (2012).

En general, se registró un menor número de expresiones referenciales de sujeto en el estudio escrito (1.106) que en el estudio oral (1.214). Esta diferencia proviene de los HNs al producir 276 menos expresiones referenciales en el discurso escrito que en el oral; mientras que los HNNs tienen un comportamiento opuesto al producir 168 más instancias en el estudio escrito que en el estudio oral (ver Tabla 5.1.3, Cap. 5).

La hipótesis 1 de este estudio se ve confirmada ya que los datos revelan que tanto los HNs como los HNNs se adhieren a la GH en el discurso escrito. Esto se debe a que la GH se propone como una jerarquía universal por Gundel et ál. (1993), quienes afirman que la GH es adecuada para describir el uso de las expresiones referenciales en las cinco

lenguas examinadas por los autores (inglés, español, chino mandarín, ruso y japonés)²³. No todas las lenguas necesitan de todos los estados, en el español, estos se ven reducidos a cinco, ya que no se diferencia entre los estados REF y TYP. La universalidad de la GH proviene de la universalidad de las habilidades cognitivas que la rigen, las cuales están relacionadas con las nociones de cohesión y coherencia textual, prominencia, enfoque, memoria, atención y la implicación de los participantes, nociones todas que guían las elecciones de los HNs y de los HNNs (Lubbers Quesada, 2005:153). En el presente análisis, los resultados demuestran que dicha universalidad también rige la modalidad escrita, tanto para HNs como HNNs.

Como predicho por la hipótesis 1, los estados colocados en los extremos de la jerarquía (INF / TYP) son aquellos que demuestran una mayor adherencia a la GH, aunque los estados FAM y UNI también se adhieren a la jerarquía con claridad. La mayor variabilidad en los datos obtenidos se observa en los estados ACT y ACTNR. Los resultados del estudio escrito se desvían de las hipótesis propuestas por ByLQ, pero estos no son del todo diferentes de aquellos del estudio oral. En el estado ACT, la hipótesis 2B del estudio oral de ByLQ predecía el uso de pronombres explícitos; sin embargo, la forma referencial más utilizada en las narrativas escritas (para los HN y HNN) y en las orales (para los HNNs y en segundo lugar para los HNs) fueron las FNs definidas. En el estudio oral, únicamente el 41% de los HNs seleccionó un pronombre explícito y este porcentaje se ve reducido al 23% en las narrativas escritas de los HNs. En el estado ACTNR la hipótesis 2C del estudio de ByLQ predecía el uso de FNs generales, y también

²³ En un estudio subsecuente, Gundel, Bassene, Gordon, Humnick y Khalfaoui (2010) aplicaron la GH a otras cuatro lenguas, confirmando de nuevo la universalidad de la GH. Las lenguas fueron las siguientes: eegimaa, kumyk, ojibwe, y el árabe de Túnez.

aquí vemos que menos de la mitad de los HNs de las narrativas orales seleccionaron esta forma (44%) y solo el 10% de los estudiantes avanzados. Las FNs definidas ocuparon el segundo lugar para los dos grupos de hablantes del estudio oral de ByLQ y fueron la forma más seleccionada para los dos grupos en las narrativas escritas. Tal vez esta variabilidad en los ECs ACT y ACTNR provenga del hecho de que, en estos estados, los niveles de atención y memoria son más difíciles de evaluar ya que la recuperabilidad del referente se encuentra en una situación intermedia en la jerarquía, donde hay varios factores que intervienen durante el proceso de resolución referencial, entre ellos, la distancia con la mención anterior, los factores pragmáticos y la función del referente en la oración. Adicionalmente, cuando hay tres o cuatro referentes que compiten en una misma escena por el papel de antecedente, se observa que en estos ECs hay una mayor dificultad por parte de los hablantes/escritores para seleccionar una expresión referencial de manera exitosa con respecto al oyente/lector. Es aquí cuando el uso de expresiones más elaboradas, como lo son las FNs definidas y los nombres propios, contribuyen hacia una mayor cohesión textual, la cual disminuye el costo de procesamiento cognitivo al eliminar referentes de manera explícita y precisa.

Las desviaciones que se observan en los dos estudios para los estados ACT y ACTNR sugieren que para estos dos ECs, las formas esperadas son las siguientes: ACT {Overt, DefNP} y ACTNR {GenNP, DefNP} o, simplemente, ACTNR {DefNP}, puesto que en la presente investigación se propone la eliminación de GenNP como categoría, la cual fue introducida por ByLQ (2012) para diferenciar entre las FNs y expresiones no pronominales de carácter genérico como *los dos*, *uno de ellos*, etc.

Con respecto al EC FAM, la predicción de la GH revisada para el español también es por el uso de FNs definidas. La diferencia entre ACTNR y FAM tiene que ver con el nivel de especificación semántica: en el EC ACTNR los hablantes utilizan expresiones más genéricas (aunque siguen siendo definidas) como *el otro*, *los dos hombres*, *la otra mujer*; mientras que en el estado FAM las FNs tienen un mayor nivel de especificación, por ejemplo, *el hombre*, *la joven*, *el policía*, lo cual contribuye a la recuperabilidad del referente en un EC donde los niveles de atención y memoria son más bajos. En este EC observamos un comportamiento contrario entre HNs y HNNs puesto que el primer grupo utiliza menos FNs definidas y el segundo hace un mayor uso de ellas. Sin embargo, el uso de nombres propios tiene una tendencia opuesta, y al analizar las dos formas conjuntamente vemos que no hay ninguna diferencia entre las modalidades de producción escrita y oral.

En el estado UNI, el cual también predice el uso de FNs definidas, también nos encontramos con un comportamiento opuesto entre HNs y HNNs: los HNs utilizan menos FN definidas y más nombres propios, mientras que los HNNs utilizan más FN definidas y menos nombres propios. Sin embargo, ninguno de estos resultados es significativo; por lo tanto, no se observan diferencias entre las narrativas orales de ByLQ y las escritas del presente estudio en el EC UNI, ya que en los dos estudios las formas más elaboradas fueron las más utilizadas.

En el estado TYP, donde se espera el uso de FNs indefinidas, observamos que los dos grupos de hablantes hacen un menor uso de esta forma en las narrativas escritas y un mayor uso de expresiones generales. En este estado, los HNs del estudio escrito utilizan más FNs definidas que los HNNs. Este comportamiento es muy diferente al de las

narrativas orales, puesto que en ellas no se presentó ningún caso de expresiones generales. La menor proporción de FNs indefinidas en el estudio escrito conlleva a que se observe una menor adherencia a la GH; sin embargo, al eliminar GenNP al comparar las narrativas escritas entre HNs y HNNs, todas las expresiones generales se vuelven a clasificar como IndNP. Esto implica que, en las narrativas escritas, los HNNs observen el 100% de adherencia a la GH y los HNs el 88%. El 12% restante está distribuido entre DefNP y Dem, confirmando la hipótesis 1 con respecto a la mayor adherencia en los estados extremos de la jerarquía.

Concluimos entonces que, aunque en las narrativas escritas se siguen las predicciones de la GH, los datos de este estudio también revelan que la adherencia a la GH se ve afectada por la modalidad de producción, siendo la adherencia menos fuerte en el discurso escrito. Evidencia de esto se encuentra en el mayor uso de expresiones más elaboradas en las narrativas escritas (FNs definidas y nombres propios) y en un menor uso de expresiones más mínimas (sujetos nulos y pronombres explícitos). Por lo tanto, vemos que el poder de predicción de la GH es mayor en las narrativas orales que en las escritas, confirmando la hipótesis 3 de este estudio. En la tabla 7.2.1, más abajo, se presenta un resumen de los datos que demuestran esta tendencia (las casillas sombreadas indican un menor uso de las formas Null y Overt y un mayor uso de las formas DefNP y PrpN en las narrativas escritas. Los porcentajes resaltados en negrilla son estadísticamente significativos – ver Capítulo 5). En esta tabla se resumen únicamente los ECs INF, ACTR y ACT, puesto que es en estos donde la predicción es por expresiones de sujeto más mínimas como Null y Overt:

Tabla 7.2.1 – DefNP y PrpN vs Null y Overt en los ECs INF, ACT y ACTR

Habla L1/L2	Modo de producción (%)								EC
	Null		Overt		DefNP		PrpN		
	Oral	Escrito	Oral	Escrito	Oral	Escrito	Oral	Escrito	
L1	91	81	5	4	2	11	1	3	INF
L2	65	80	26	9	3	6	4	5	
L1	82	55	8	5	2	20	1	16	ACTR
L2	59	54	34	15	2	18	5	7	
L1	12	6	41	23	26	53	13	9	ACT
L2	2	9	33	23	44	53	17	12	

El uso más elaborado de las expresiones referenciales en las narrativas escritas también está relacionado con la necesidad de crear cohesión en un texto. La cohesión es una noción cognitivo-discursiva que se ve afectada por la selección de la forma referencial: cuando un hablante/escritor no utiliza la forma adecuada en una narrativa, es posible que se produzca una ruptura en la cohesión del texto, debido a que no se forma un vínculo entre el elemento cohesivo (la anáfora) y el nodo (el antecedente). Saunders (1999:178) indica que, al ser el español una lengua [+pro-drop], el mantenimiento de un personaje en la narrativa debe hacerse a través del uso de un sujeto nulo. Si se utiliza una expresión más elaborada, esta crea un nuevo nodo y, por lo tanto, se producen menos vínculos en la narrativa, lo cual conlleva a una menor cohesión textual ya que el oyente no es capaz de conectar el nuevo nodo con el anterior. Sin embargo, en los datos escritos encontramos algunos casos donde el sujeto nulo puede causar un mayor costo cognitivo en el proceso de identificación del referente. Así un referente sea accesible cognitivamente con una expresión más mínima, si el procesamiento es más costoso que con una expresión más elaborada vemos que los escritores tienden a seleccionar la opción más elaborada, lo cual puede deberse al hecho de que el escritor no tiene la posibilidad de asegurarse del nivel de atención y memoria del lector. Esto suele suceder en casos de posible ambigüedad al haber más de un referente posible en la escena y donde la

morfología verbal no permite la identificación. En el ejemplo (17) de la sección 6.2, reproducido aquí como (1), el referente *Charlot* en L25 se clasifica en el EC ACTR porque un sujeto nulo hubiera permitido la recuperación del referente, sin embargo, al haber dos hombres más en escena se requiere de un esfuerzo cognitivo para hacer la correcta identificación, este esfuerzo se ve atenuado con el uso de una forma más elaborada, en este caso un nombre propio:

(1) Ejemplo [P5.L1TC]

L23. Por otro lado, *Charlot* continua en el banco

L24. y dos hombres se sientan a su lado.

L25. *Charlot* en un descuido aprovecha para beberse la bebida de uno de ellos con lo cual se genera un

Por el contrario, los HNNs tienden a sobreutilizar las formas elaboradas causando rupturas en la cohesión, como lo vimos en el ejemplo (13) de la sección 6.1, reproducido aquí como (2), con la repetición del pronombre personal *ella* en L21. Estos casos no ocurren con los HNs, asociándose por lo tanto con falta de adquisición del sujeto nulo:

(2) Ejemplo [P32.L2NTC]

L20. *Ella* le vendó los ojos al hombre y

L21. *ella* se fue.

También vemos que, en la modalidad escrita, los HNNs se acercan al habla nativa, como predicho por la hipótesis 2. Esto puede deberse a que los escritores tienen más oportunidad de pensar en sus respuestas y corregirlas, así las narrativas escritas sean de carácter espontáneo, como lo son en esta investigación. En la siguiente tabla (Tabla 7.2.2), vemos un resumen del acercamiento al habla nativa por parte de los HNNs. En la tabla se exponen únicamente aquellas formas predichas por la GH, con excepción de

ACTNR, donde también se incluye DefNP al ser la forma más utilizada en las narrativas escritas de este estudio:

Tabla 7.2.2 - Acercamiento al habla nativa por parte de los HNNs

INF – Null: <ul style="list-style-type: none"> ○ Oral: 91% (L1) vs 65% (L2) ○ Escrito: 81% (L1) vs 80% (L2) 	ACTR – Null: <ul style="list-style-type: none"> ○ Oral: 82% (L1) vs 54% (L2) ○ Escrito: 54% (L1) vs 55% (L2)
ACT – Overt: <ul style="list-style-type: none"> ○ Oral: 41% (L1) vs 33% (L2) ○ Escrito: 23% (L1) vs 23% (L2) 	ACTNR – GenNP: <ul style="list-style-type: none"> ○ Oral: 44% (L1) vs 10% (L2) ○ Escrito: 28% (L1) vs 13% (L2)
ACTNR – DefNP: <ul style="list-style-type: none"> ○ Oral: 30% (L1) vs 26% (L2) ○ Escrito: 44% (L1) vs 54% (L2) 	FAM – DefNP/PrpN: <ul style="list-style-type: none"> ○ Oral: 83% (L1) vs 97% (L2) ○ Escrito: 86% (L1) vs 92% (L2)
UNI – DefNP: <ul style="list-style-type: none"> ○ Oral: 83% (L1) vs 53% (L2) ○ Escrito: 68% (L1) vs 68% (L2) 	TYP – IndNP*: <ul style="list-style-type: none"> ○ Oral: 90% (L1) vs 77% (L2) ○ Escrito: 62% (L1) vs 69% (L2) *No incluye GenNP reclasificado como IndNP

Como vemos en la Tabla 7.2.2, en varias instancias el acercamiento al habla nativa elimina enteramente las diferencias entre los HNs y los HNNs. En el estudio oral de ByLQ (2012) se encontraron diferencias en todos los ECs, entre los HNs y los HNNs, en su uso y selección de expresiones referenciales de sujeto. Estas diferencias son significativas en los siguientes estados: INF (Null, Overt y PrpN: p entre 0.0117 y $<.001$, HNs vs HNNs —todos los grupos²⁴) (p. 154); ACTR (Null: p entre 0.0080 y $<.0001$, HNs vs HNNs —todos los grupos; Overt: $p=0.0004$, HNs vs HNNs avanzados) (p. 155); ACTNR (GenNP: $p=0.0013$, HNs vs HNNs avanzados) (p. 158); FAM²⁵ (PrpN p entre 0.0043 y 0.0009, HNs vs HNNs —todos los grupos) (p. 159), y UNI (DefNP chi

²⁴ En su artículo, ByLQ (2012) en algunas ocasiones dan un rango de p entre todos los aprendientes (principiantes, intermedios y avanzados) y los HNs, sin indicar el valor de p específico entre cada grupo de HNN y los HNs, y en algunos casos sin especificar entre formas referenciales.

²⁵ ByLQ (2012:159) indican que los HNs seleccionan DefNP en muchas más ocasiones que todos los HNNs en su estudio, pero no indican si esta diferencia es significativa: L1 67% vs HNNs avanzados 60%.

cuadrado $p=0.0167$, HNs vs HNNs —todos los grupos). En el presente estudio, estas diferencias significativas son eliminadas en todos los ECs entre los dos grupos de participantes (HNs y HNNs), para las formas pronosticadas por la GH. En muchos casos, proporcionalmente no hay ninguna, o muy mínima, diferencia en la selección de las formas referenciales, como es el caso en los estados INF (L1 81% vs L2 80%), ACT (L1 23% vs L2 23%), ACTNR (L1 54% vs L2 55%) y UNI (L1 68% vs L2 68%). Esto revela un acercamiento al habla nativa cuando la modalidad de producción es escrita, posiblemente porque una modalidad más cuidada resulta en un discurso más cercano al potencial máximo de las habilidades lingüísticas, como también lo sugiere Ellis (1985:49) al analizar la manera en cómo los aprendientes de una L2 monitorean su lenguaje en un discurso planeado.

El acercamiento al habla nativa en las narrativas escritas también se observa en la codificación de las propiedades pragmáticas de las formas referenciales de sujeto por parte de los HNNs. Por ejemplo, en el estado INF, ByLQ no encontraron datos por parte de los HNNs donde estos seleccionaran una forma más elaborada en lugar de una forma mínima por razones pragmáticas (p. 154-155); sin embargo, la codificación de los factores pragmáticos sí se registró en los HNNs en las narrativas escritas y esta es una diferencia importante puesto que es evidencia de que, en el lenguaje escrito, el escritor tiene más oportunidad de producir un lenguaje más cercano al nativo. En el ejemplo (5) de la sección 6.1, reproducido aquí abajo como (3), el HNN repite el nombre propio (*Charlie*) como énfasis pragmático al indicar la nueva escena donde Chaplin ve al hombre y a la mujer en L14. Como mencionando en la sección 6.1, sabemos que este hablante ha adquirido el sujeto nulo a través de su uso correcto en L12 y L13:

(3) Ejemplo [P24.L2TC]

- L11. Ahora vemos que *Charlie Chaplin* viene caminando por el parque y
- L12. sin darse cuenta [\emptyset] camina arriba de un aspersor y
- L13. [\emptyset] se moja todo.
- L14. *Charlie* ve al hombre con la mujer misteriosa sentados en otro banco del parque y

Otra diferencia por resaltar entre las narrativas escritas y las orales es con respecto al uso de los demostrativos. Esta diferencia puede repercutir en las formas esperadas para el EC UNI en la GH. En las dos modalidades de producción se encontraron pocos casos de demostrativos y estos son usados con mayor frecuencia por los HNs (ver Tabla 6.6.1). Gundel et ál. (1993) dicen que, en español, los demostrativos *aquel/ese/este/este N* solo requieren activación y para los demostrativos *ese N/aquel N* el estado FAM es suficiente. En los datos del presente estudio se encontraron tres instancias de demostrativos en el EC UNI, lo cual no ocurrió en el estudio oral, y tampoco son reportados en los datos de Gundel et ál. Vimos en la Sección 6.6 que en todas las ocasiones las expresiones demostrativas son utilizadas para comenzar la descripción de la película, como en el siguiente ejemplo:

(4) Ejemplo [P22.L2TC]

- L1. *Este videoclip* comienza mostrando una familia descansando el parque.

Las tres instancias se clasifican dentro del EC UNI porque el referente es identificable a través de la sola expresión, por lo que el demostrativo actúa como una expresión definida capaz de seleccionar una entidad única. Este tipo de referencia exofórica lleva al lector fuera del texto y por lo tanto no se considera cohesiva (no une dos elementos en un texto). No obstante este tipo de expresión referencial no se considere esperada en el estado UNI al no ocurrir con frecuencia, es una forma adecuada y satisface

las características de este EC, lo cual conlleva a concluir que, para los demostrativos, el estado UNI es suficiente puesto que lo importante es que, a través de la expresión referencial, se pueda hacer una asociación única con el referente.

Como mencionado anteriormente, en este trabajo la perseveración no fue considerada como una variable independiente. Sin embargo, en los datos se encontró evidencia de que la repetición de ciertas formas y de ciertas estructuras tuvo un posible efecto en la selección de las formas referenciales. Vimos algunos ejemplos con respecto a la repetición del sujeto inanimado *la película* en el estado INF, así como ejemplos de paralelismo estructural de FNs definidas en algunas ocasiones. Es posible que este paralelismo permita una mayor cohesión, lo cual reduce el costo de procesamiento cognitivo, pero se requiere de un análisis más enfocado en este factor para llegar a conclusiones definitivas al respecto.

Hay claras diferencias entre una narrativa escrita y el discurso oral. A pesar de estas diferencias entre las dos modalidades de producción, no se considera necesario plantear dos jerarquías según el tipo de discurso, ya que hemos observado que en la modalidad escrita se siguen los principios básicos de la GH, pero sí se considera importante que la GH abarque aquellas formas que se predicen para cada EC provenientes también del lenguaje escrito. De esta manera, la jerarquía puede incorporar las diferentes modalidades de lenguaje disponibles a los hablantes, ampliando así su poder de predicción, específicamente el uso de FNs definidas para los estados ACT y ACTNR. Sin embargo, queda claro que la GH no explica todas las restricciones que se presentan durante el discurso para la selección de las formas referenciales. Principalmente, hemos visto que hay factores pragmáticos (como lo son el foco contrastivo, o el cambio de escena o nueva

acción) además de factores cognitivos de atención y memoria, que explican el uso de expresiones más elaboradas, cuando la jerarquía predice formas más mínimas. La GH tampoco considera la selección de ciertas formas para crear cohesión textual y disminuir el costo cognitivo durante el proceso de resolución referencial.

Para concluir, los resultados obtenidos en esta investigación nos permiten contestar la primera pregunta de investigación de la siguiente manera: para las formas pronosticadas por la GH adaptada al español por ByLQ, no se encontraron diferencias significativas en la producción escrita entre HNs y HNNs, con excepción del EC TYP (los HNs utilizan más DefNPs). Principalmente, las tendencias revelan que tanto HNs como HNNs siguen los principios de la GH, aunque lo hacen adhiriéndose con menor fuerza en las narrativas escritas que en las orales de ByLQ, al seleccionar con mayor frecuencia expresiones más elaboradas y al reducir el uso de expresiones más mínimas. Los estados con mayor claridad siguen siendo aquellos encontrados en los extremos de la jerarquía, INF y TYP, aunque se manifiesta también en los estados UNI y FAM. Además, hay evidencia de un acercamiento al habla nativa por parte de los HNNs debido a la mayor atención y cuidado que se presta en la producción escrita. Se confirman de esta manera las hipótesis 1, 2 y 3 planteadas para esta disertación.

7.3 El papel del Terreno Común en la narrativa escrita

Los objetivos con respecto al TC en esta investigación se establecieron en relación con la posibilidad de manipular este factor para aumentar o disminuir el nivel de atención y memoria de los participantes en las narraciones escritas y observar cuáles ECs, y correspondientes expresiones referenciales, se verían afectados, así como determinar

posibles diferencias entre HNs y HNNs. No se hizo una comparación con las narraciones orales puesto que en el estudio de ByLQ no se aisló esta variable y por lo tanto no se estableció de manera consistente el TC entre el participante y el entrevistador; aunque, en la mayoría de los casos, a los participantes se les dijo que el entrevistador no había visto la película (M. Lubbers Quesada, comunicación personal, marzo 2018).

No se obtuvieron suficientes resultados significativos que nos permitan llegar a conclusiones decisivas; por lo cual, no podemos confirmar la hipótesis 4 con respecto a la posibilidad de aislar el TC como variable independiente. Aun así, el análisis cualitativo revela algunas tendencias de los hablantes al intentar codificar el TC que demuestran que este hace parte integral de la codificación de los ECs de la GH.

Como vimos en el Capítulo 2, para establecer el TC entre dos interlocutores y poder utilizar de manera adecuada un referente definido que codifique un conocimiento compartido, se necesita de tres procesos de copresencia: (1) la pertenencia a una comunidad; (2) la copresencia física, y (3) la copresencia lingüística (Clark y Marshall, 1981). En el trabajo de Schreuder y Buttrick (1992) que revisamos en el Capítulo 2, sí se pudo comprobar la influencia del TC en las selecciones de los participantes porque en aquella investigación se estudiaron las expresiones deícticas, las cuales se basan en la copresencia física y este es el más fuerte de los tres procesos heurísticos para determinar la existencia del TC (Clark y Marshall, 1981:38). Cuando las expresiones referenciales no son de carácter deíctico, el TC se basa, principalmente, en la copresencia lingüística y en el conocimiento mutuo que proviene del pertenecer a una comunidad. Esto lo vemos claramente en Gundel (2010:154), donde la autora propone una serie de pautas o

directrices para la codificación de los diferentes ECs y vemos que el TC está integrado de la siguiente manera en algunos de los estados de la GH:

- En el estado INF, el referente está en foco si el oyente está mirando con gran atención al referente o si este ha sido introducido en una posición prominente en el enunciado anterior. Esto implica bien sea una copresencia física o una copresencia lingüística.
- En el estado ACT, el referente está presente en el contexto extralingüístico inmediato o ha sido mencionado en la oración inmediatamente precedente. De nuevo tenemos copresencia física y lingüística.
- En el estado FAM, el referente hace parte del conocimiento general cultural mutuo o ha sido mencionado en el discurso precedente. Esto implica una copresencia lingüística y la pertenencia a una comunidad.

Vemos entonces que el TC es establecido a través de diferentes procesos heurísticos de copresencia y es de esta manera que este factor afecta la selección de las formas, por ejemplo, el uso de una forma mínima asume conocimiento mutuo con respecto al referente. ByLQ (2012:162) dicen que el uso de expresiones referenciales mínimas es motivado no solamente por el EC, sino también por otros factores, entre ellos las presuposiciones que se hacen a través del conocimiento mutuo del discurso anterior. Aunque las investigadoras hablan del discurso anterior, este conocimiento mutuo se puede extender más allá del discurso y ser afectado por factores como el haber visto la película juntos (copresencia física), el que un referente sea conocido a raíz de conocimiento por pertenencia a una comunidad (como en el caso del personaje de Chaplin) o incluso por las instrucciones dadas para completar una tarea en una encuesta.

El análisis cualitativo de las narrativas escritas revela tres comportamientos con respecto al TC, que demuestran su influencia en el discurso:

1. El escritor utiliza una pre-introducción en la narrativa como estrategia para establecer TC con el lector.
2. En algunos casos, el TC no afecta la forma referencial sino el EC de dicha forma.
3. El lector utiliza el TC en su proceso de resolución referencial.

La primera tendencia es casi exclusiva de los HNs y ocurrió con tres de los 10 participantes en las narrativas con TC. La hipótesis 5 del estudio plantea que el factor TC desempeñará un papel más importante en las narrativas escritas de los HNs que de los HNNs. Esta hipótesis se cumple, pero no de la forma esperada. Se predijo que los HNs utilizarían un mayor número de expresiones definidas que los HNNs en las narrativas con TC, y que harían un mayor uso de cláusulas con información de fondo en narrativas sin TC. No se encontró evidencia para apoyar estas predicciones, en cambio, vemos que algunos HNs establecen el TC en la introducción de la narrativa al reconocer con su interlocutor ficticio que ambos han visto la película. De todas maneras, esta introducción es el único indicio de codificación explícita del TC con su interlocutor ficticio en toda la narrativa, puesto que el uso de expresiones definidas para el EC TYP, planteado por la hipótesis 4, solo ocurre en estas pre-introducciones. Vimos ejemplos de este comportamiento en la sección 6.7 y aquí abajo reproducimos uno de ellos, donde el hablante introduce a la familia con un DefNP:

(5) Ejemplo [P10.L1.TC]

L1. ¿No te acuerdas? Sí, hombre, está la familia (DefNP/TYP) en un banco de un parque,

Hay un caso de un HNN donde también se establece el TC en una pre-introducción, pero el escritor no utiliza expresiones referenciales de sujeto, por lo tanto, las expresiones utilizadas no son analizables en este trabajo:

(6) Ejemplo [P26.L2.TC]

L1. Acabo de mirar el videoclip de *la coqueta* y el travieso *Charlie* que me recomendaste. ¡Qué chistoso!

Este comportamiento de los HNs es comparable con los resultados de los monólogos del estudio de Smith, Noda, Andrews y Jucker (2005), que revisamos en el Capítulo 2. En los monólogos, el hablante no tiene oportunidad de confirmar el TC con el oyente, acercándose esta situación a las narraciones escritas del presente trabajo. La estrategia utilizada por los participantes en el estudio de Smith et ál. es la creación de una audiencia imaginaria o el asumir una audiencia informada, y esto lo hacen en la introducción de las narrativas. Las narrativas de Smith et ál. fueron conversaciones orales, por lo cual podemos concluir que este comportamiento de pre-introducción para el establecimiento del TC sigue el mismo patrón en ambas modalidades de producción del lenguaje.

La segunda tendencia es con respecto a la influencia del TC sobre el EC. Una vez que el escritor establece el TC, pasa a narrar la película y vuelve a introducir a los personajes. Esto lo vemos ejemplificado en (7):

(7) Ejemplo [P9.L1.TC]

L1. Te acuerdas de la película que vimos? Si, la del hombre que quer'ia ligar con un chica joven, en el

L2. parque.... Que no te acuerdas? quieres que la recuerde? Pues el resumen es que hay *un hombre aburrido*

Como consecuencia de esta pre-introducción en L1, el referente *el hombre* es activado en la memoria del lector, por lo cual su subsecuente introducción en la narrativa ya no está en el EC TYP sino en el EC FAM. En este caso, el TC no afecta la forma referencial puesto que el escritor utiliza una FN indefinida (*un hombre aburrido*), como es lo esperado al introducir un personaje nuevo en la narrativa, sino el EC del referente en la mente del lector.

El tercer comportamiento observado se refiere al hecho de que, aunque no haya evidencia de que el TC sea codificado expresamente por el escritor, es solamente a través de un conocimiento mutuo que el lector puede recuperar a la entidad intencionada por el escritor. Esto sucede tanto en las narrativas escritas de los HNs (8) como de los HNNs (9):

(8) Ejemplo [P8.L1.TC]

- L5. el hombre_i se va tras de ella
- L6. y [\emptyset]_i le ofrece una bebida,
- L7. mientras que el_i le trae la bebida
- L8. aparece otro hombre_j
- L9. (Chaplin)_j, también a conquistarla,
- L10. [\emptyset]_{ij} tienen una discusión

(9) Ejemplo [P26.L2.TC]

- L1. Acabo de mirar el videoclip de la coqueta y el travieso Charlie que me recomendaste. ¡Qué chistoso!
- L2. Cuando [\emptyset] comenzó con la señora durmiéndose al lado de su esposo, yo no esperaba lo que iba a pasar.
- L3. Este cortometraje es muy tradicional en términos de su comedia y los papeles masculinos y femeninos.
- L4. Cuando se durmieron las dos señoras
- L5. *el señor* (DefNP/FAM) se fue a seguir a la coqueta sin pensarlo.

En L10 del ejemplo (8), el sujeto nulo se refiere a los dos referentes (*Charlie* y *el hombre*) que han sido sujetos en las cláusulas anteriores, sin embargo, el personaje de la

mujer coqueta también se encuentra activado. La presencia de TC en la narrativa potencialmente elimina una posible ambigüedad causada por el uso del sujeto nulo, pero es el lector quien puede hacer uso del TC para identificar al referente intencionado por el escritor, y no el escritor quien lo ha codificado en su narrativa.

En el ejemplo (9) el participante establece el TC en la introducción. Se introduce al esposo en L2 y a Charlie en L1. Al mencionar al señor en L5 con una FN definida se crea un nuevo nodo que no está vinculado a la mención anterior del referente (*el esposo*) en L2. La falta de este vínculo cohesivo hace que el lector pueda utilizar el TC y el contexto para evitar seleccionar incorrectamente a Charlie como antecedente.

En conclusión, los resultados obtenidos nos permiten contestar la segunda pregunta de investigación de la siguiente manera: a través de los resultados cuantitativos no podemos valorar el impacto del TC en las narrativas escritas y, por lo tanto, las hipótesis 4 y 5 no se pueden confirmar; sin embargo, el análisis cualitativo nos permite observar que este factor hace parte integral de la GH y se codifica, principalmente, a través de pre-introducciones en las narrativas y no a través del uso de expresiones definidas.

7.4 Observaciones generales con respecto a la relación entre la adquisición del español como L2 y la cognición

Este trabajo contribuye hacia los estudios cognitivos del lenguaje, en cuanto que hemos podido observar que hay factores internos al lenguaje, así como factores universales de cognición, que guían la adquisición de una L2. El objetivo del lenguaje es la comunicación; las expresiones referenciales de sujeto surgen de la cognición como unidades simbólicas compuestas de forma y significado, por lo que tienen poder

comunicativo donde se evalúa la disponibilidad cognitiva del referente. Según Tomasello (2003:199), los hablantes harán lo que sea necesario para compartir el mismo marco de atención con el oyente, lo cual incluye el uso de construcciones nominales (así como de gestos deícticos). En el caso de la escritura, esto no deja de ser cierto: los datos de la presente investigación muestran que los escritores evalúan el conocimiento compartido con el lector, así como sus posibles niveles de memoria y atención en un momento dado, especialmente usando como base el discurso inmediatamente anterior. En otras palabras, tanto hablantes como escritores evalúan la disponibilidad cognitiva del referente en la mente del lector, lo cual involucra su accesibilidad, el nivel de topicalidad y de “lo dado”.

Las jerarquías cognitivas, como aquellas propuestas por Gundel et ál (1993), establecen niveles graduados de dicha disponibilidad cognitiva, desde la presuposición que hace un hablante/escritor de que algo se encuentra en el foco de atención inmediato del oyente/lector, hasta referentes que son completamente nuevos. Los HNNs de nivel avanzado tienen la misma habilidad cognitiva, al ser esta una habilidad universal, y utilizan herramientas lingüísticas, como lo son las expresiones referenciales, para codificar los diferentes niveles o grados de accesibilidad. El modo de producción escrito parece reducir significativamente las diferencias en el uso de estas herramientas entre HNs y HNNs, posiblemente debido a que en la escritura, al ser un modo más cuidado, el escritor tiene la oportunidad de revisar su lenguaje, incluso en narrativas de carácter espontáneo, como es el caso en el presente estudio.

Sin embargo, aun a este nivel avanzado de adquisición, hay algunas deficiencias residuales que permanecen, las cuales afectan la competencia discursiva de los HNNs. Este tipo de deficiencias han sido observadas en varios estudios donde se ha comparado

la adquisición del parámetro del sujeto nulo con las propiedades discursivas pragmáticas por parte de aprendientes de español como L2 (Pérez-Leroux y Glass, 1999, Licerias y Días, 1999, Lozano, 2009, *inter alia*). Este resultado lo podemos observar en el EC de tipo identificable, en donde se encontró una diferencia significativa en el uso de frases nominales definidas entre los dos grupos de participantes: los casos de FNs definidas utilizadas por los HNs representan el 10% de las ocurrencias, lo cual corresponde con cinco instancias, todas para introducir por primera vez a una entidad en la narrativa: *los tres caballeros, el policía, la policía y la mujer/los hombres*; mientras que los HNNs únicamente utilizaron FN indefinidas. Esto parecería ir en contra de la GH; sin embargo, este uso generalmente anafórico es pragmáticamente apropiado cuando los referentes tienen una función secundaria o terciaria en la narrativa, como sucede en todas las anteriores instancias. Es posible que el uso del artículo definido tenga como función pragmática la generalización de estas entidades como parte del fondo; es decir, del conocimiento compartido entre el escritor y el lector, en vez de considerarlos personajes como tal, Saunders (1999:102) también llegó a esta conclusión al encontrar en sus datos que el 66.7% de los hablantes nativos en su estudio seleccionaron FN definidas al introducir a la vendedora como personaje terciario en la narrativa: *la vendedora, la señorita de la caja*. A diferencia del estudio de Saunders, donde los HNNs avanzados también utilizaron FNs definidas en un 75% de los casos, aproximándose al uso nativo, en mis datos de las narrativas escritas esto no ocurre, demostrando algún error residual en la adquisición de la propiedad pragmática de topicalidad según la función de la entidad en la narrativa y su prominencia discursiva. Se necesitaría de más datos en estudios futuros para evaluar más a fondo estas diferencias.

Cuando una entidad se encuentra en un marco de atención compartido se considera tópico y, según la GH, este nivel de accesibilidad se codifica en español a través de expresiones de sujeto nulas o a través de pronombres, mientras que las entidades fuera de este marco de atención requieren de un mayor esfuerzo lingüístico para ser llevadas al foco de atención (Tomasello, 2003:267), y este mayor esfuerzo es codificado en las frases nominales. Los datos del presente estudio demuestran una frecuencia mayor en el uso de expresiones más elaboradas lingüísticamente (frases nominales definidas y nombres propios) que en el estudio de datos orales de ByLQ. Hemos podido observar esta tendencia tanto en los HNs como en los HNNs, y es posible que se deba a que en la modalidad escrita el costo cognitivo de accesibilidad de un referente sea mayor que en la modalidad oral, o también debido a la imposibilidad de confirmar el marco de atención compartido entre escritor y lector.

Las formas lingüísticas surgen de la cognición y en consecuencia indican algo acerca de la cognición. La presente investigación ha mostrado cómo la escritura puede afectar cognitivamente la selección de las formas referenciales para identificar a entidades en una narrativa. Asimismo, hemos podido comparar el uso y selección de las formas referenciales de sujeto, según el nivel de atención y memoria, entre HNs de español y HNNs avanzados de español como L2. La presente investigación ha presentado datos que demuestran que a un alto nivel de adquisición, los recursos lingüísticos usados por los HNNs tienden a acercarse al habla nativa en la modalidad escrita, aun cuando todavía podemos observar algunas diferencias al nivel pragmático-discursivo.

7.5 Limitaciones del estudio

Las principales limitaciones del estudio son de naturaleza metodológica. En primer lugar, para poder recoger narrativas escritas que pudieran compararse con las narrativas orales previamente estudiadas, fue necesario utilizar el recuento de una película; sin embargo, el narrar una película por escrito es, sin duda, una tarea artificial. Para intentar ofrecer a los participantes un elemento más natural, se les pidió que le narraran la película a un amigo y que describieran al final lo que más les había gustado de ella. En algunos casos esta estrategia resultó efectiva y evidencia de ello lo vemos sobre todo en las narrativas con TC, cuando el escritor intenta establecer el hecho de que ambos han visto la película en las pre-introducciones, pero en la mayoría de los casos los participantes simplemente narraron la película sin hacer referencia a un amigo ficticio y sin expresar lo que más les había gustado. De todas maneras, siempre habrá en las investigaciones lingüísticas de esta naturaleza algún factor artificial difícil de eliminar, lo cual no invalida las investigaciones.

Para poder recoger un número suficiente de narrativas, se decidió utilizar una herramienta en línea; a su vez, esta herramienta permitió diseñar una tarea de escritura no-planeada al incluir un límite de tiempo de 15 minutos para la producción espontánea. Además, el hecho de que la investigadora no estuviera presente tal vez permitió que los participantes estuvieran más relajados. La desventaja principal es con respecto al TC. Al no haber visto la película al mismo tiempo que los participantes y, aunque se les pidió que pensarán en un amigo ficticio que también había visto la película, la falta de copresencia física con un interlocutor real pudo haber afectado la manipulación del TC y, en consecuencia, la posibilidad de aislar esta variable de forma exitosa.

Otra limitación es con respecto a los participantes. El haber utilizado diferentes participantes en el estudio oral de ByLQ y en el estudio escrito para el presente trabajo, no tenemos una exacta equivalencia entre los resultados de las narrativas orales y de las narrativas escritas, principalmente en cuanto a la competencia lingüística de los HNNs. Tampoco podemos estar seguros de que el conocimiento del personaje de Chaplin sea el mismo entre todos los hablantes, lo cual pudo haber afectado el uso del nombre propio en las narrativas, aunque este sea mundialmente conocido. Sin embargo, este efecto se ve neutralizado por el uso de FN definidas cuando se desconoce el nombre del personaje. Las formas DefNP y PrpN se consideran expresiones más elaboradas, por lo tanto los resultados para efectos de la presente investigación no se ven afectados negativamente.

7.6 Investigaciones futuras

Aunque la adherencia a la GH se ve demostrada en la presente investigación en relación con las narrativas escritas, también observamos que dicha adherencia varía dependiendo de la modalidad: la GH tiene menos poderes predictivos en las narrativas escritas que en las orales. En este trabajo solo analizamos narrativas espontáneas; futuros trabajos podrían utilizar datos de narrativas planeadas tanto de HNs como de HNNs y de esta manera determinar cómo afecta el nivel de planeación las selecciones de los hablantes. Es posible que veamos un mayor grado de desviación de la GH, aunque es probable que la tendencia siga siendo la de adherencia puesto que la GH es universal. Además, sería de interés observar si a mayor nivel de planeación, también se aprecia un mayor acercamiento al habla nativa por parte de los HNNs, reduciendo así el déficit residual en la adquisición. De esta manera, trabajos futuros podrían considerar otros

géneros de producción (oral y escrito): conversaciones o diálogos, narraciones personales, artículos de periódico y revistas, entre otros. Estos estudios podrían incluir una comparación de los diferentes géneros.

En este trabajo se analizaron únicamente expresiones referenciales de sujeto, pero hemos visto que el introducir a un personaje en otras posiciones sintácticas, el EC del referente se ve afectado. Por ejemplo, al mencionar a un personaje en posición de objeto, este es llevado al foco de atención. Trabajos futuros podrían analizar el EC de las expresiones referenciales en otras posiciones sintácticas.

Otra área de investigación es con respecto a la cohesión narrativa. Hemos visto que, en ocasiones, el uso de sujetos más elaborados permite una mayor cohesión al disminuir el costo cognitivo de procesamiento. Habría que elaborar este punto, incluyendo los datos de las narrativas orales para determinar si hay diferencias significativas en la manera en que se construye la cohesión en un texto según la modalidad.

Asimismo, vimos cómo la perseverancia parece contribuir hacia una menor carga cognitiva a través de la repetición de ciertas formas estructurales, lo cual sucedió primordialmente con las formas más elaboradas. Previos estudios en este campo se han centrado en la perseverancia de sujetos nulos y pronombres en el discurso oral, principalmente entre dos interlocutores. Expandir el análisis a otras formas de sujeto en el discurso escrito ofrece la oportunidad de informar los efectos de este fenómeno psicolingüístico.

Por último, a través de una metodología diferente que permita aislar el TC con mayor claridad, un estudio futuro podría retomar este factor y determinar de manera más concluyente el efecto que este tiene en la codificación lingüística de los ECs.

BIBLIOGRAFÍA

- Abreu, L. (2012). Subject Pronoun Expression and Priming Effects Among Bilingual Speakers of Puerto Rican Spanish. En K. Geeslin y M. Díaz-Campos (eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, 1-8. Somerville, MA: Cascadilla Proceedings Project.
- Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L. y Clifton, C. Jr. (2002). Null vs overt pronouns and Topic-Focus articulation in Spanish. *Italian Journal of Linguistics Rivista di linguistica*, 14(2), 151-169.
- Ariel, M. (1988). Referring and Accessibility. *Journal of Linguistics*, 24(1), 65-87.
- Ariel, M. (1990). Accessing Noun-phrase Antecedents. *Linguistics*, 24, 65-87. New York: Routledge.
- Bel, A., Perera, J y Salas, N. (2010). Anaphoric devices in written and spoken narrative discourse data from Catalan. *Written Language & Literacy*, 13 (2), 236-259.
- Blackwell, S. E. (2001). Testing the new-Gricean pragmatic theory of anaphora: The influence of consistency constraints on interpretations of coreference in Spanish. *Journal of Pragmatics*, 33, 901-941.
- Blackwell, S. E. (2003). *Implicatures in Discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Blackwell, S. E., & Lubbers Quesada, M. (2012). Third-Person Subjects in Native Speakers' and L2 Learners' Narratives: Testing (and Revising) the Givenness

- Hierarchy for Spanish. En K. Geeslin y M. Díaz-Campos (eds.), *14th Hispanic Linguistics Symposium*, 142-164. Somerville: Cascadilla.
- Bourdin, B. y Fayol, M. (1994). Is Written Language Production More Difficult Than Oral Language Production? A Working Memory Approach. *International Journal of Psychology*, 29 (5), 591-620.
- Bourdin, B. y Fayol, M. (2002). Even in adults, written production is still more costly than oral production. *International Journal of Psychology*, 37 (4), 219-227.
- Cadierno, T. e Hijazo-Gascón. (2013). A. Cognitive linguistic approaches to second language Spanish. En K.L. Geeslin (ed.), *The Handbook of Spanish Second Language Acquisition* (pp.96-110). Chichester, Inglaterra: John Wiley & Sons, Inc.
- Cadierno, T. y Pedersen, J. (2014). Introducción a la lingüística cognitiva: Orígenes y postulados básicos. En S.S. Fernández y J. Falks (eds.), *Temas de gramática española para estudiantes universitarios: una aproximación cognitiva y funcional* (pp.15-35). Bern, Suiza: Peter Lang D.
- Carlson, T.B. (1992). *Context for Comprehension*. En H.H. Clark (Ed.), *Arenas of Language Use*, 60-77. Chicago: The University of Chicago Press.
- Carminati. (2002). *The Processing of Italian Subject Pronouns*. (Disertación Doctoral). University of Massachusetts.
- Chafe, W. L. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. En C. Li (Ed.) *Subject and Topic*, 25-55. New York Academic Press.
- Chafe, W. L. (1994). Speaking and Writing. En *Discourse, Consciousness and Time*, 41-50. Chicago/Londres: The University of Chicago Press.

- Chafe, W. y Tannen, D. (1987) The Relation between Written and Spoken Language. *Annual Review of Anthropology*, 16, 383-407.
- Clark, H.H. (1975). Bridging. En P. N. Johnson-Laird y P. C. Wason (Eds.), *Thinking: Readings in cognitive science*, 411–420. Cambridge: Cambridge University Press.
- Clark, H.H. y Carlson, T. (1981). Context for Comprehension. En J. Long y A. Baddeley (eds.), *Attention and Performance IX*, 313-330. New Jersey, Hillsday: Lawrence Erlbaum Associates.
- Clark, H. H. y Marshall, C. R. (1981). Definite reference and mutual knowledge. En *Elements of discourse understanding*, 10-63. Nueva York: Press Syndicate of the University of Cambridge.
- de Beaugrande, R. (2004). Language, discourse, and cognition: Retrospects and prospects. En T. Virtanen (ed.), *Approaches to Cognition through Text and Discourse*, 1-16. Berlín: Mouton de Gruyter.
- Dussias, P. (2003). Cognitive Perspectives on the Acquisition of Spanish as a Second Language. En B. Lafford y R. Salaberry (eds.), *Spanish Second Language Acquisition* (pp. 432-456). Washington, D.C.: Georgetown University Press.
- Ellis, R. (1985). A Variable Competence Model of Second Language Acquisition. En *IRAL*, 23(1-4),47-70.
- Figueras, C. (2002). La jerarquía de la accesibilidad de las expresiones referenciales en español. *Revista Española de Lingüística*, 32(1), 53-96.
- Filiaci, F. (2010). Null and overt biases in Spanish and Italian: A cross-linguistic comparison. En C. Borgonovo et ál (eds.), *Selected Proceedings of the 12th Hispanic*

- Linguistics Symposium*, (pp.171-182). Somerville, MA: Cascadilla Proceedings Project.
- Fillmore, C.J. (1975). An Alternative to Checklist Theories of Meaning. *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, 123-131.
- García Salido, M. (2011). La distinción de ísis/anáfora y su aplicación a las formas de persona del español. *Revista de filología del español*, 65-88.
- García-Alcaraz, E. y Bel, A. (2011). Selección y distribución de los pronombres en el español L2 de los hablantes de árabe. *Revista de Lingüística y Lenguas Aplicadas*, 6, 165-179.
- Geeslin, K., Linford, B., Fafulas, S., Long, A. y Díaz-Campos, M. (2013). The L2 Development of Subject Form Variation in Spanish: The individual vs. the group. En J. C. Amaro et ál (eds.), *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, 156-174. Somerville, MA: Cascadilla Proceedings Project.
- Givón, T. (1983). Introduction. En G. Talmy (ed.), *Topic Continuity in Discourse: A Quantitative Cross-language Study*, 5-41. Amsterdam/Philadelphia: John Benjamins Franklin.
- Gudmestad, A. (2014). *Variationist approaches to second language Spanish*. En K.L. Geeslin (ed.), *The Handbook of Spanish Second Language Acquisition* (pp.80-93). Chichester, Inglaterra: John Wiley & Sons, Inc.
- Gudmestad, A. y Geeslin, K.L. (2010). Exploring the Roles of Redundancy and Ambiguity in Variable Subject Expression: A Comparison of Native and Non-native Speakers. En C. Borgonovo et ál (eds.), *Selected Proceedings of the 12th Hispanic*

- Linguistics Symposium*, (pp.270-283). Somerville, MA: Cascadilla Proceedings Project.
- Gudmestad, A., House, L. y Geeslin, K.L. (2013). What a Bayesian Analysis Can Do for SLA: New Tools for the Sociolinguistic Study of Subject Expression in L2 Spanish. *Language Learning*, (63)3, 371–399.
- Gundel, J. K. (1985). Shared knowledge and Topicality. *Journal of Pragmatics*, 9, 83-107.
- Gundel, J.K. (2010). Reference and Accessibility from a Givenness Hierarchy Perspective. *International Review of Pragmatics*, 2, 148-168.
- Gundel, J.K., Bassene, M., Gordon, B., Humnick, L., y Khalfaoui, A. (2010). Testing predictions of the Givenness Hierarchy framework: A crosslinguistic investigation. *Journal of Pragmatics*, (42), 1770–1785.
- Gundel, J. K., Hedberg, N. y Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, 69(2), 274-307.
- Halliday, M.A.K. y Hasan, R. (1976). Chapter 1- Introduction. *Cohesion in English*, 1-30. Londres/Nueva York: Routledge.
- Huang, Y. (1991). A neo-Gricean pragmatic theory of anaphora. *Journal of Linguistics* 27, 301-335.
- Isabelli, C.A. (2003). The Acquisition of the Null Subject Parameter Properties in SLA: Some Effects of Positive Evidence in a Naturalistic Learning Context. *Hispania*, 87(1), 150-162.

- Jegerski, J., VanPatten, B. y Keating, G. (2011). Cross-linguistic Variation and the Acquisition of Pronominal Reference in L2 Spanish. *Second Language Research*, 27(4),481-507.
- Jorgensen, S. W. (1998) What are referring expressions. En M. Dascal (ed.), *Pragmatics & Cognition*, 99-120. John Benjamins Publishing Company: Amsterdam/Philadelphia.
- Kellogg, R.T. (2001). Competition for Working Memory among Writing Processes. *The American Journal of Psychology*, 114 (2), 175-191.
- Levinson, S. C. (2000). The Pragmatics of local anaphora. En *Presumptive Meanings: The Theory of Generalized Conversational Implicature*, 267-273. Cambridge, MA: MIT Press.
- Liceras, J.M. y Díaz, L. (1999). Topic-drop versus pro-drop: null subjects and pronominal subjects in the Spanish L2 of Chinese, English, French, German and Japanese speakers. *Second Language Research*, 15(1),1-40.
- Lieven, E. y Tomasello, M. (2008). Children´s First Language Acquisition from a Usage-Based perspective. En P. Robinson, y N. Ellis (eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp.168-196). Nueva York: Routledge.
- Lozano, C. (2008). ¿Déficits de representación o de procesamiento en una segunda lengua? Evidencia de un estudio de resolución de anáfora con griegos adultos aprendices de español. En R. Monroy y A. Sánchez (eds.), *25 Años de Lingüística Aplicada en España: Hitos y retos*, (pp.855-866). Murcia: Editum.

- Lozano, C. (2009). Selective deficits at the syntax-discourse interface. Evidence from the CEDEL2 Corpus. En N. Snape, Y.I. Leung y M. Sharwood Smith (eds.), *Representational Deficits in SLA: Studies in Honor of Roger Hawkins*, 127-166. Amsterdam/Philadelphia: John Benjamins Franklins.
- Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntax-discourse interface: Advanced English learners of Spanish in the CEDEL2 corpus. En M. Alonso-Ramos (ed.), *Spanish Learner Corpus Research: Current trends and future perspectives* (pp.235-265). Amsterdam/Philadelphia: John Benjamins.
- Lubbers Quesada, M. (2015). *The L2 acquisition of Spanish subjects*. Boston/Berlin: Walter de Gruyter Inc.
- Lubbers Quesada, M. y Blackwell, S. E. (2009). The L2 Acquisition of Null and Overt Spanish Subject Pronouns: A Pragmatic Approach. En J. Collentine et ál. (eds.), *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, (pp.117-130). Somerville, MA: Cascadilla Proceedings Project.
- Martín Rojo, L. y Meeuwis, M. (1993). Referentes del sujeto pronominales y tácitos en la conversación en español: un enfoque pragmático. En H. Haverkate, K. Hengeveld y G. Mulder (eds.) *Diálogos Hispánicos 12: Aproximaciones pragmlingüísticas al español*, 87-118. Amsterdam: Rodopi.
- Matos Amaral, P. y Schwenter, S.A (2005). Contrast and the (Non-) Occurrence of Subject Pronouns. *Selected Proceedings of the 7th Hispanic Linguistics Symposium*, D. Eddington (ed.), 116-127. Somerville, MA: Cascadilla Proceedings Project.
- McCutchen, D. (1996). A Capacity Theory of Writing: Working Memory in Composition. *Educational Psychology Review*, 8 (3), 299-325.

- Montalbetti, M.M. (1984). *After Binding. On the Interpretation of Pronouns*. (Disertación Doctoral). Massachusetts Institute of Technology.
- Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntax convergence. *Bilingualism: Language and Cognition*, 7, 125-142.
- Montrul, S. y Rodríguez Louro, C. (2006). Beyond the syntax of the null subject parameter: A look at the discourse-pragmatic distribution of null and overt subjects by L2 learners of Spanish. En V. Torrens and L. Escobar (eds.), *The acquisition of syntax in Romance languages*, 401-418. Amsterdam/Filadelfia: John Benjamins.
- Nuyts, J. (2010). Cognitive Linguistics and Functional Linguistics. En D. Geeraerts y H. Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 543-565. Oxford: Oxford University Press.
- Ochs, E. (1979). Planned and Unplanned Discourse. En T. Givón (ed.), *Syntax and Semantics: Volume 12*, 51-80. Nueva York/San Francisco/Londres: Academic Press.
- Olson, D.R. (1996). Towards a psychology of literacy: On the relations between speech and writing. *Cognition* 60, 83-104.
- Olson, D.R. (2016). *The Mind on Paper*. Cambridge: Cambridge University Press.
- Otheguy, R. (2014). Remarks of Pronominal Perseveration and Functional Explanation. En R. Torres Cacoullos, N. Dion y A. Lapierre. (eds.). *Linguistic Variation: Confronting Fact and Theory*, 373-396.
- Perales, S. y Portillo, R. (2007). Sobre las propiedades referenciales de los sujetos nulos y pronominales del español oral y escrito. En E. Balsameda Maestu (ed.), *Las destrezas orales en la enseñanza del español L2-LE: XVII Congreso Internacional*

- de la Asociación del Español como Lengua Extranjera (ASELE)*, (2), 889-900.
Logroño, España: Universidad de la Rioja.
- Pérez-Leroux, A.T, y Glass, W.R. (1999). Null Anaphora in Spanish Second Language Acquisition: Probabilistic versus Generative Approaches. *Second Language Research*, 15 (2), 220-249.
- Prince, E. F. (1981). Toward a Taxonomy of Given-New Information. En P. Cole (ed.), *Radical pragmatics*, 223-233. New York, NY: Academic Press.
- Reinhart, T. (1976). *The Syntactic Domain of Anaphora*. (Disertación doctoral).
Massachusetts Institute of Technology.
- Reinhart, T. (1986). Center and periphery in the grammar of anaphora. En B. Lust (ed.), *Studies in the Acquisition of Anaphora (Vol. 1): Defining the Constraints*, 123-150.
Dordrecht: Reidel.
- Ritchie, W. C. y Bhatia, T. K. (1996). Second language acquisition: Introduction, foundations, and overview. En W. C. Ritchie y T. K. Bhatia (eds.), *Handbook of Second Language Acquisition*, (pp.1-35). San Diego: Academic Press.
- Rothman, J. e Iverson, M. (2007). On Parameter Clustering and Resetting the Null-Subject Parameter in L2 Spanish: Implications and Observations. *Hispania*, 90 (2), 328-341.
- Rothman, J. (2009). Pragmatic Deficits with Syntactic Consequences? L2 Pronominal Subjects and the Syntax-Pragmatics Interface. *Journal of Pragmatics*. 41, 951-973.
- Rothman, J. y Pascual y Cabo, D. (2014). Generative approaches to Spanish second language acquisition. En K.L. Geeslin (ed.), *The Handbook of Spanish Second Language Acquisition*, 46-63. Chichester, Inglaterra: John Wiley & Sons, Inc.

- Sánchez, L. y Toribio, A. J. (2003). Current Issues in the Generative Study of Spanish Second Language Syntax. En B. Lafford y R. Salaberry (eds.), *Spanish Second Language Acquisition*, 189-232. Washington, D.C.: Georgetown University Press.
- Saunders, K. J. (1999). *Null and Overt References in Spanish Second Language Acquisition: A Discourse Perspective*. (Disertación Doctoral). The University of Texas.
- Schreuder, R. y Buttrick, S. (1992). Common Ground and the Understanding of Demonstrative Reference”. En H.H. Clark (ed.), *Arenas of Language Use*, 78-99. Chicago: The University of Chicago Press.
- Smith, S.W. y Jucker, A.H. (1998). Interactive aspects of reference assignment in conversation. En M. Dascal (ed.), *Pragmatics & Cognition*, 153-187. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Smith, S.W. y Jucker, A.H. y Müller, S. (2000). ‘Some artist guy’: the role of salience and common ground in the formulation of referring expressions in conversational narratives. En T.E. Németh (ed.), *Pragmatics in 2000: Selected Papers from the 7th International Pragmatics Conference*, 2, 528–542. Antwerp: International Pragmatics Association.
- Smith, S.W., Noda, H.P., Andrews, S. y Jucker, A.H. (2005). Setting the Stage: How Speakers Prepare Listeners for the Introduction of Referents in Dialogues and Monologues. *Journal of Pragmatics*, 37, 1865-1895.
- Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P. y Stevenson, M. (2003). First Language and Second Language Writing: The Role of

- Linguistic Knowledge, Speed of Processing, and Metacognitive Knowledge.
Language Learning, 53 (1), 165-202.
- Sorace, A. y Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian.
Second Language Research, 22(3), 339–368.
- Tannen, D. (1980). Spoken/Written Language and the Oral/Literate Continuum.
Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society, 207-218
- Tannen, D. (1982)a. The Oral/Literate Continuum in Discourse. En D. Tannen (ed.),
Spoken and Written Language: Exploring Orality and Literacy, 1-16. Norwood:
 Ablex Publishing Corporation.
- Tannen, D. (1982)b. Oral and Literate Strategies in Spoken and Written Narratives.
Language, 58 (1), 1-21.
- Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics*, 4,
 142–163.
- Tomasello, M. (1998). Introduction: A Cognitive-Functional Perspective on Language
 Structure. En M. Tomasello (ed.), *The New Psychology of Language: Cognitive and
 Functional Approaches to Language Structure* (pp.7-23). New Jersey/London:
 Lawrence Erlbaum Associates.
- Tomasello, M. (2005). *Constructing a Language*. Cambridge/Londres: Harvard
 University Press.
- Tomasello, M. (2006). First steps towards a usage-based theory of language acquisition.
 En Geeraerts, D. (ed.), *Cognitive Linguistics: Basic Readings* (pp.439-457).
 Berlin/New York: Mouton de Gruyter.

- Torres Cacoullous, R. y Travis, C. E. (2014). Prosody, priming and particular constructions: The patterning of English first-person singular subject expression in conversation. *Journal of Pragmatics*, 63:19-34.
- Virtanen, T. (2004). Text, discourse and cognition: An introduction. En T. Virtanen (ed.) *Approaches to Cognition through Text and Discourse*, 1-16. Berlín: Mouton de Gruyter.
- Zubizarreta, M.L. (1998). *Prosody, focus, and word order*. Cambridge, MA: MIT press.
- Zulaica-Hernández, I. (2009). On the Cognitive Status of Antecedents in Spanish Discourse Anaphora. *Proceedings of the First International Conference on Corpus Linguistics (CILC I)*, 646-663.
- Zyzik, E. (2014). *Functional approaches to second language Spanish*. En K.L. Geeslin (ed.), *The Handbook of Spanish Second Language Acquisition* (pp.30-45). Chichester, Inglaterra: John Wiley & Sons, Inc.
- Zyzik, E. (2017). Subject expression in L2 Spanish: Convergence of generative and usage-based perspectives? *Second Language Research* 33 (1), 33-59.

Apéndice A: Prueba adaptada DELE del Instituto Cervantes²⁶

Las siguientes fueron las preguntas incluidas en la encuesta para los hablantes no-nativos. En negrilla se presenta la respuesta correcta. También se dio la opción de contestar *I don't know the answer* ('No sé la respuesta').

1. ____ un chico muy abierto. Le encanta hablar con todo el mundo: **es** / está
2. La mujer ____ vimos ayer, es la mujer de Juan: quien / **que**
3. Sentí que alguien ____ a la puerta: llamara / **llamaba**
4. Por mucho que ____, no conseguirás aprobar: **estudies** / estudias
5. Salgo para Madrid mañana ____ mediodía: **al** / en el
6. Se me ____ las llaves: he perdido / ha perdido / **han perdido**
7. Espero que ____ más suerte la próxima vez: **tengas** / hayas tenido / tendrás
8. Ojalá ____ al concierto, pero estuve enfermo: fuera / **hubiera ido** / vaya
9. Me di cuenta de que había olvidado las llaves ____ llegué a la oficina: aun cuando / antes de / **tan pronto como**
10. Date prisa que los invitados ____ llegar: están a punto de / se ponen a / vienen a
11. No me gusta este apartamento: es pequeño, frío y, ____, caro: **además** / entonces / igualmente
12. Por mucho que ____, no lo terminarás: te hubieras esforzado / **te esfuerces** / te esfuerzas
13. Las flores ____ me felicitó el cumpleaños Juan son muy bonitas: de las que / **con las que** / a las cuales
14. Cuando llegues, ya ____ la cena: prepararé / **habré preparado** / habría preparado
15. No tenía hambre, de manera que ____ sin cenar a dormir: se vaya / se fuera / **se fue**
16. Hola, Carlota, ¿Qué haces tú ____ aquí?: **por** / hacia / para / en
17. Me amenazó ____ no dejarme nunca más los apuntes de clase: a / de / **con** / por
18. ____ al retraso de las ayudas económicas, mereció la pena el proyecto: a pesar de / **pese** / si bien / aun
19. Al final no fuimos a Madrid. Nos ____ pero no pudimos ir porque Javier se puso enfermo: había gustado / gustó / gustaría / **habría gustado**
20. De repente ____ una idea para solucionar el problema de las tuberías: ocurrió / se ocurrió / me ocurrió / **se me ocurrió**

²⁶ Diploma de español: nivel superior. Instituto Cervantes, 11 de mayo de 2007.

Apéndice B: Expresiones generales – estado cognitivo ACTNR

En la siguiente tabla, vemos todas las expresiones generales utilizadas en el estado ACTNR de las narrativas escritas. La tercera columna muestra la reclasificación a DefNP o IndNP:

Tabla Apéndice B - Expresiones generales utilizadas por HNs y HNNs: estado cognitivo ACTNR

# Participante HN	GenNP	Reclasificación DefNP/IndNP	# Participante HNN	GenNP	Reclasificación DefNP/IndNP
1	uno	IndNP	21	el otro	DefNP
7	una de las escenas	IndNP	33	los dos hombres	DefNP
8	una	IndNP	36	una	IndNP
12	uno de los hombres	IndNP	38	el otro	DefNP
14	una de las mujeres	IndNP	39	una mujer	IndNP
14	los dos	DefNP	39	la otra mujer	DefNP
14	uno de ellos	IndNP	39	el otro hombre	DefNP
14	otro	IndNP	24	uno de ellos	IndNP
14	el otro hombre	DefNP			
16	los 2	DefNP			
16	los 2 enamorados	DefNP			
19	una de ellas	IndNP			
19	la otra	DefNP			
19	el otro hombre	DefNP			