



UNIVERSIDAD DE QUINTANA ROO

DIVISIÓN DE CIENCIAS POLÍTICAS Y HUMANIDADES

---

El estudio de la complejidad sintáctica en  
España y en Hispanoamérica

---

**Monografía en la modalidad de investigación documental**

Para obtener el grado de  
**Licenciada en Humanidades**

PRESENTA  
**Suemy Nayeli Cal Santamaria**

ASESORES  
Mtra. María Elena Sánchez Arroba  
Dra. Edith Hernández Méndez  
Dr. Moisés Damián Perales Escudero  
Dra. Amparo Reyes Velázquez  
Dr. Martín Ramos Díaz



Chetumal, Quintana Roo, México, noviembre de 2019



UNIVERSIDAD DE QUINTANA ROO

DIVISIÓN DE CIENCIAS POLÍTICAS Y HUMANIDADES

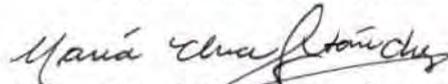
El estudio de la complejidad sintáctica en España y en Hispanoamérica

Presenta:  
Suemy Nayeli Cal Santamaria

Monografía para obtener el grado de Licenciada en Humanidades

COMITÉ DE SUPERVISIÓN

Asesor:

  
Mtra. María Elena Sánchez Arroba

Asesor:

  
Dra. Edith Hernández Méndez

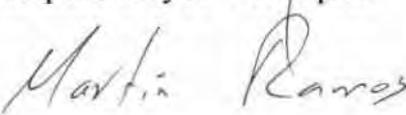
Asesor:

  
Dr. Moisés Damían Perales Escudero

Suplente:

  
Dra. Amparo Reyes Velázquez

Suplente:

  
Dr. Martín Ramos Díaz

Chetumal, Quintana Roo, México, noviembre de 2019



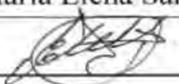
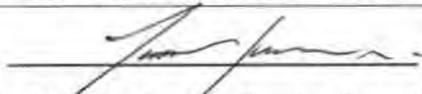
## Declaratoria de originalidad

**Suemy Nayeli Cal Santamaria**

En mi calidad de autora de la monografía titulada “El estudio de la complejidad sintáctica en España y en Hispanoamérica”, que presento para obtener el título de Licenciada en Humanidades, declaro bajo protesta de decir verdad que:

1. Este trabajo monográfico es inédito y original, de mi propia autoría intelectual.
2. Que si bien contiene parcialidades del contenido de obras, las mismas son citadas y respaldas en el reconocimiento del derecho moral de los autores; por lo que no es una traducción ni una versión mejorada de otro documento publicado o aún sin publicar.
3. No ha sido utilizada anteriormente para obtener algún grado académico, ni ha sido publicado por cualquier medio.
4. En todas las citas y las paráfrasis que utilizo, identifico las fuentes originales e incluyo las referencias completas en el apartado correspondiente.
5. Identifico la procedencia de las tablas y figuras (gráficas, mapas, diagramas, esquemas ilustraciones, arte digital, fotografías u otros) previamente publicadas.
6. Todos los contenidos de esta monografía están libres de derechos de autor y asumo la responsabilidad de cualquier litigio o reclamación relacionada con derechos de propiedad intelectual, exonerando de toda responsabilidad a la Universidad de Quintana Roo.
7. Reconozco que la Universidad de Quintana Roo no comparte necesariamente las afirmaciones que en esta monografía se plantean.

Chetumal, Quintana Roo, a 20 de noviembre de 2019.

 <hr/> Br. Suemy Nayeli Cal Santamaria	 <hr/> Mtra. María Elena Sánchez Arroba
	 <hr/> Dra. Edith Hernández Méndez
	 <hr/> Dr. Moisés Damán Perales Escudero

## **Carta de cesión de derechos**

**Suemy Nayeli Cal Santamaria**

En mi calidad de autora de la monografía titulada “El estudio de la complejidad sintáctica en España y en Hispanoamérica”, presentada para obtener el título de Licenciada en Humanidades, es de mi plena voluntad:

1. Autorizar a la Universidad de Quintana Roo para que se encargue de la divulgación total o parcial de este documento, en formato impreso o digital, sin limitación en el tiempo, por los medios que dicha institución decida, y con fines académicos exclusivamente.
2. Aceptar que los lectores puedan descargar, almacenar, copiar y distribuir gratuitamente la versión final aprobada de la tesis, siempre y cuando se realice sin fines comerciales, no se generen obras derivadas y se mencione la autoría de la obra.
3. Reutilizar la versión final aprobada de la monografía con propósitos académicos y a publicar la obra, en formato impreso o digital, siempre y cuando no se generen nuevos derechos que impidan a la Universidad de Quintana Roo continuar con la divulgación de la obra.
4. Aceptar que si la monografía es publicada con fines comerciales, esta no debe denotar, contener, insertar o incluir en ninguna parte interna o externa de la publicación el escudo, emblema, logotipo o nombre de la Universidad de Quintana Roo. En caso contrario, debo obtener previamente la autorización por escrito del representante legal de la Universidad.
5. Autorizo en este acto a la Universidad de Quintana Roo para que difunda mi información personal, tales como mi nombre y seudónimo, correo electrónico a través del repositorio del sistema bibliotecario ([www.risisbi.uqroo.mx](http://www.risisbi.uqroo.mx)), por lo que, la eximo de cualquier responsabilidad y/o futura reclamación presente por la protección de datos personales señalados en la Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados.

Chetumal, Quintana Roo, a 20 de noviembre de 2019.



Br. Suemy Nayeli Cal Santamaria

## AGRADECIMIENTOS

Quiero expresar mi gratitud, primeramente, a Dios por permitirme cumplir una de mis metas profesionales, y por supuesto, a aquellas personas que de alguna manera me brindaron su apoyo durante este difícil, pero no imposible, proceso para mí:

A la Mtra. María Elena Sánchez Arroba, asesora principal de esta monografía, mi infinita gratitud por sus enseñanzas, su tiempo y su paciencia a lo largo de la redacción de este trabajo; pues con su experiencia y sus conocimientos, me orientó e hizo posible la conclusión de este trabajo.

A cada uno de los integrantes de mi Comité de Supervisión, les agradezco su colaboración; sus comentarios y sus observaciones fueron valiosos para enriquecer la versión preliminar de esta monografía. Dra. Edith Hernández Méndez, Dr. Moisés Damián Perales Escudero, Dra. Amparo Reyes Velázquez y Dr. Martín Ramos Díaz, muchas gracias.

A mis padres, les estaré eternamente agradecida por su apoyo moral, económico e incondicional durante todo este tiempo. Igualmente, a mis hermanas y a mi compañero de aventuras, quienes siempre creyeron en mí y me motivaron a seguir adelante.

Por último, agradezco a todos los maestros de la DCPH, por los conocimientos compartidos, por su guía y por su dedicación durante el trayecto de mi formación académica en la Universidad de Quintana Roo.

# ÍNDICE DE CONTENIDO

<b>ÍNDICE DE CONTENIDO</b> .....	VI
<b>ÍNDICE DE FIGURAS</b> .....	VII
<b>ÍNDICE DE TABLAS</b> .....	VIII
<b>INTRODUCCIÓN</b> .....	IX
<b>CAPÍTULO 1. ESTUDIOS SOBRE COMPLEJIDAD SINTÁCTICA EN ESPAÑA</b> .....	12
1.1 Estudios sobre complejidad sintáctica realizados en Granada .....	12
1.2 Estudios sobre complejidad sintáctica realizados en las Islas Canarias .....	19
1.3 Estudios sobre complejidad sintáctica realizados en Madrid .....	25
<b>CAPÍTULO 2. ESTUDIOS SOBRE COMPLEJIDAD SINTÁCTICA EN HISPANOAMÉRICA</b> .....	30
2.1 Estudios sobre complejidad sintáctica realizados en Centroamérica y en el Caribe .....	30
2.2 Estudios sobre complejidad sintáctica realizados en Chile .....	39
2.3 Estudios sobre complejidad sintáctica realizados en Argentina .....	56
2.4 Estudios sobre complejidad sintáctica realizados en otros países de Sudamérica .....	61
<b>CAPÍTULO 3. REFLEXIONES CRÍTICAS SOBRE LOS ESTUDIOS DE COMPLEJIDAD SINTÁCTICA REALIZADOS FUERA DE MÉXICO</b> .....	69
<b>CONCLUSIONES</b> .....	74
<b>REFERENCIAS</b> .....	76

## ÍNDICE DE FIGURAS

Figura 1. ....	26
Figura 2. ....	65

## ÍNDICE DE TABLAS

Tabla 1.....	12
Tabla 2.....	13
Tabla 3.....	14
Tabla 4.....	15
Tabla 5.....	18
Tabla 6.....	18
Tabla 7.....	20
Tabla 8.....	21
Tabla 9.....	22
Tabla 10.....	22
Tabla 11.....	22
Tabla 12.....	23
Tabla 13.....	24
Tabla 14.....	27
Tabla 15.....	30
Tabla 16.....	31
Tabla 17.....	33
Tabla 18.....	36
Tabla 19.....	37
Tabla 20.....	43
Tabla 21.....	45
Tabla 22.....	46
Tabla 23.....	47
Tabla 24.....	49
Tabla 25.....	51
Tabla 26.....	53
Tabla 27.....	54
Tabla 28.....	57
Tabla 29.....	58
Tabla 30.....	60
Tabla 31.....	63
Tabla 32.....	67
Tabla 33.....	67
Tabla 34.....	68

## INTRODUCCIÓN

El presente trabajo monográfico, bajo la modalidad de investigación documental, se enfoca en el estudio de la complejidad sintáctica en España y en Hispanoamérica desde las diversas posturas teóricas y metodológicas aplicadas por los autores que se han interesado en el análisis de este tema. El interés por este trabajo surgió, debido a que hasta la actualidad, no se han realizado investigaciones sobre complejidad sintáctica en la lengua española en la zona sureste de México. Cabe mencionar, sin embargo, una tesis de Maestría en Educación, cuyo interés fue el efecto del índice de complejidad sintáctica en la lengua materna, el español, en la producción de estructuras sintácticas complejas en una segunda lengua, en este caso, el inglés, por parte de estudiantes de licenciatura de la Universidad de Quintana Roo (Cohuo, 2015).

La complejidad o madurez sintáctica, términos usados indistintamente en este trabajo, comprendida como “la capacidad para producir unidades sintácticas estructuralmente complejas” (Véliz 1988: 108), es un concepto que se desarrolló inicialmente en investigaciones de lengua inglesa, en la década de los setenta, por el autor Kellogg Hunt, quien utilizó un método de medición cuantitativo, sustentado en la aplicación de sus conocidos índices primarios e índices secundarios. Años más tarde, este concepto fue validado en trabajos de lengua española por las autoras A. Torres (1993, 1997) y Véliz (1988, 1999).

El estudio del aspecto sintáctico en la lengua escrita o en la lengua oral permite conocer el nivel de competencia lingüística de los participantes, así como sus dificultades o sus fortalezas para la producción discursiva; por tanto, los resultados de esas investigaciones permiten contribuir en la implementación de mejores técnicas de enseñanza de la lengua española con el fin de que los alumnos incrementen sus habilidades sintácticas, y sean capaces de redactar textos acordes con su nivel educativo. Pese a la importancia de esta temática, en los estudios realizados para determinar el índice de complejidad sintáctica de la lengua española, no se observa un gran desarrollo. Ese fue uno de los motivos que me condujo a elaborar este trabajo monográfico.

Pues aunque existen investigaciones que incluyen los estudios previos que se han realizado sobre esta temática, lo hacen de manera breve y muy general, como en el caso de Bartolomé (2016) y A. Torres (1993), entre otros; por lo cual, se consideró necesario contar con una investigación que se concentre en el análisis de los estudios sobre complejidad sintáctica en español, para

seleccionar el enfoque metodológico y práctico más adecuado en futuras investigaciones con población estudiantil del sureste mexicano, específicamente, del estado de Quintana Roo.

El propósito principal de esta monografía fue elaborar una investigación documental en la que se incluyeran diversos estudios realizados sobre complejidad sintáctica en la lengua española. Me propuse como objetivos: 1) describir minuciosamente los estudios previos sobre complejidad sintáctica realizados fuera de México; y 2) reflexionar con una postura crítica sobre esos estudios, con base en las teorías en los cuales se basaron, la metodología aplicada, los resultados obtenidos y las conclusiones. Es importante mencionar que los estudios realizados en México se omitieron de este trabajo, por cuestiones de tiempo y espacio.

Si bien no se incluyen en este trabajo monográfico las investigaciones producidas en México, sí existen aportes de autores mexicanos. Durante la recolección de material bibliográfico, se encontró un total de 12 estudios sobre complejidad sintáctica realizados en México, 7 artículos y 5 tesis; se observó que la producción de estos trabajos mayormente se presenta en el centro del país, específicamente en Querétaro y en la ciudad de México. El primer trabajo consultado fue el de Herrera (1991), quien investigó la madurez sintáctica de escolares de la ciudad de México; y el más actual fue producido por Alarcón y Auza (2015), quienes analizaron las funciones de los nexos empleados por los niños queretanos que inician su educación primaria. Cabe destacar que la Dra. Alarcón es una investigadora que en los últimos años ha producido diversos aportes en México sobre este tema, mediante la dirección de tesis (Becerra, 2014; López, 2013; Maya, 2014) y la publicación de artículos científicos sobre complejidad sintáctica en español (Alarcón y Auza, 2015; Alarcón y García, 2015).

Aunque inicialmente, se contempló incluir en esta monografía los estudios producidos en México, el número de trabajos analizados se limitó conforme se fue desarrollando esta investigación, según el orden cronológico de las investigaciones realizadas en determinadas regiones hispanohablantes. Teniendo en cuenta que el primer trabajo sobre complejidad sintáctica se preparó en España, se optó por incluir esta región en el primer capítulo, y después de presentar los avances previos de esta investigación (“El estudio de la complejidad sintáctica en España”) en el X Encuentro de Estudiantes de Humanidades, realizado del 18 al 21 de septiembre de 2018 en la Universidad de Quintana Roo, Unidad Académica Chetumal, gracias a los comentarios de la Dra. Edith Hernández Méndez, se decidió excluir a México de Hispanoamérica en la redacción del segundo capítulo, porque ya se contaba con material suficiente de las dos regiones mencionadas.

Se decidió iniciar esta investigación con la lectura de los trabajos pioneros en el mundo hispanohablante, lo cual no impide que en el futuro, se realice otro trabajo monográfico sobre el estudio de la complejidad sintáctica en español, pero basado exclusivamente en las investigaciones producidas en México.

Esta investigación documental se dividió en tres capítulos. En el primer capítulo, se incluyen los estudios realizados en España, entre los cuales, destacan el de Salvador (1985, 1995) y Torres (2001), elaborados en Granada. Así también, el de A. Torres (1993), quien realizó un estudio sobre madurez sintáctica en Tenerife, y cuatro años más tarde, en las Islas Canarias, A. Torres (1997); finalmente, el de Bartolomé (2016) que es el trabajo más reciente hecho en Madrid.

En el segundo capítulo, se encuentran los estudios realizados en Hispanoamérica, esta región posee mayor cantidad de investigaciones sobre complejidad sintáctica, la mayoría fueron realizadas en el país de Chile, en Centroamérica y en el Caribe; no obstante, Hispanoamérica también cuenta con trabajos realizados en Argentina y en otros países de Sudamérica. En Chile, destacan: Balboa, Crespo y Rivadeneira (2012); Coloma, Fernández y Peñaloza (2007); Hugo (2014); Véliz (1988, 1999); y Véliz, Muñoz y Echeverría (1985). El trabajo de Véliz fue considerado pionero al aplicar los índices de Hunt, por lo que sus trabajos fueron sumamente importantes para las investigaciones posteriores.

En Centroamérica y en el Caribe, destacan los trabajos realizados por: Montenegro (1988, 1991), en Santiago de los Caballeros; Espinet (1996) y Rodríguez (1991), en Puerto Rico; y Sánchez (1998), en Costa Rica. La región de Argentina contó con: Delicia (2011); Delicia y Muse (2013); y con Delicia, Fernández, Muse y Porporato (2012). En la sección de otros países de Sudamérica, se incluyeron los trabajos de: Duarte, Espejo y Hernández (2013), en Colombia; e Ibarra (2018) y Trelles (2015), en el Perú. Finalmente, en el capítulo tres, se desarrolla una reflexión crítica sobre los estudios de complejidad sintáctica realizados fuera de México. Y por último, se presentan las conclusiones de este trabajo.

Este estado del arte<sup>1</sup> en el que se pretendió integrar las investigaciones sobre complejidad sintáctica en la lengua española, producidas fuera de México, puede ser una fuente de consulta para alumnos y profesores de la Licenciatura en Humanidades, así como para todo aquel que se interese por el tema de la complejidad sintáctica.

---

<sup>1</sup> Galeano y Vélez afirman que un estado del arte “es una investigación documental sobre la cual se recupera y trasciende reflexivamente el conocimiento acumulado sobre determinado objeto de estudio” (como se citó en Gómez, Galeano y Jaramillo, 2015, p. 424).

# CAPÍTULO 1. ESTUDIOS SOBRE COMPLEJIDAD SINTÁCTICA EN ESPAÑA

La complejidad sintáctica es un tema, cuyo estudio aún no ha alcanzado un gran desarrollo en la lengua española. No obstante, en España sí se han realizado algunas investigaciones lingüísticas, las cuales se sintetizarán en el presente capítulo, por orden cronológico y de acuerdo con la provincia donde se hayan realizado. Por tanto, primero se presentarán los estudios realizados en la provincia de Granada, la cual se ubica en la región de Andalucía; posteriormente, los estudios realizados en las Islas Canarias; y, por último, el estudio más reciente realizado en la provincia de Madrid, que se ubica en la región de Castilla la Nueva.

## 1.1 Estudios sobre complejidad sintáctica realizados en Granada

En la década de los ochenta, Salvador (1985) analizó la relación existente entre los índices de complejidad sintáctica de textos escritos por alumnos del Ciclo Medio de Educación General Básica y las variables independientes de clase social, sexo y curso o nivel instructivo. En esta investigación, basada en la tesis doctoral de Salvador (como se citó en Salvador, 1985), la muestra quedó conformada por 436 alumnos, pertenecientes a los cursos de tercero, cuarto y quinto grado: 218 hombres y 218 mujeres. Con base en el nivel académico y en la profesión de sus padres, se distribuyó a los alumnos en tres clases sociales: 57 en la clase alta, 138 en la clase media y 241 en la clase baja (véase tabla 1).

Tabla 1.

Muestra de los estratos que conforman la población escolar del Ciclo Medio de Educación General Básica

Variable	Nivel	N	%
Curso	3°	114	26.1
	4°	160	36.7
	5°	162	37.2
Sexo	Varón	218	50
	Hembra	218	50

Clase social	Alta	57	13
	Media	138	31.7
	Baja	241	55.3

Fuente: Salvador (1985, p. 63).

La recolección del corpus se llevó a cabo en un día hábil escolar, sin contabilizar el tiempo que tardaron los alumnos en la redacción de un texto de tema libre, el cual se les solicitó a los estudiantes seleccionados, en una hoja de veintiuno por trece cm. Como era de esperarse, los discursos predominantes fueron los narrativos y los descriptivos. Para la evaluación de las composiciones escritas por los estudiantes de la ciudad de Granada, Salvador (1985) creó cuatro índices de complejidad sintáctica: numérica, cualitativa, estructural oracional y estructural proposicional (véase tabla 2), basándose en el índice diseñado por Simon (1973), “cuya fórmula establece una relación entre el número de proposiciones y el número de oraciones de un texto...” (como se citó en Salvador, 1985, p. 61). Por consiguiente, en el análisis de los textos, se definió cada uno de los elementos incluidos en los índices, los cuales se diseñaron con base en la teoría gramatical de tipo tradicional, adoptada por la Real Academia de la Lengua.

Tabla 2.  
Fórmula de los índices de complejidad sintáctica diseñados por Salvador (1985)

N	Fórmula
1	$CP_1 = \frac{\text{n.º de proposiciones}}{\text{n.º de oraciones}} \times 10$
2	$CP_2 = \frac{O - SPL + 2(YXTON + CDON) + 3(SBON - 1) + 4(SBON - 2) + 5(SBON - 3)}{OR - T}$
3	$CP_3 = \frac{CDON + YXTPON}{SBON}$
4	$CP_4 = \frac{CDAS + YXTAS}{SBDAS}$

Nota:  $CP_1$  = índice de complejidad numérica,  $CP_2$  = índice de complejidad cualitativa,  $CP_3$  = índice de complejidad estructural oracional, y  $CP_4$  = índice de complejidad estructural proposicional.

Fuente: Elaboración propia con base en Salvador (1985, pp. 64-65).

Según los resultados de este estudio, las variables clase social y sexo no mostraron índices significativos; el índice medio de complejidad sintáctica de los niños fue superior al de las niñas, por lo que no se validó la hipótesis inicial en la que se planteaba que los textos escritos por las niñas serían más complejos que los escritos por los niños. Tampoco se confirmó que la complejidad sintáctica en los escritos de los alumnos de clase alta fuera mayor que la de los alumnos de clase media, ni que la de estos, fuera más alta que la de los alumnos de clase baja; pues el índice medio

más alto que se obtuvo correspondió a los escritos de los alumnos de clase media. Por el contrario, la variable curso sí fue significativa en esta investigación, pues el índice de complejidad sintáctica se incrementó significativamente de un nivel instructivo a otro (véase tabla 3).

Tabla 3.  
Interacción curso-clase social (CP1)

		A	M	B	F	P
3.º	X	16.8	17.8	20.4	2.24	.06
	G	4.82	5.17	11.5	-	-
4.º	X	20.12	21.09	20.18	-	-
	G	6.57	5.74	6.43	-	-
5.º	X	25.2	23.5	21.7	-	-
	G	6.81	7.52	6.27	-	-

Fuente: Salvador (1985, p. 69).

Salvador (1985) destacó que los alumnos del tercer curso de Educación General Básica utilizaron más estructuras coordinadas y yuxtapuestas en sus escritos, y por esa razón, no alcanzaron el índice medio correspondiente a la estructura binaria, ni por coordinación. Los alumnos del cuarto curso mostraron un equilibrio entre las oraciones coordinadas y las subordinadas, por lo que la estructura numérica característica de sus escritos fue la binaria, y su estructura cualitativa, la coordinación. Los alumnos de quinto curso superaron el índice de complejidad cualitativa de tipo coordinativo, ya que utilizaron más estructuras subordinadas en sus textos. Con respecto a la validez de los índices utilizados para medir la complejidad sintáctica, el autor considera que sus resultados fueron indicios suficientes para aceptar esas fórmulas en este tipo de investigaciones; puesto que los índices de complejidad obtenidos fueron más altos en las estructuras complejas, en contraste con los índices de las estructuras simples, los cuales fueron bajos. Aun así, sugiere continuar el trabajo de validación de estos índices, recurriendo a nuevas estrategias experimentales.

Una década más tarde, en 1995, Salvador realizó un nuevo estudio en la ciudad de Granada, en el que estableció el índice de complejidad sintáctica de textos escritos por alumnos del ciclo superior de la Educación General Básica, y analizó la relación existente entre los índices obtenidos y las variables curso, sexo y clase social. El propósito principal del autor fue verificar si el uso de oraciones subordinadas propiciaba algún incremento de madurez sintáctica en la expresión escrita según el grado cursado, e identificar si el nivel social y el sexo de los estudiantes determinaban variaciones significativas en el uso de estructuras sintácticas oracionales. Salvador (1995) se propuso, además, aportar conocimientos que permitiesen mejorar la competencia lingüística de

esos estudiantes, mediante la implementación de normas didácticas para la enseñanza de la expresión escrita en el ciclo superior.

La muestra de esta investigación quedó conformada por trescientos alumnos procedentes de tres colegios: uno público, Gallego Burín; y dos privados, Regina Mundi y Padre Manjón. Uno se localizaba en zona céntrica, y dos, en barrios periféricos de la ciudad. Se seleccionaron cien alumnos de sexto curso, cien de séptimo y cien de octavo, cincuenta hombres y cincuenta mujeres en cada caso; el 33% del total de la muestra correspondía a la clase alta, y el 67%, a la clase baja. Esta clasificación se asignó con base en la profesión y en el nivel de estudio de los padres.

Conjugando ambos criterios, se establecieron dos niveles en la variable “clase social”: 1) Clase alta: Profesiones relacionadas con los cuadros superiores de mando en la industria, la administración y los servicios. Estudios con titulación superior o de grado medio; 2) Clase baja: Obreros, empleados, con o sin estudios primarios. (Salvador, 1995, p.185)

Para obtener el corpus, Salvador (1995) solicitó la intervención de los maestros de cada grupo, para que ellos les aplicaran la prueba a los alumnos; la cual consistió en la escritura de una composición de un tema elegido por los estudiantes, sin importar el tiempo que tardaran en redactarlo. Para el análisis de los textos recolectados, el autor consideró la oración y la proposición<sup>2</sup> como unidades sintácticas, y estableció las clases de oraciones subordinadas que se considerarían en este estudio (véase tabla 4), con base en la terminología gramatical propuesta por la Administración Educativa (como se citó en Salvador, 1995). “La conceptualización de los términos deriva de una teoría gramatical, elaborada por lingüistas de prestigio y avalada por la Real Academia de la Lengua (1973)” (Salvador, 1995, p. 186).

Tabla 4.  
Variables dependientes consideradas para el análisis de los textos escritos

SBON-T: Oraciones compuestas por subordinación.
SBON-1: Subordinación de 1er. grado.
SBON-2: Subordinación de 2º grado.
SBON-3: Subordinación de 3er. grado.
SBDAS-T: Propositiones subordinadas.
SUST-T: Subordinadas sustantivas.
SUST-S: Sustantivas de sujeto o atributo.
SUST-CP: Sustantivas de complemento verbal.
SUST-CN: Sustantivas de complemento adnominal.
ADJVAS: Subordinadas adjetivas.
CC-T: Subordinadas circunstanciales.

<sup>2</sup> “Esta unidad se define como una secuencia, estructurada en sujeto y predicado, que se combina con otras para construir una unidad superior: la oración compuesta” (Salvador, 1995, p. 186).

Fuente: Salvador (1995, p. 186).

Como se observa en la tabla 4, Salvador (1995) estableció tres tipos de oraciones subordinadas, tomando como referencia el índice de subordinación diseñado por Loban (como se citó en Salvador, 1995). De acuerdo con ese índice, cuando la proposición dependía directamente de la principal, se trataba de una subordinación de primer grado; cuando la proposición dependía de otra subordinada, se trataba de una subordinación de segundo grado; y cuando la proposición dependía de una subordinada, que a su vez dependía de otra, se trataba de una subordinación de tercer grado. Por su parte, las proposiciones subordinadas se clasificaron con base en la función que desempeñaban en la proposición principal: cuando sustituían a un sustantivo o a un adjetivo, se denominaban subordinadas sustantivas o subordinadas adjetivas; cuando funcionaban como complemento circunstancial o sustituían a un adverbio, se denominaban subordinadas circunstanciales o subordinadas adverbiales.

Con base en los resultados obtenidos a través del análisis de varianza, Salvador (1995) concluyó que las variables sexo y curso sí fueron significativas en esta investigación. En cuanto a la variable sexo, el autor estableció que esa variable sí era significativa, debido a que las composiciones de las niñas presentaron índices más altos de complejidad sintáctica que las de los niños, por lo que se validó la hipótesis inicial en la que se planteaba una superioridad en la producción de oraciones subordinadas por parte de las niñas. Por su parte, la variable curso fue considerada como la más significativa, ya que el uso de estructuras sintácticas de subordinación se incrementó según el curso al que pertenecía el alumno, es decir, el nivel de complejidad fue más alto en octavo; más bajo, en séptimo que en octavo; y más bajo, en sexto que en séptimo. Cabe mencionar que la evolución de la madurez sintáctica se presentó mayormente en oraciones subordinadas de segundo y tercer grado, en las sustantivas subjetivas y en las sustantivas de complemento adnominal. Por el contrario, la variable clase social no fue significativa en esta investigación, pues no hubo diferencias estadísticas que favorecieran a los alumnos de clase alta en las composiciones analizadas.

Partiendo de lo expuesto anteriormente, Salvador (1995) sostuvo que el índice de complejidad sintáctica se contrasta más por la variable curso, que por el sexo o la clase social. Finalmente, propuso apoyar en su fase diagnóstica y en su acción correctora, al diseño curricular del área de lengua, con el fin de mejorar el grado de competencia lingüística de los alumnos, tanto en la producción como en la comprensión de textos, ya sean orales o escritos. Pues “sólo si se tiene

un desarrollo suficiente en la aplicación de las relaciones de subordinación, podrá el alumno abordar la lectura comprensiva, haciendo del lenguaje escrito una experiencia de comunicación” (Salvador, 1995, p. 193).

Continuando con los estudios orientados a escolares, M. Torres (2001) realizó en la provincia de Granada una investigación sobre los índices de madurez sintáctica de alumnos procedentes del Curso de Orientación Universitaria, para analizar si existía alguna relación entre los índices obtenidos y las variables sexo, nivel sociocultural, relación con los medios de comunicación, opción (Ciencias o Letras) y nivel académico; aunque en realidad, su objetivo principal fue el análisis de las dos últimas variables. Para conformar la muestra, seleccionó cuarenta estudiantes: veinte pertenecían al área de Ciencias, y veinte, al área de Letras; todos, procedentes del Instituto de Educación Secundaria Pedro Antonio de Alarcón de Guadix de la generación 97-98.

En cuanto a las variables establecidas, M. Torres (2001) diseñó una encuesta para obtener información de carácter sociológico que le permitiera definir qué variables analizar. Asignó el nivel sociocultural con base en la instrucción, la profesión y el ingreso económico de los padres: bajo, medio-bajo, medio-alto y alto. Midió la relación entre el alumno y los medios de comunicación, creando tres niveles con su valor correspondiente: si el contacto informativo del alumno se limitaba a la televisión, nivel bajo; si además de la televisión, tenía acceso a la radio, nivel medio; y si el alumno tenía acceso a la radio, la televisión y la prensa, nivel alto. Asignó el nivel académico con base en los resultados obtenidos en su primera evaluación: bajo, medio-bajo, medio-alto y alto.

Para obtener el corpus, M. Torres (2001) solicitó a los estudiantes seleccionados que en una sesión normal de clase, escribieran una composición que tuviera como mínimo doscientas cincuenta palabras, con un tema previamente establecido: imaginar cuál sería la situación familiar y profesional de cada uno, pasados unos seis o siete años. Para medir el grado de madurez sintáctica de los textos recolectados, la investigadora aplicó dos métodos de medición, con el fin de obtener datos más reales y completos. El primero se basó en la teoría de Hunt (como se citó en M. Torres, 2001), quien propuso dos tipos de índices sintácticos para medir la complejidad oracional: índices primarios e índices secundarios. No obstante, este estudio se centró únicamente en los índices primarios (véase tabla 5), y aplicó su terminología sintáctica basada en dos unidades: la unidad terminal y la cláusula.

Tabla 5.  
Primer método de medición: índices primarios diseñados por Kellogg Hunt

Índice	Fórmula
Promedio de longitud de la unidad terminal	(PAL/U-T)
Promedio de longitud de las cláusulas	(PAL/CL)
Promedio de cláusulas por unidad terminal	(CL/U-T)

Nota: PAL= palabras (total de palabras), U-T= unidad terminal, CL= cláusula.

Fuente: Elaboración propia con base en M. Torres (2001, p. 97).

Para el segundo método, M. Torres (2001) tuvo como referencia la medición de dos unidades básicas de la sintaxis: la oración simple, constituida por un núcleo verbal, y la oración compleja, constituida por dos o más núcleos verbales, en la cual incluyó el conteo de proposiciones dependientes (véase tabla 6).

Tabla 6.  
Segundo método de medición: oración simple y oración compleja

Índice	Fórmula
Porcentaje de estructuras simples: P/ES	ES por 100 entre TO
Porcentaje de estructuras complejas: P/EC	EC por 100 entre TO
Porcentaje de proposiciones dependientes: P/PR	PD por 100 entre (PP más PD)

Nota: ES= estructuras simples, EC= estructuras complejas, TO= total de oraciones, PD= proposición dependiente, PP= proposición principal.

Fuente: Elaboración propia con base en M. Torres (2001, p. 108).

En el primer análisis de textos, M. Torres (2001) primero contó las palabras, después separó las unidades terminales y las cláusulas y las contó; por último, aplicó las fórmulas de la tabla 5. En el segundo análisis, primero contó las oraciones simples y las complejas; posteriormente, analizó las estructuras complejas y determinó la cantidad de proposiciones dependientes de cada texto. Para la segmentación de oraciones, siguió el modelo de análisis sintáctico propuesto por Corral (como se citó en M. Torres, 2001): “1. Oraciones simples. 2. Oración compleja con proposiciones insertadas: llevan transpositor y aquí estarían las completivas, complementarias, adverbiales, causales, finales, comparativas y consecutivas. 3. Coordinación: copulativas y disyuntivas. 4. Interordinadas: adversativas, concesivas, condicionales, ilativas” (p. 98).

Con base en los resultados y en el análisis estadístico aplicado, esta autora concluyó que con el primer método de medición de madurez sintáctica, las variables sexo y nivel académico sí fueron significativas en su investigación. En el promedio de longitud de la unidad terminal y en el promedio de cláusulas por unidad terminal, los índices más altos de madurez sintáctica fueron los de los varones. En el promedio de longitud de las cláusulas, el índice más alto de madurez sintáctica

se obtuvo en el nivel académico alto, y el más bajo, en el nivel académico medio-alto. M. Torres (2001) establece estas diferencias: “entre el nivel académico bajo y el medio alto: 5.99 vs. 5.11,  $p=0.03^3$ ; entre el nivel académico medio bajo- y el medio-alto: 6.09 vs. 5.11,  $p=0.02$ ; entre el nivel académico medio-alto y el alto: 5.11 vs. 6.53,  $p=0.01$ ” (p. 103).

Con el segundo método de medición, sólo la variable sexo resultó significativa en el porcentaje de estructuras complejas y en el de proposiciones dependientes; pues al igual que en el primer análisis, los índices de madurez sintáctica de los varones fueron los más altos. Aunque con este método, la variable nivel académico no mostró datos estadísticamente significativos, cabe destacar que el nivel académico alto también presentó un mayor índice de complejidad sintáctica en el porcentaje de estructuras complejas y en el de proposiciones dependientes. Las variables restantes no resultaron significativas con ninguno de los dos métodos. De manera que la única variable significativa con ambos métodos fue el sexo, pues las composiciones redactadas por los varones presentaron mayor madurez sintáctica en ambos casos; debido a esto, la autora concluyó que no hubo relación asociativa entre los índices de complejidad sintáctica de textos escritos por alumnos de la provincia de Granada y las variables nivel académico y opción (Ciencias/Letras), las cuales habían sido el objetivo principal de la investigación.

## **1.2 Estudios sobre complejidad sintáctica realizados en las Islas Canarias**

En la década de los noventa, A. Torres (1993) escribió su tesis doctoral en la cual estableció el índice de madurez sintáctica de ciento cuarenta alumnos provenientes del cuarto y el octavo curso de Educación General Básica y del Curso de Orientación Universitaria<sup>4</sup> de la zona metropolitana de Tenerife. Su objetivo fue conocer el desarrollo sintáctico de la lengua escrita de los estudiantes, de acuerdo con su nivel de escolaridad, tipo de centro (enseñanza pública y privada), sexo y nivel sociocultural. Esta autora se propuso, además, comparar sus resultados con los de otros trabajos

---

<sup>3</sup> “En el análisis de varianza, se establecen tres grados de significación estadística; de menor a mayor, se presenta así:  $p0,05$ ;  $p0,01$ ;  $p0,001$ ” (M. Torres, 2001, p. 103).

<sup>4</sup> “Según la Ley Orgánica de Ordenación del Sistema Educativo (LOGSE), el curso de cuarto corresponde al último curso del segundo ciclo de la Educación Primaria; octavo coincide con el último grado del primer ciclo de la Educación Secundaria Obligatoria; COU, con el último curso del Bachillerato” (A. Torres, 1993, p. 65).

realizados en Hispanoamérica en los que también se hubiese aplicado la fórmula de Hunt, para determinar si en el mundo hispánico, existían características particulares o ciertas constantes respecto a este tema.

Para llevar a cabo esta investigación, A. Torres (1993) seleccionó sólo los centros de Educación General Básica que incluyeran el octavo curso y los centros de Bachillerato Unificado y Polivalente que impartieran el Curso de Orientación Universitaria, porque quiso abarcar tanto la enseñanza general básica como el bachillerato. Por tanto, la muestra se limitó a ocho colegios de Tenerife, de los cuales, cinco pertenecían al sector público y tres, al sector privado (véase tabla 7).

Tabla 7.  
Número de centros seleccionados para la muestra

	EGB	BACHILLERATO
PÚBLICA	4	1
PRIVADA	2	1

Fuente: A. Torres (1992, p. 64).

De acuerdo con la tabla 8, y teniendo en cuenta que de cada curso, se seleccionaron 10 alumnos, la muestra quedó conformada por 40 alumnos del cuarto curso de Educación General Básica de centros públicos: 20 pertenecían al género femenino, y 20, al género masculino, todos, con un promedio de nueve años y ocho meses de edad. Del mismo curso y nivel académico, pero en centros privados, se seleccionó a 20 alumnos: 10 hombres con un promedio de nueve años y siete meses de edad, y 10 mujeres con un promedio de nueve años y ocho meses de edad.

Además, se incluyó a 40 alumnos del octavo curso de Educación General Básica de centros públicos: 20 hombres y 20 mujeres con una edad promedio de catorce años y un mes en ambos casos. Del mismo curso y nivel educativo, pero en un centro privado, se seleccionó a 20 alumnos, de los cuales, 10 eran hombres con un promedio de catorce años y un mes de edad, y 10 eran mujeres de trece años y cinco meses de edad promedio. Y por último, a 10 alumnos del curso de Orientación Universitaria: 5 hombres con un promedio de edad de dieciocho años y tres meses, y 5 mujeres con un promedio de edad de diecisiete años y ocho meses. Del mismo curso y nivel académico, pero en un centro privado, se seleccionó a 10 alumnos: 5 hombres con un promedio de edad de diecisiete años y seis meses, y 5 mujeres con un promedio de edad de diecisiete años y siete meses (véase tabla 8).

Tabla 8.  
Número de sujetos según los cursos y la edad promedio

CURSOS	PÚBLICA	PRIVADA
4.º	40 20 hombres: 9 años y 8 meses 20 mujeres: 9 años y 8 meses	20 10 hombres: 9 años y 7 meses 10 mujeres: 9 años y 8 meses
8.º	40 20 hombres: 14 años y 1 mes 20 mujeres: 14 años y 1 mes	20 10 hombres: 14 años y 1 mes 10 mujeres: 13 años y 5 meses
COU	10 5 hombres: 18 años y 3 meses 5 mujeres: 17 años y 8 meses	10 5 hombres: 17 años y 6 meses 5 mujeres: 17 años y 7 meses

Fuente: Elaboración propia con base en A. Torres (1993).

De los ciento cuarenta alumnos seleccionados, el 57.2% tiene un perfil sociocultural bajo, el 25.7% cuenta con un perfil sociocultural medio, y sólo el 17.1% se clasificó con un perfil sociocultural alto. El nivel sociocultural se asignó con base en dos indicadores: 1) el grado de escolaridad de sus padres y 2) su ocupación. El primer indicador se asignó de acuerdo con "... la clasificación establecida por Moreno Becerra (1985: 125), quien indica que la educación en Canarias puede dividirse en tres niveles: 1.- Instrucción primaria o menos, 2.- Estudios secundarios (terminados o incompletos), 3.- Estudios superiores (terminados o incompletos)" (A. Torres 1993, pp.74-75); y el segundo indicador, de acuerdo con "... la tipología de ocupaciones de los padres de los escolares siguiendo las instrucciones de la Encuesta de Población Activa (INE, 1987: 30)" (A. Torres, 1993, p. 75).

Para la recolección del corpus, se le solicitó a cada uno de los alumnos seleccionados, una composición de tema libre que incluyera como mínimo doscientas palabras, y para el análisis de dichos textos, se tuvo como referencia la teoría de Hunt (como se citó en A. Torres, 1993), quien propuso dos tipos de índices sintácticos para medir la complejidad oracional: índices primarios e índices secundarios. En el primer grupo, Hunt estableció tres índices primarios para medir la complejidad sintáctica de los escritos: 1) promedio de longitud de la unidad terminal<sup>5</sup>, 2) promedio de longitud de las cláusulas<sup>6</sup>, y 3) promedio de cláusulas por unidad terminal<sup>7</sup> o índice de

5 "... establece una relación asociativa entre madurez sintáctica y unidades terminales más largas" (A. Torres, 1993, p. 55).

6 "... se apoya en que existe una relación asociativa entre el aumento del número de palabras por cláusula y mayor complejidad sintáctica" (A. Torres, 1993, p. 56).

7 "... se basa en que la madurez sintáctica depende del número de subordinaciones que se incrusten en la unidad terminal" (A. Torres, 1993, p. 57).

subordinación (véase tabla 9). En el segundo grupo, estableció dos índices secundarios, los índices secundarios clausales: 4) promedio de cláusulas adjetivas por unidad terminal, 5) promedio de cláusulas sustantivas por unidad terminal, y 6) promedio de cláusulas adverbiales por unidad terminal (véase tabla 10); así como los índices secundarios no clausales: 7) promedio de adjetivos calificativos por unidad terminal, 8) promedio de adjetivos posesivos por unidad terminal, 9) promedio de frases preposicionales por unidad terminal, y 10) promedio de aposiciones por unidad terminal (véase tabla 11).

Tabla 9.  
Fórmula de los índices primarios diseñados por Kellogg Hunt

N	FÓRMULA
1.	$L / U-T$ (Total de palabras / total de unidades terminales)
2.	$L / CL$ (Total de palabras / total de cláusulas)
3.	$CL / U-T$ (Total de cláusulas / total de unidades terminales)

Nota: L= longitud (total de palabras), U-T= unidad terminal, CL= cláusula.

Fuente: Elaboración propia con base en A. Torres (1993, pp. 55-57).

Tabla 10.  
Fórmula de los índices secundarios clausales diseñados por Kellogg Hunt

N	FÓRMULA
4.	$CL\ ADJ / U-T$
5.	$CL\ SUST / U-T$
6.	$CL\ ADV / U-T$

Nota: CL ADJ= cláusula adjetiva, CL SUST= cláusula sustantiva, CL ADV= cláusula adverbial U-T= unidad terminal.

Fuente: Elaboración propia con base en A. Torres (1993, p.61).

Tabla 11.  
Fórmula de los índices secundarios no clausales diseñados por Kellogg Hunt

N	FÓRMULA
7.	$CAL / U-T$
8.	$POS / U-T$
9.	$FR\ PREP / U-T$
10.	$APOSICIONES / U-T$

Nota: CAL= calificativos, POS= posesivos, FR PREP= frases preposicionales, APOSICIONES= aposiciones, U-T= unidad terminal.

Fuente: Elaboración propia con base en A. Torres (1993, p. 62).

Como resultado de este estudio, A. Torres (1993) concluyó que las variables tipo de centro, sexo y nivel sociocultural no fueron significativas en esta investigación; dado que no presentaron una relación asociativa con los índices de madurez sintáctica, como se esperaba, excepto en algunos casos, aunque con diferencias poco significativas. Por ejemplo, con la variable tipo de centro, la

relación se produjo sólo con los índices sintácticos: 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas, y 7) promedio de adjetivos calificativos por unidad terminal; con la variable sexo, se produjo una intrascendente relación con los índices sintácticos: 3) promedio de cláusulas por unidad terminal, 5) promedio de cláusulas sustantivas por unidad terminal, y 6) promedio de cláusulas adverbiales por unidad terminal. En cuanto a la variable nivel sociocultural, sólo se produjo relación con el índice 7) promedio de adjetivos calificativos por unidad terminal, pues ese fue el único índice que aumentó de acuerdo con los niveles socioculturales de cada grado (véase tabla 12).

Tabla 12.

Medias de los índices primarios y secundarios que no presentaron relación asociativa con las variables centro, sexo y nivel sociocultural

VARIABLE	ÍNDICE	4°		8°		COU	
		ENSEÑANZA	ENSEÑANZA	ENSEÑANZA	ENSEÑANZA	ENSEÑANZA	ENSEÑANZA
CENTRO	L U-T	9.06	9.56	11.03	13.97	12.98	15.39
	L CL	-	-	-	-	-	-
	CAL U-T	.26	.34	.47	.65	.68	.77
SEXO	CL U-T	1.76	1.80	2.03	2.31	2.01	2.34
	CL SUST	.30	.37	.42	.67	.37	.45
	CL ADV	.31	.34	.40	.43	.38	.48
NIVEL SOCIO-CULTURAL	CAL U-T	-	-	-	-	-	-

Nota: L= longitud, U-T= unidad terminal, CL= cláusula, CAL= calificativos, CL SUST= cláusula sustantiva, CL ADV= cláusula adverbial.

Fuente: Elaboración propia con base en A. Torres (1993).

En contraste con las anteriores variables, el nivel de escolaridad sí fue significativo en esta investigación, pues los índices de madurez sintáctica de los escolares se fueron incrementando a medida que los alumnos avanzaban de grado, lo cual coincidió con la hipótesis inicial planteada por A. Torres (1993): “El desarrollo sintáctico de los escolares -su madurez sintáctica- va aumentando de acuerdo con el crecimiento cronológico y mental que, se supone, experimentan al avanzar en su proceso de escolarización” (p.41). El crecimiento fue progresivo en los tres cursos estudiados, excepto en el caso de las cláusulas sustantivas (cuarto .34, octavo .54, COU .41) y en las aposiciones (cuarto .0356, octavo .1138, COU 0.506), cuyos resultados demuestran que los alumnos de octavo superan a los del Curso de Orientación Universitaria.

Después de haber comparado los resultados obtenidos en este estudio, con los de otras investigaciones sobre madurez sintáctica de escolares realizadas en el mundo hispánico, A. Torres (1993) señaló que tanto en los trabajos realizados en inglés como en los realizados en español, los alumnos presentan un aumento progresivo según se avanza de nivel escolar; además, demostró que existen unas constantes que se mantienen relativamente fijas en español en cuanto a los tres índices primarios de madurez sintáctica.

En 1997, A. Torres realizó una nueva investigación; sin embargo, en esta ocasión, ya no se centró en estudiantes, sino en escritores de las Islas Canarias. Con este estudio, A. Torres (1997) buscó establecer el grado de complejidad sintáctica de textos narrativos y descriptivos escritos por autores canarios, tratando de comprobar si existía alguna relación asociativa entre estos índices y las variables modo de discurso y sexo; de igual manera, esta autora se propuso comprobar la fiabilidad de los resultados obtenidos, comparando las diferencias entre las medias de textos producidos por el mismo autor, pero de diferente modo de discurso. La muestra estuvo integrada por 13 escritores, 8 hombres y 5 mujeres; se eligieron dos textos de cada modalidad por cada participante, se tuvo en cuenta que cada uno incluyera más de doscientas palabras, no obstante, para el análisis de los datos, solamente se utilizaron cuarenta textos del total. De manera que el corpus quedó conformado por cuarenta textos, veinte narrativos y veinte descriptivos, los cuales incluían en total 8,591 palabras (véase tabla 13).

Tabla 13.  
Total de palabras, de unidades terminales y de cláusulas según el modo de discurso

	PALABRAS	U-T	CLÁUSULAS
TEXTOS NARRATIVOS	4.252	291	678
TEXTOS DESCRIPTIVOS	4.339	242	600

Fuente: A. Torres (1997, p. 110).

Para el análisis de los textos seleccionados, la autora se apoyó nuevamente en los presupuestos de Kellogg Hunt (como se citó en A. Torres, 1997), como en su estudio de 1993 (véase página 19). En la segmentación del corpus, A. Torres (1997) contó las palabras de cada narración para obtener el total de unidades terminales y de cláusulas, así como el promedio de cláusulas por unidad terminal; posteriormente, identificó los índices secundarios clausales, y finalmente, los no clausales o modificadores nominales. Siguiendo lo establecido por Moreno (como se citó en A. Torres, 1997), para el análisis estadístico de los datos obtenidos en este estudio, fue necesario cumplir con dos objetivos: 1) describir y resumir los datos, y 2) hacer estimaciones

de significación y fiabilidad. Para cumplir con tales objetivos, la autora tuvo en cuenta la media, la desviación estándar y el rango.

Los resultados de la investigación de A. Torres (1997) mostraron que la variable modo de discurso sí produjo una relación asociativa con los índices establecidos. En el caso de los índices primarios, esa variable sólo fue significativa para el promedio de longitud de las cláusulas; y en el de los índices secundarios clausales, para el promedio de cláusulas adjetivas por unidad terminal. En el caso de los índices secundarios no clausales o modificadores nominales, la variable modo de discurso fue significativa para el promedio de adjetivos calificativos por unidad terminal, y para el de frases preposicionales por unidad terminal. Cabe mencionar que con los índices primarios, los resultados significativos se presentaron en ambos tipos de texto, narrativos y descriptivos; pero con los índices secundarios, sólo hubo resultados significativos en los textos descriptivos.

Por el contrario, la variable sexo no fue significativa en este estudio; pues aunque los varones obtuvieron resultados más altos en los índices primarios y en los secundarios, esas diferencias no fueron estadísticamente significativas al aplicar la prueba de medias. Respecto a la fiabilidad de los resultados, la autora concluyó que los índices propuestos por Hunt son medidas válidas y estables cuando se aplican a textos narrativos y a textos descriptivos de escritores canarios; puesto que no se observan diferencias significativas entre las medias del texto narrativo y las del texto descriptivo producido por un mismo autor.

### **1.3 Estudios sobre complejidad sintáctica realizados en Madrid**

Recientemente, Bartolomé (2016) determinó en su tesis doctoral el índice de madurez sintáctica de composiciones argumentativas y composiciones descriptivas escritas por estudiantes de centros educativos de la provincia de Zamora. Su objetivo principal fue investigar si existía alguna correlación entre los resultados obtenidos y la influencia de la variable horas de exposición a las tecnologías de la información y la comunicación; asimismo, investigó si esa correlación también se producía con las variables curso, sexo, centro (público o privado) y ubicación (rural o urbana). Además, comparó sus resultados con los de otras investigaciones realizadas con la lengua española en la década de los noventa, para saber si se había producido un aumento del índice de complejidad

sintáctica desde esa época hasta la fecha de su estudio, o por el contrario, si este había disminuido por el uso de la tecnología.

En cuanto a las variables establecidas, Bartolomé (2016) diseñó un cuestionario de opciones múltiples para obtener los datos sociológicos de cada participante, el cual incluía dos preguntas principales sobre qué medios de internet eran los más utilizados por los alumnos, y cuántas horas a la semana les dedicaban; puesto que esa información era relevante para su objetivo de investigación (véase figura 1). De acuerdo con los resultados del cuestionario, asignó tres distinciones al número de horas que los alumnos dedicaban a navegar en internet y en las redes sociales, ya fuera en un dispositivo móvil, en una tableta o en su computadora: P, R y M, abreviaturas de “poco”, “regular” y “mucho”. Si los estudiantes se exponían de 0 a 9 horas a las tecnologías de la información y la comunicación, Bartolomé consideraba ese tiempo como “poco”; si el tiempo era entre 10 y 19 horas por semana, lo consideraba “regular”; y cuando el tiempo dedicado a las tecnologías de la información y la comunicación era de 20 o más horas por semana, lo calificaba como “mucho”.

Figura 1.  
Preguntas principales del cuestionario diseñado

8. ¿Cuáles de estos medios de internet utilizas?					
<input type="checkbox"/> Correo electrónico	<input type="checkbox"/> Facebook	<input type="checkbox"/> Tuenti	<input type="checkbox"/> Twitter	<input type="checkbox"/> Foros	
<input type="checkbox"/> Blogs	<input type="checkbox"/> Chats (Messenger, WhatsApp, Viber, Line)		<input type="checkbox"/> Página web personal		
<input type="checkbox"/> Otros: _____					
9. ¿Cuántas horas dedicas a la semana a estos medios de internet?					
<input type="checkbox"/> Cinco	<input type="checkbox"/> Diez	<input type="checkbox"/> Quince	<input type="checkbox"/> Veinte	<input type="checkbox"/> Veinticinco	<input type="checkbox"/> Treinta
<input type="checkbox"/> Más de treinta: _____					

Fuente: Bartolomé (2016, p. 143).

La muestra quedó conformada por 191 alumnos, con un rango de edad entre nueve, trece y diecisiete años; procedentes del cuarto curso de Educación Primaria, segundo curso de Educación Secundaria Obligatoria y segundo curso de Bachillerato. Para obtener el corpus, la investigadora pidió a los profesores de lengua española que solicitaran a sus estudiantes una composición manuscrita de modo argumentativa, y otra, de modo descriptivo; ambas con una extensión mínima de doscientas palabras. Sin embargo, el promedio de palabras fue de 176 para los textos

argumentativos, y 182, para los descriptivos. El corpus total fue de 382 composiciones; las cuales incluían 68,254 palabras, 5,415 unidades terminales y 10,839 cláusulas (véase tabla 14).

Tabla 14.  
Corpus global desglosado según el modo de discurso

	Palabras	Unidades-t	Cláusulas
Textos argumentativos	33.520	2.480	5.530
Textos descriptivos	34.684	2.935	5.309
Total	68.204	5.415	10.839

Fuente: Bartolomé (2016, p. 142).

Los alumnos debían escoger para sus composiciones, alguno de los cuatro temas propuestos por Bartolomé (2016). Para el modo argumentativo: 1) no puedo vivir sin mi teléfono móvil ni internet, 2) el móvil, el iPad o el ordenador no son importantes para mí, 3) es normal que los padres castiguen a sus hijos y 4) me parece mal que no se pueda utilizar el móvil en el colegio/instituto; para el modo descriptivo: 1) describe tu serie de televisión favorita, 2) habla sobre tu libro favorito, 3) describe tu videojuego favorito y 4) describe tu pueblo o ciudad favoritos. Para el análisis del corpus, la autora partió de las fórmulas propuestas por Hunt (como se citó en Bartolomé, 2016), por lo que aplicó sus índices primarios y los secundarios, siguiendo su terminología sintáctica que se basaba en la unidad terminal y en la cláusula. No obstante, definió estas unidades según los conceptos planteados por Véliz y Torres (como se citó en Bartolomé, 2016).

Por consiguiente, primero dividió cada texto en unidades terminales y en cláusulas, después contó el total de palabras; y posteriormente, contó la cantidad de oraciones subordinadas sustantivas, adjetivas y adverbiales. Por último, contó los adjetivos calificativos, los posesivos, las aposiciones y las frases preposicionales complementos de sustantivos. Para el procesamiento de los datos, la autora realizó un análisis estadístico inferencial<sup>8</sup>, y clasificó las correlaciones como débiles, moderadas y altas. Consideró una correlación como débil, cuando el resultado obtenido era entre 0 y 0.35; moderada, cuando el resultado variaba entre 0.35 y 0.75; y alta, cuando oscilaba entre 0.75 y 1. Estableció como negativa, una correlación entre mayor tiempo de exposición a internet y menores índices de madurez sintáctica; y como positiva, una correlación entre menor tiempo de exposición a internet y mayores índices de madurez sintáctica.

---

<sup>8</sup> Para realizar este análisis estadístico, la autora aplicó la técnica denominada regresión logística, la cual permitió determinar qué variables contribuyeron a clasificar a los sujetos como pertenecientes a un grupo u otro, según la intensidad de su uso de las tecnologías de la información y la comunicación (Bartolomé, 2016, p. 163).

Según los resultados del estudio, el modo de discurso de textos argumentativos no fue significativo; pues en ese tipo de composiciones, sólo hubo ocho resultados significativos en las variables centro, sexo y ubicación, de los cuales, seis pertenecían a correlaciones negativas, y dos, a correlaciones positivas. Estas últimas sólo se presentaron con los índices sintácticos: 1) promedio de cláusulas adverbiales por unidad terminal en alumnos de cuarto curso de Educación Primaria, y 2) promedio de longitud de las cláusulas en alumnos de segundo curso de Educación Secundaria Obligatoria; ambas correlaciones positivas fueron moderadas. Por el contrario, el modo de discurso de textos descriptivos sí fue significativo; pues los resultados mostraron 11 resultados significativos, 5 de los cuales pertenecían a correlaciones negativas, y 6, a correlaciones positivas. Y en este caso, a diferencia del modo argumentativo que distribuía sus resultados en variables indistintas, cuatro de las seis correlaciones positivas fueron débiles, y se presentaron en el mismo índice sintáctico en el mismo curso: promedio de cláusulas adverbiales por unidad terminal en alumnos de cuarto curso de Educación Primaria.

En cuanto a la correlación entre los índices de madurez sintáctica y la influencia de la variable horas de exposición a las tecnologías de la información y la comunicación, Bartolomé (2016) concluyó que esa variable no influyó en los índices de complejidad sintáctica, porque el valor de correlación (débil, moderada y alta) no fue uniforme según los grados de exposición a la tecnología (poco, regular y mucho); no obstante, aclaró que desde un punto de vista cuantitativo, se observó que los estudiantes que pasaban más horas expuestos al uso de las nuevas tecnologías tendían a escribir menos palabras por unidad terminal y por cláusula, y menos adjetivos calificativos y cláusulas adjetivas por unidad terminal; por lo que sugirió realizar en el futuro, un estudio con mayor número de composiciones para verificar estos resultados.

Las variables curso, sexo, centro (público o privado) y ubicación (rural o urbana) sí fueron significativos, pues un 97% de los índices de madurez sintáctica de los alumnos se incrementó conforme avanzaban de curso, y en el 77% de los casos, los estudiantes de centros privados produjeron textos con mayor complejidad sintáctica que los de centros públicos. También hubo diferencias significativas entre los textos escritos por alumnos de diferente género: en cuarto grado de Educación Primaria, las niñas obtuvieron índices más altos de complejidad sintáctica en sus composiciones; no obstante, en la secundaria y en el bachillerato, los niños obtuvieron mejores resultados. Con respecto a la variable ubicación, el 93% de los índices de complejidad sintáctica más altos pertenecieron a alumnos de escuelas urbanas.

Para comparar los resultados de este estudio con los de investigaciones anteriores, Bartolomé (2016) analizó los índices de madurez sintáctica unificando las composiciones argumentativas con las descriptivas, pues tuvo en cuenta que los textos de los demás estudios tenían diversas modalidades discursivas. Al realizar el análisis agrupando ambos tipos de texto, observó que por contar con mayor número de composiciones, las correlaciones mostraban resultados más significativos, por lo cual concluyó que el tamaño del corpus es importante en este tipo de investigaciones. Después de cotejar sus resultados con los de Olloqui, Herrera y Torres (como se citó en Bartolomé, 2016), la autora estableció que no se observó gran diferencia entre su estudio y los de investigaciones realizadas antes de la revolución tecnológica, y por esa razón, no descarta realizar más adelante, un estudio cualitativo sobre este tema, para mostrar resultados concretos de la influencia de la variable horas de exposición a las tecnologías de la información y la comunicación, en la complejidad sintáctica de composiciones escritas.

## **CAPÍTULO 2. ESTUDIOS SOBRE COMPLEJIDAD SINTÁCTICA EN HISPANOAMÉRICA**

Hispanoamérica también cuenta con aportes sobre complejidad sintáctica en la lengua española. Para este trabajo se consultaron 17 estudios producidos en países hispanohablantes, los cuales se expondrán en el presente capítulo de manera cronológica, según la región o según el país donde se realizaron. Así, se iniciará con las investigaciones producidas en Centroamérica y en el Caribe (Puerto Rico, República Dominicana y Costa Rica), se continuará con las concernientes al país de Chile, con los estudios realizados en Argentina; y finalmente, se presentarán los estudios que fueron elaborados en otros países de Sudamérica (Colombia y Perú).

### **2.1 Estudios sobre complejidad sintáctica realizados en Centroamérica y en el Caribe**

En la República Dominicana, Montenegro (1988) midió la competencia lingüística de discursos escritos por alumnos de educación primaria y de educación secundaria, con el fin de analizar la relación existente entre los índices obtenidos y las variables edad, curso y estrato social. La muestra de esta investigación incluyó 288 estudiantes procedentes de cuarto, sexto y octavo de primaria, así como de segundo y cuarto de secundaria. Sus edades promedio, ordenadas de menor a mayor, según el grado cursado, fueron las siguientes: nueve años siete meses para cuarto de primaria, once años cuatro meses para sexto de primaria, trece años cuatro meses para octavo de primaria, quince años cuatro meses para segundo de secundaria y diecisiete años ocho meses para cuarto de secundaria (véase tabla 15). Con base en el tipo de institución a la que acudían y teniendo en cuenta que tuvieran el mismo sistema de educación, se determinó el estrato social: escuela pública y colegio privado, 144 alumnos en cada caso.

Tabla 15.  
Número de alumnos según la edad y el estrato social

Nivel	Curso	Escuela Pública	Colegio Privado	Edad Promedio
-------	-------	-----------------	-----------------	---------------

Primaria	4°	30	30	9 años, 7 meses
	6°	30	30	11 años, 4 meses
	8°	30	30	13 años, 4 meses
Secundaria	2°	30	30	15 años, 4 meses
	4°	24	24	17 años, 8 meses
	Total	144	144	

Fuente: Elaboración propia con base en Montenegro (1988).

Para la recolección del corpus, Montenegro (1988) solicitó a los alumnos que escribieran una composición narrativa y, del total obtenido, eligió treinta narraciones de cada curso, a excepción del cuarto grado de secundaria, en el que sólo obtuvo veinticuatro narraciones. Si bien la variable sexo no fue considerada en esta investigación; la investigadora tuvo en cuenta que el 50% del total de su muestra fuera escrito por hombres, y el otro 50%, por mujeres. En cuanto a la metodología aplicada para el análisis de los textos, la investigadora tuvo como referencia los tres índices primarios propuestos por Hunt (como se citó en Montenegro, 1988): 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas, y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22).

Con respecto a la variable edad, los resultados de esta investigación mostraron que el promedio global de los tres índices primarios se incrementó en ambas instituciones educativas, según el curso estudiado. En el caso del índice 1) promedio de longitud de la unidad terminal y el 2) promedio de longitud de las cláusulas, el incremento más elevado de los alumnos se presentó con el cambio de nivel educativo: en el primer índice, la diferencia fue de 2.05 (octavo de primaria, 9.30 y segundo de secundaria, 11.35), y en el segundo, la diferencia fue de 0.98 (octavo de primaria, 6.53 y segundo de secundaria, 7.51). Esta tendencia no se observó en el índice 3) promedio de cláusulas por unidad terminal; aun así, el promedio aumentó a medida que se cambiaba de curso. Como consecuencia, Montenegro (1988) determinó que las variables edad y curso sí produjeron una relación asociativa con los índices establecidos, pues tal como se observa en la tabla 16, el índice de complejidad sintáctica se incrementó significativamente de un grado a otro, y el índice más alto se presentó en los alumnos de cursos superiores.

Tabla 16.  
Resultados generales de los índices primarios, según las variables edad, curso, sexo y estrato social

Índices	Sexo	PAL/UT			PAL/CL			CL/UT		
		Pública	Privada	Total	Pública	Privada	Total	Pública	Privada	Total
Primaria	4°	7.55	8.20	7.88	6.07	6.36	6.22	1.23	1.30	1.27

	6°	8.61	8.91	8.76	6.38	6.63	6.51	1.37	1.31	1.34
	8°	9.36	9.23	9.30	6.64	6.42	6.53	1.42	1.45	1.43
Secundaria	2°	11.66	11.04	11.35	7.49	7.52	7.51	1.55	1.47	1.51
	4°	13.14	12.10	12.62	8.37	8.09	7.97	1.67	1.51	1.59

Nota: PAL/UT: promedio de longitud de la unidad terminal, PAL/CL: promedio de longitud de las cláusulas, CL/UT: promedio de cláusulas por unidad terminal.

Fuente: Elaboración propia con base en Montenegro (1988).

En la tabla 16, se observa que la variable estrato social mostró incrementos poco significativos entre los alumnos procedentes de escuela pública y los del colegio privado, en los tres índices primarios. Por esta razón, la investigadora no estableció una relación asociativa entre esta variable y los índices establecidos, lo cual le permitió corroborar su hipótesis inicial: “La adquisición de la sintaxis no está sujeta a la variable socio-económica; por lo tanto, no se hallarán diferencias sustanciales entre los sujetos de la escuela pública y los del colegio privado” (Montenegro, 1998, p. 543). Por otra parte, Montenegro (1988) propuso que sus resultados contribuyeran en la implementación de mejores técnicas de enseñanza de la lengua española, con el fin de incrementar la comprensión y la producción de discursos escritos de los alumnos.

Tres años más tarde, la misma autora, Montenegro (1991), analizó la relación existente entre los índices de complejidad sintáctica de textos escritos por alumnos de escuelas primarias ubicadas en la ciudad de Santiago y las variables independientes edad, curso, sexo y estrato social. La muestra de este estudio quedó conformada por ciento veinte alumnos de sexto y octavo grado, los de nivel avanzado tenían un promedio de trece años, cuatro meses de edad, y los de menor grado, un promedio de once años, cuatro meses de edad. Se estableció el estrato social con base en el tipo de escuela a la que acudían los alumnos, pública o privada, teniendo en cuenta que las escuelas participantes pertenecieran a las regiones de mayor o menor posición económica, y que contaran con el mismo sistema educativo. Para la obtención del corpus, Montenegro (1991) solicitó a cada uno de los alumnos que realizara dos composiciones escritas con los temas previamente establecidos; en la primera, podían escribir sobre algún paseo familiar o amistoso, y en la segunda, podían redactar sobre su película, libro o serie favorita. Para el análisis de dichos textos, la autora tuvo como referencia la teoría de Kellogg Hunt (como se citó en Montenegro, 1991), en la cual se proponen índices primarios e índices secundarios para medir la complejidad oracional; sin embargo, el estudio de Montenegro (1991) se centró únicamente en analizar los índices primarios (véase tabla 9, p. 22).

Los resultados indican que el crecimiento fue progresivo en los dos cursos estudiados, tanto en escuela pública como en escuela privada; el total global en el caso del promedio de longitud de unidades terminales fue de 9.30 en octavo, en contraste con 8.76 en sexto; en el promedio de longitud de la cláusula, fue de 6.53 en octavo, en contraste con 6.51 en sexto; y en el promedio de cláusulas por unidad terminal, fue de 1.43 en octavo, en contraste con 1.34 en sexto (véase tabla 17). Estos resultados condujeron a la autora a concluir que las variables edad y curso sí fueron significativas en esta investigación, pues conforme los alumnos tenían mayor edad y más alto grado escolar, realizaban composiciones más complejas. En este caso, el nivel de complejidad fue más alto en octavo que en sexto grado, lo cual coincidió con la hipótesis inicial planteada por Montenegro (1991): “Existe un crecimiento sintáctico que depende de la edad de los individuos; por lo tanto, los índices de madurez sintáctica aplicados a textos producidos por escolares de diferentes edades variarán y serán significativamente más altos en los niños de cursos superiores” (p. 115).

Tabla 17.  
Resultados generales de los índices primarios, según las variables edad, curso, sexo y estrato social

Índices	Sexo	Escuela pública			Escuela privada			Total global
		F	M	Total	F	M	Total	
PAL/UT	6°	8.47	8.76	8.61	9.72	8.10	8.91	8.76
PAL/UT	8°	9.37	9.35	9.36	8.66	9.80	9.23	9.30
PAL/CL	6°	6.26	6.49	6.38	6.83	6.43	6.63	6.51
PAL/CL	8°	6.39	6.88	6.64	6.08	6.76	6.42	6.53
CL/UT	6°	1.36	1.37	1.37	1.35	1.27	1.31	1.34
CL/UT	8°	1.47	1.36	1.42	1.42	1.45	1.44	1.43

Nota: PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas, CL/UT= promedio de cláusulas por unidad terminal, F= femenino y M= masculino.

Fuente: Elaboración propia con base en Montenegro (1991).

Por el contrario, las variables de estrato social y sexo no fueron significativas en esta investigación, pues según el análisis de varianza, ambos factores presentaron diferencias poco significativas entre el desarrollo sintáctico de textos producidos por alumnos procedentes de una escuela pública, y el de textos producidos por alumnos de una escuela privada. Como puede apreciarse en la tabla 17, la variable estrato social presentó resultados inconsistentes en los tres índices analizados; pues en el índice promedio de longitud de las unidades terminales, los alumnos de octavo grado procedentes de una escuela pública presentaron resultados más altos que los de

escuela privada (9.36, en contraste con 9.23); no obstante, en sexto grado, los promedios más altos fueron los de los alumnos de escuela privada (8.91, en contraste con 8.61).

Asimismo, en el índice promedio de longitud de las cláusulas, fueron los alumnos de octavo grado de una escuela pública, quienes nuevamente obtuvieron los promedios más elevados (6.64, en contraste con 6.42); en los alumnos de sexto grado, también se repitió el patrón observado en el índice anterior, ya que los alumnos de escuela privada fueron quienes obtuvieron promedios más altos (6.63, en contraste con 6.38). No obstante, en el índice promedio de cláusulas por unidad terminal, cambió esta tendencia; los índices más altos para el octavo grado ya no fueron los de los alumnos de escuela pública, sino los de escuela privada (1.44, en contraste con 1.42), y para el sexto grado, ya no fueron los de los alumnos de escuela privada, sino los de los alumnos de escuela pública (1.37, en contraste con 1.31). De acuerdo con estos resultados, se confirmó una vez más otra de las hipótesis de Montenegro (1991), quien sostuvo que "...la adquisición de la sintaxis no está totalmente sujeta a las variables socioeconómicas" (p. 116).

En lo que respecta a la variable sexo, la autora observó que la tendencia inicial apuntaba a que quienes utilizaban más palabras por unidad terminal y por cláusula eran los hombres; sin embargo, este panorama cambió en la escuela privada, donde las mujeres de sexto grado obtuvieron índices más altos que los hombres (véase tabla 17). Para concluir, la autora estableció los índices de Hunt como medidas válidas para el análisis cuantitativo de composiciones narrativas, y sugirió, continuar en el futuro con el estudio de la madurez sintáctica, para tratar de mejorar la expresión escrita de los estudiantes.

En Puerto Rico, Rodríguez (1991) realizó una investigación sobre los índices de madurez sintáctica de alumnos procedentes de un colegio de educación primaria ubicado en la ciudad de Caguas, para analizar si existía alguna relación entre los índices obtenidos y las variables grado escolar, sexo y modo del discurso (narrativo y descriptivo). La autora también se propuso verificar la verosimilitud de los índices aplicados en su estudio, para medir la complejidad sintáctica de textos escritos en lengua española. La muestra quedó conformada por treinta alumnos, pertenecientes a segundo, cuarto y sexto grado; hubo 10 alumnos de cada curso, 5 hombres y 5 mujeres en cada caso. Todos pertenecían a un colegio privado, cuyos alumnos pertenecían a familias de estrato socioeconómico medio y medio alto, llamado San José.

El corpus de este estudio se recolectó en febrero de 1988, época en la cual se solicitó a los participantes realizar una composición narrativa sobre sus vacaciones de navidad y una

composición descriptiva sobre su juego favorito. El análisis del corpus se hizo con base en la teoría de Hunt (como se citó en Rodríguez, 1991), quien propuso dos tipos de índices sintácticos para medir la complejidad oracional: índices primarios e índices secundarios. No obstante, en este estudio, se aplicaron únicamente los índices primarios: 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas, y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22). Por consiguiente, para iniciar con la segmentación del corpus, la investigadora estableció primero el promedio de palabras escritas en cada composición, de modo que según los resultados, el número de palabras fue incrementando según el grado cursado: segundo grado, 745 palabras en textos narrativos, en contraste con 598 palabras en textos descriptivos; cuarto grado, 1,376 palabras en textos narrativos, en contraste con 1,072 palabras en textos descriptivos; y sexto grado, 2,026 palabras en textos narrativos, en contraste con 1,571 en textos descriptivos.

De acuerdo con los resultados obtenidos a través del análisis de varianza, Rodríguez (1991) concluyó que las variables sexo y modo de discurso no fueron significativas en esta investigación; dado que no presentaron una relación asociativa con los índices de madurez sintáctica, como se esperaba, excepto en algunos casos, aunque con diferencias poco significativas. Por ejemplo, con la variable sexo, la relación se produjo sólo con el índice sintáctico promedio de longitud de las cláusulas; en este caso, las mujeres utilizaron más palabras por cláusulas, 6.12, en contraste con 5.49. Con la variable modo de discurso, se produjo una relación intrascendente con los índices sintácticos promedio de longitud de las cláusulas y promedio de cláusulas por unidad terminal; en este último, los alumnos produjeron más cláusulas en las composiciones descriptivas, 1.64, en contraste con 1.31.

La variable grado de escolaridad sí fue significativa en este estudio, ya que los índices de madurez sintáctica de los escolares se fueron incrementando, a medida que los alumnos avanzaban de grado. El crecimiento fue progresivo para los tres cursos en los índices: promedio de longitud de la unidad terminal (segundo, 6.873; cuarto, 8.544; sexto, 10.56) y promedio de cláusulas por unidad terminal (segundo, 1.29; cuarto, 1.45; sexto, 1.66). Por tanto, Rodríguez (1991) estableció que mientras más alto el curso escolar de un alumno, mayor será el promedio de estos índices en sus composiciones. Respecto a la fiabilidad de los índices aplicados, la autora comparó sus resultados con los del estudio de Véliz realizado en Chile (como se citó en Rodríguez, 1991), lo cual le permitió observar que los alumnos puertorriqueños presentaron índices promedio más altos

que los de los alumnos chilenos, aunque con diferencias poco significativas. En consecuencia, planteó los índices de Hunt como medidas válidas para la complejidad sintáctica de textos, y sugirió incluir nuevas variables y nuevos tipos de discurso en futuras investigaciones; además de análisis cualitativos.

Otro estudio realizado en Puerto Rico fue el de Espinet (1996), quien estableció el índice de complejidad sintáctica de textos narrativos, descriptivos, expositivos y argumentativos producidos por escritores profesionales puertorriqueños, con el fin de comprobar si existía alguna relación asociativa entre estos índices y las variables modo de discurso y sexo; de igual manera, esta autora se propuso comprobar la fiabilidad y la validez de los índices aplicados en este trabajo, el cual se basó en su tesis doctoral (como se citó en Espinet, 1996). En esta investigación, la muestra quedó integrada por cuarenta composiciones publicadas entre 1983 y 1991, se eligieron 10 textos diferentes de cada modalidad, y se tuvo en cuenta que de cada modo discursivo, 5 textos fueran escritos por hombres y 5, por mujeres. De la muestra total, sólo nueve textos pertenecían a prosa periodística, y los demás, a prosa literaria: cuentos, novelas, ensayos y obras de teatro.

Para el análisis de los textos seleccionados, la autora se apoyó en los presupuestos de Hunt (como se citó en Espinet, 1996), aplicando únicamente los índices primarios. Por consiguiente, la segmentación del corpus inició con el conteo de palabras escritas en cada composición, para obtener el promedio de longitud de la unidad terminal y de cláusulas, así como el promedio de cláusulas por unidad terminal. Para determinar si los resultados obtenidos arrojaban diferencias estadísticamente significativas, los datos se sometieron a un análisis de varianza. Según los resultados de este estudio, los textos de modo descriptivo fueron los que presentaron índices más altos de complejidad sintáctica en los tres índices primarios analizados: 21.01, 8.99 y 2.39 (véase tabla 18). Sin embargo, no fueron estadísticamente significativos ni el índice promedio de longitud de la unidad terminal, ni el índice promedio de cláusulas por unidad terminal. Por tanto, Espinet (1996) estableció que la relación asociativa solamente se produjo con el índice promedio de longitud de las cláusulas, el cual sí presentó resultados significativos.

Tabla 18.  
Resultados generales de los índices primarios, según la variable modo de discurso

ÍNDICES	MODOS			
	NARRAT.	DESCRIPT.	EXPOSIT.	ARGUMENTANT.
PAL/UT	17.46	21.01	20.02	17.22
PAL/CL	8.02	8.99	8.46	7.42
CL/UT	2.13	2.39	2.36	2.34

Nota: PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas y CL/UT= promedio de cláusulas por unidad terminal.

Fuente: Espinet (2007, p. 45).

En cuanto a los resultados de la variable sexo, la autora estableció que los textos escritos por informantes hombres presentaron promedios más altos de complejidad sintáctica que los escritos por las informantes mujeres en los tres índices aplicados. El incremento en favor del sexo masculino se observó en las cuatro modalidades discursivas (véase tabla 19), excepto en los siguientes casos en los que las mujeres presentaron resultados más altos: promedio de longitud de la unidad terminal en textos argumentativos (18.56, en contraste con 15.87); promedio de longitud de las cláusulas en textos descriptivos (9.18, en contraste con 8.80); y promedio de cláusulas por unidad terminal en textos narrativos (2.55, en contraste con 1.70), así como en textos argumentativos (2.62, en contraste con 2.06).

Tabla 19.

Resultados generales de los índices primarios, según las variables sexo y modo de discurso

ÍNDICES	MODOS							
	NARRAT.		DESCRIPT.		EXPOSIT.		ARGUMENT.	
	F	M	F	M	F	M	F	M
PAL/UT	14.74	20.19	14.72	27.30	19.04	21.00	18.56	15.87
PAL/CT	7.37	8.67	9.18	8.80	8.44	8.47	7.05	7.78
CL/UT	2.55	1.70	1.59	3.20	2.22	2.50	2.62	2.06

Nota: PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas, CL/UT= promedio de cláusulas por unidad terminal, F= femenino y M= masculino.

Fuente: Espinet (2007, p. 45).

Según el análisis estadístico, Espinet (1996) concluyó que la variable sexo no arrojó resultados significativos en la modalidad expositiva en ninguno de los tres índices primarios. Sin embargo, sí hubo relación asociativa con la variable sexo en los siguientes índices: promedio de longitud de la unidad terminal en textos descriptivos (27.30, en contraste con 14.72) y en textos narrativos (20.19, en contraste con 14.74); promedio de longitud de las cláusulas en textos argumentativos (7.78, en contraste con 7.05); y promedio de cláusulas por unidad terminal en textos descriptivos (3.20, en contraste con 1.59) y en textos narrativos (1.70, en contraste con 2.55).

En conclusión, la autora estableció que la relación asociativa entre el índice de complejidad sintáctica de textos producidos por escritores puertorriqueños y las variables analizadas (modo de discurso y sexo) se produjo sólo en algunos casos; pues aunque en algunos índices, hubo

diferencias, no todas fueron significativas al realizar el análisis estadístico. Respecto al modo discursivo, destacó que en los tres índices analizados, el nivel de complejidad fue más alto para las composiciones descriptivas que para las composiciones narrativas, las expositivas y las argumentativas. En el caso de la variable sexo, las composiciones de los hombres presentaron índices más altos de complejidad sintáctica que las de las mujeres. En cuanto a la fiabilidad de los resultados, la autora concluyó que los índices propuestos por Hunt son medidas válidas y estables cuando se aplican a composiciones de adultos cultos, por lo que sugiere seguir aplicando esta teoría en futuras investigaciones relacionadas con el estudio cuantitativo de la lengua escrita.

En Costa Rica, Sánchez (1998) centró su investigación en niños de educación preescolar originarios del Valle Intermontano Central. Con este estudio, el autor buscó establecer el grado de complejidad sintáctica de producciones orales realizadas por preescolares vallecentrales, tratando de comprobar si existía alguna relación asociativa entre estos índices y las variables sexo y estrato social; esta última se asignó con base en el tipo de institución a la que pertenecían los alumnos (marginal, urbana y privada). La muestra estuvo conformada por 30 preescolares: 7 niñas y 3 hombres de una institución pública marginal, 6 niñas y 4 niños de una institución pública urbana, y 4 niñas y 6 niños de una institución privada; 17 preescolares del sexo femenino y 13 preescolares del sexo masculino, en total.

El corpus final incluyó 43 textos narrativos, obtenidos a través de conversaciones orales dirigidas, las cuales fueron grabadas y transcritas por estudiantes del curso denominado: Artes del Lenguaje I para preescolar. Posteriormente, el investigador procedió con la segmentación y el análisis estadístico de los textos, de acuerdo con los tres índices primarios propuestos por Hunt (como se citó en Sánchez, 1998): 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22). Cabe mencionar que los datos fueron analizados con un programa cuantificador de estructuras sintácticas, creado por un funcionario del Centro de Informática de Costa Rica, Franklin Ureña Camacho.

Los resultados de la investigación de Sánchez (1998), respecto a la variable sexo, mostraron promedios más altos por parte de los hombres en los índices promedio de longitud de la unidad terminal (6.39, en contraste con 6.34) y promedio de cláusulas por unidad terminal (1.87, en contraste con 1.82); por tanto, en el promedio de longitud de las cláusulas, el promedio más alto fue el de las mujeres. Sin embargo, en ninguno de los tres casos, hubo diferencias significativas.

En cuanto a la variable estrato social, el investigador observó que los niños procedentes de instituciones marginales obtuvieron promedios más bajos que los de escuelas públicas urbanas y los de escuelas privadas. Esta tendencia se presentó, al igual que la variable anterior, en los índices promedio de longitud de la unidad terminal (marginal 6.01, urbana 6.049, privada 6.57) y promedio de cláusulas por unidad terminal (marginal 1.63, privada 1.87, urbana 2.06), excepto en el promedio de longitud de las cláusulas, cuyos índices más altos fueron los de alumnos del grupo marginal. Este promedio fue el único que arrojó resultados significativos.

Dadas las consideraciones anteriores, Sánchez (1998) concluyó que ninguna de las dos variables establecidas fue significativa en su investigación, puesto que no presentaron una relación asociativa con los índices de madurez sintáctica utilizados. A pesar de los resultados poco significativos, el investigador consideró que la variable estrato social influyó más en el desarrollo sintáctico, que el sexo; pues al contrastar los datos considerando solamente dos grupos sociales, público (urbano y marginal) y privado, este último fue el que presentó los resultados más altos en los tres índices primarios (6.57, 3.52 y 1.87, en contraste con 6.23, 3.41 y 1.83). Para finalizar, Sánchez (1998) sugirió combinar los estudios cuantitativos con investigaciones cualitativas, con el fin de conocer la competencia comunicativa de los educandos, y mejorarla.

## **2.2 Estudios sobre complejidad sintáctica realizados en Chile**

Véliz, Muñoz y Echeverría (1985) expusieron los resultados de su proyecto de investigación financiado en el año de 1982 por la Universidad de Concepción. En ese estudio, aplicaron el método de combinación de oraciones, el cual establece “que la producción de unidades T significativamente más largas puede ser enseñada y aprendida en forma sistemática mediante ejercicios que consisten en generar oraciones complejas a partir de series de oraciones simples que se presentan como estímulo” (Véliz et al., 1985, p. 109). El objetivo principal de los autores fue demostrar la eficacia de este método para incrementar la madurez sintáctica, y mejorar la calidad general de la composición escrita; pues se consideró que la metodología tradicional utilizada para la enseñanza del español no había aportado resultados significativos en esa área de la lingüística.

Para conformar la muestra de este estudio, los autores seleccionaron 57 alumnos inscritos en la clase de Gramática y Redacción Española, pertenecientes al ciclo escolar de 1983, los cuales cursaban segundo año de Pedagogía. Véliz et al. (1985) segmentaron el grupo en dos partes: el grupo control, integrado por 29 alumnos, y el grupo experimental, conformado por 28 alumnos. Este último incluía a los estudiantes que aplicaron el método de combinación de oraciones durante el curso, y el otro grupo se refería a los que trabajaron con el método tradicional; cuyo objetivo era familiarizar al alumno con el dominio de una estructura sintáctica determinada. Según la hipótesis planteada, el grupo experimental obtendría índices más altos de complejidad sintáctica, y un progreso significativo en la calidad de sus composiciones, a diferencia del grupo tradicional.

La recolección del corpus estuvo a cargo de dos de los investigadores, quienes diseñaron la prueba pretest y la postest, las cuales consistían en solicitar una composición al iniciar el curso (pretest), y otra, al finalizar (postest). Los alumnos debían desarrollar un texto argumentativo, habiendo elegido alguno de los temas previamente establecidos por los autores de la investigación, tema A y tema B. Si escribían sobre el tema A en el pretest, desarrollaban el B en el postest, y viceversa; de manera que todos los participantes pudieran desarrollar ambos temas. Para medir la madurez sintáctica de los datos obtenidos, Véliz et al. (1985) tomaron como referencia las tres medidas principales establecidas por Kellogg Hunt: 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22). Para evaluar la calidad general de la composición escrita, los autores establecieron dos métodos: 1) un juicio holístico o impresionista, y 2) una evaluación analítica. Ambos métodos consistían en asignar puntaje a los textos; sin embargo, el primero se basaba en apreciaciones generales, y el segundo, en criterios previamente establecidos.

Según los resultados obtenidos, Véliz et al. (1985) establecieron que el grupo experimental presentó incrementos significativos en el promedio de longitud de la unidad terminal (2.521, en contraste con 0.37) y en el promedio de cláusulas por unidad terminal (2.3, en contraste con 1.14), por lo que se demuestra la eficacia de este método para incrementar la madurez sintáctica en los alumnos. Aunque el promedio de longitud de las cláusulas presentó un incremento (grupo experimental, 0.367 y grupo control, 0.83), los resultados fueron poco significativos. Ambos grupos mejoraron significativamente la calidad de sus composiciones; no obstante, según los evaluadores, el grupo experimental presentó puntajes más elevados, en contraste con los del grupo control: 2.85, en contraste con 2.64. Por tanto, los autores de este estudio concluyeron que el

desarrollo de la madurez sintáctica de estudiantes universitarios, cuya lengua materna es el español, puede ser intervenido y que tanto el método de combinación de oraciones, como los métodos propuestos por Hunt, son eficaces para evaluar e incrementar la madurez sintáctica de estudiantes universitarios.

Tres años más tarde y aplicando nuevamente la metodología de Kellogg Hunt; Véliz (1988) realizó otro estudio en la ciudad de Concepción, en el cual determinó el índice de madurez sintáctica de cincuenta alumnos, pertenecientes a cuarto, sexto y octavo grado de educación básica, y a segundo y cuarto grado de educación media. Su objetivo principal fue conocer el desarrollo sintáctico de la composición escrita de los estudiantes, de acuerdo con su grado de escolaridad. Asimismo, se propuso verificar que los índices aplicados fueran medidas válidas y confiables para medir la complejidad sintáctica de textos escritos en lengua española. La muestra de esta investigación quedó integrada por estudiantes de un centro educativo privado, pertenecientes al estrato socioeconómico medio-alto; la autora seleccionó 10 alumnos de cada grado escolar con un rango de edad entre nueve años, nueve meses, y diecisiete años, dos meses de edad.

Para obtener el corpus, Véliz (1988) solicitó a los participantes que escribieran dos composiciones narrativas referentes a dos temas previamente asignados: 1) contar la trama de la película “El extraterrestre”, y 2) contar una historia, según su imaginación, sobre las aventuras de un terrestre en otros planetas. Para analizar el corpus obtenido, de acuerdo con la fórmula de Hunt, la autora dividió primero cada texto en unidades terminales y en cláusulas; posteriormente calculó el promedio de longitud de la unidad terminal, el promedio de longitud de las cláusulas y el promedio de cláusulas por unidad terminal. Por último, aplicó un análisis de varianza para establecer si las medias de cada grupo arrojaban resultados estadísticamente significativos, y para probar la fiabilidad de los índices, realizó el cálculo de los coeficientes de estabilidad establecido por Brown (como se citó en Véliz, 1988).

De acuerdo con los resultados de este estudio, todos los alumnos incrementaron su complejidad sintáctica conforme avanzaron de grado escolar; como consecuencia, Véliz (1988) estableció que cuando los alumnos desarrollan su madurez sintáctica, “tienden a escribir unidades-t y cláusulas más largas y a incrustar, en una proporción mayor, cláusulas en el interior de las unidades-t...” (Veliz, 1988, p. 115). Por tanto, se validó la hipótesis inicial planteada respecto a la variable grado escolar; pues el índice medio de complejidad sintáctica fue significativamente más alto en los alumnos de cursos superiores, que en los de cursos inferiores. En cuanto a la validez de

los índices de madurez sintáctica aplicados en este estudio, la autora sostuvo que hubo una alta correlación entre el promedio de longitud de unidad terminal y su complejidad sintáctica, por lo cual concluyó que los índices propuestos por Hunt sí fueron instrumentos válidos y confiables, al aplicarse en la lengua española.

Véliz realizó un nuevo estudio en 1999, el cual ya no se centró en estudiantes, sino en adultos lingüísticamente competentes: escritores, periodistas, académicos, entre otros. Con este estudio, Véliz (1999) buscó establecer el grado de complejidad sintáctica de textos argumentativos, narrativos y descriptivos escritos por autores contemporáneos, tratando de comprobar si existía alguna correlación entre estos índices y la variable modo de discurso; asimismo, se propuso identificar los rasgos sintácticos más relevantes que permitieran caracterizar cada tipo de texto. Además, comparó sus resultados con los obtenidos en su estudio anterior (Véliz et al., 1991), de manera que pudiera contrastar la complejidad sintáctica de discursos producidos por adultos, con la de discursos producidos por escolares.

El corpus quedó conformado por sesenta textos: veinte narrativos, veinte argumentativos y veinte descriptivos. Estos variaban entre cuentos, novelas, ensayos, reportajes, crónicas periodísticas, editoriales, artículos de revistas, manuales de estudio, entre otros, e incluían entre trescientas y trescientas cincuenta palabras. Para el análisis de los textos seleccionados, la autora se apoyó nuevamente en los presupuestos de Kellogg Hunt (como se citó en Véliz et al., 1999) para medir la complejidad oracional: índices primarios e índices secundarios. Primero segmentó los textos en unidades terminales, posteriormente identificó en cada unidad la cláusula subordinante, así como las posibles cláusulas subordinadas adjetivas, subordinadas sustantivas y subordinadas adverbiales.

De acuerdo con los resultados obtenidos a través del análisis de varianza, Véliz (1999) concluyó que el modo de discurso sí incide significativamente en la complejidad sintáctica de las composiciones escritas. Así pues, los textos escritos en modo argumentativo presentaron índices más altos de complejidad, debido a que en este tipo de composiciones, los autores incluyeron cláusulas subordinadas con mayor frecuencia. Por el contrario, los textos narrativos presentaron los índices más bajos de complejidad, puesto que en este modo de discurso, hubo estructuras oracionales menos complejas que en las otras dos modalidades. De manera que, los textos descriptivos quedaron en un nivel medio de complejidad.

En cuanto a los rasgos sintácticos más relevantes de cada modo discursivo, Véliz (1999) estableció que para el modo argumentativo, el rasgo más relevante fue la incidencia en la incrustación de cláusulas subordinadas sustantivas y cláusulas subordinadas adjetivas. Para el modo narrativo, lo fue el hecho de usar cláusulas subordinadas sustantivas, subordinadas adjetivas y subordinadas adverbiales de manera proporcional, y para el modo descriptivo, fue sobresaliente el menor uso de cláusulas subordinadas adjetivas, un mayor uso de cláusulas subordinadas sustantivas, así como la mayor longitud que alcanzan como promedio sus cláusulas. Para concluir, después de contrastar los resultados de los adultos con los de los escolares, la autora demostró que los textos producidos por escritores lingüísticamente competentes presentaron una sintaxis más compleja, que los producidos por estudiantes de cuarto grado de educación media. Además, que en ambos estudios, la variable modo de discurso influyó significativamente en el grado de complejidad sintáctica de las composiciones escritas.

En el año 2007, Coloma, Fernández y Peñaloza realizaron un estudio con la finalidad de conocer el desarrollo sintáctico de oraciones subordinadas producidas oralmente por niños chilenos mayores de siete años, y determinar si existía correlación alguna entre la variable edad y los índices obtenidos. La muestra de esta investigación estuvo conformada por 59 alumnos con nivel socioeconómico medio-alto, procedentes de un colegio particular. Del total de la muestra, 30 alumnos cursaban el tercer grado de educación básica: 17 mujeres y 13 hombres, con una edad promedio de ocho años, cinco meses; y 29 alumnos cursaban quinto grado de educación básica: 18 mujeres y 11 hombres, con una edad promedio de diez años, nueve meses (véase tabla 20).

Tabla 20.  
Muestra total según los cursos y la edad promedio

Grado	Mujeres	Hombres	Total	Edad promedio
3°	17	13	30	8 años 5 meses
5°	18	11	29	10 años 9 meses

Fuente: Elaboración propia con base en Coloma et al. (2007).

Para la recolección del corpus, se citó a los alumnos participantes de manera individual, en un salón del colegio seleccionado, y se les relataron tres cuentos infantiles previamente seleccionados por Coloma et al. (2007): 1) El tiranosaurio, 2) La ballena gigante y 3) Robinson y Viernes. Al término de cada cuento, el investigador que se encontraba a cargo solicitó a cada alumno que contara la misma historia para poder grabarla; por consiguiente, cada narración fue

transcrita de manera íntegra, incluyéndose todas las pausas realizadas durante la grabación. Cabe mencionar que los autores de este estudio decidieron aplicar el discurso narrativo, pues según estudios previos realizados por Gutiérrez-Clellen y Pavez et al. (como se citó en Coloma et al., 2007), el uso de la sintaxis compleja y la producción de oraciones extensas se relaciona con el desempeño de este tipo de discurso.

Para el análisis del corpus recolectado, los investigadores no aplicaron ninguna teoría en específico, sin embargo, establecieron los términos que tuvieron en cuenta para analizar cada composición: oración, oración simple y oración compleja; en esta última, consideraron cláusulas subordinadas adjetivas, cláusulas subordinadas sustantivas y cláusulas subordinadas adverbiales. Por tanto, para Coloma et al. (2007):

Se entendió por oración una forma lingüística constituida por un núcleo predicativo autónomo.

Se entendió por oración simple la que presenta un núcleo predicativo (verbal o no verbal) sin presencia de otros núcleos subordinados.

Por su parte, la oración compleja es la que contiene un núcleo predicativo principal y uno o más núcleos predicativos subordinados.

La locución “hasta que” no se consideró subordinante, sino un marcador pragmático porque aparecía en oraciones simples con un carácter enfático que estaba marcado por pausa.

La expresión *para que* fue analizada como locución subjuntiva de cláusula adverbial (valor semántico de finalidad) y no como subordinante de cláusula sustantiva. (p. 37)

De acuerdo con lo anterior, para el análisis estadístico, los autores primeramente establecieron el número total de oraciones simples y complejas que constituía cada corpus; posteriormente, contaron el número total de cláusulas, y según su clasificación, determinaron el número existente de cada tipo de cláusula. Por último, para contrastar los resultados entre los niños de ocho y los de diez años, aplicaron la prueba *t* de *Student*<sup>9</sup>, y para comprobar si existía correlación entre la sintaxis compleja y la variable edad, aplicaron la prueba *r* de *Pearson*.

Según los resultados obtenidos en esta investigación, el total de oraciones producidas en ambos grupos etarios tuvo un promedio semejante: 44.2 (ocho años) y 45.7 (diez años). Los alumnos de ocho años se distinguieron por expresar más oraciones simples que los alumnos de diez años (30.56, en contraste con 28.27). Por el contrario, los alumnos de diez años produjeron más oraciones subordinadas que los alumnos de ocho años (17.4, en contraste con 13.6), y presentaron incrementos significativos en el total de cláusulas (25.4, en contraste con 17.5), en el total de

---

<sup>9</sup> Esta prueba se utiliza para determinar si hay una diferencia significativa entre las medias de dos grupos.

cláusulas subordinadas adjetivas (6.6, en contraste con 5.0) y en el total de cláusulas subordinadas sustantivas (13.0, en contraste con 7.6). En el caso de las cláusulas subordinadas adjetivas, hubo mayor producción, pero con diferencias estadísticamente poco significativas (5.7, en contraste con 4.8). Por tanto, según los resultados obtenidos a través de la prueba *r* de *Pearson*, Coloma et al. (2007) sostuvieron que los alumnos sí incrementaron su complejidad sintáctica de acuerdo con su edad, y que sí hubo correlación con los índices evaluados (véase tabla 21), excepto en el caso de las cláusulas subordinadas adverbiales.

Tabla 21.  
Correlaciones entre edad y sintaxis compleja

Total de oraciones complejas	r = 0.36
Total de cláusulas	r = 0.45
Total de cláusulas adjetivas	r = 0.34
Total de cláusulas sustantivas	r = 0.44

Fuente: Coloma et al. (2007, p. 43).

P = 0.01

Cabe destacar que Coloma et al. (2007) contrastaron sus resultados con investigaciones previas como las de Gili Gaya (1972), Véliz et al. (1985), Véliz (1988) y Gutiérrez-Clellen y Hofstetter (1994) (como se citó en Coloma et al., 2007). Sus resultados coincidieron con los de esas investigaciones en que el incremento de las oraciones complejas y el de las cláusulas subordinadas adjetivas en las composiciones de los niños surge a partir de los siete años. Todos estos autores plantearon la necesidad de elaborar futuras investigaciones en las que se incluyan otras variables enfocadas en las cláusulas adverbiales, debido a que estas construcciones sintácticas no presentaron incrementos significativos, tanto en el estudio de Coloma et al. (2007) como en el de Hofstetter.

Por su parte, Balboa, Crespo y Rivadeneira (2012) analizaron la relación existente entre los índices de complejidad sintáctica de narraciones orales producidas por alumnos de Valparaíso y la variable independiente de estrato socioeconómico. Además, analizaron sus resultados con base en la teoría del déficit de Bernstein<sup>10</sup>, y con base en la sociolingüística variacionista<sup>11</sup> propuesta por Labov (1965, 1966, 1972), Weinreich et al. (1968), Labov y Sankoff (1985), entre otros... (como

<sup>10</sup> "... sostiene que el fracaso escolar de los sujetos de estratos más bajos obedece a una suerte de déficit lingüístico, en comparación con aquellos sujetos pertenecientes a los estratos medios y altos" (Balboa et al., 2012, p. 147).

<sup>11</sup> "... busca indagar las diferencias lingüísticas que caracterizan a los diferentes grupos sociales" (Balboa et al., 2012, p. 147).

se citó en Balboa et al., 2012). La muestra quedó conformada por sesenta estudiantes que cursaban el segundo año de educación media, y tenían un promedio de quince años de edad (véase tabla 22).

Tabla 22.

Muestra de los estratos de la población escolar del Ciclo Medio de Educación General Básica

Variable	Número	Hombres	Mujeres	%
B: Clase baja-Pública	30	15	15	50
E: Clase alta-Privada	30	15	15	50
Total	60	30	30	100

Fuente: Elaboración propia con base en Balboa et al. (2012).

Balboa et al. (2012) seleccionaron 30 alumnos de colegios públicos y 30 alumnos de colegios privados, 15 hombres y 15 mujeres en cada caso; el 50% del total de la muestra correspondía a la clase baja (grupo B), y el 50%, a la clase alta (grupo E). Esta clasificación se asignó con base en el grado de escolaridad de los padres y en el estrato socioeconómico (ingreso familiar), el cual se ordenó de menor a mayor: A, B, C, D y E. Cabe mencionar que todos los alumnos seleccionados tenían un nivel académico homogéneo, puesto que sólo se incluyó a los que obtuvieron los puntajes más altos en la prueba denominada Sistema de Medición de la Calidad de la Educación, aplicada por el Sistema Nacional de Evaluación de resultados de aprendizaje del Ministerio de Educación de Chile.

Para la recolección de datos, Balboa et al. (2012) solicitaron a los participantes de manera individual que visualizaran el corto animado sin diálogos titulado: *Fly Guy*. Posteriormente, le solicitaron a cada alumno una narración oral de la historia, la cual grabaron para poder transcribirla y analizarla. Analizaron el corpus con base en la propuesta de las unidades T de Hunt y la adaptación de Véliz, y para el análisis estadístico, aplicaron la prueba de *Kruskal-Wallis* (como se citó en Balboa et al., 2012). Por tanto, primero contaron el total de palabras producidas en cada narración, seguidamente segmentaron las producciones en unidades terminales, posteriormente, identificaron el total de cláusulas de cada texto, y por último, determinaron los índices primarios y los índices secundarios o cláusales (véase tabla 9 y 10, p. 22). Primarios: 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas y 3) promedio de cláusulas por unidad terminal. Secundarios: 4) promedio de cláusulas adjetivas por unidad terminal, 5) promedio de cláusulas sustantivas por unidad terminal y 6) promedio de cláusulas adverbiales por unidad terminal.

Balboa et al. (2012) concluyeron que existen diferencias poco significativas entre los índices primarios de complejidad sintáctica de narraciones orales producidas por alumnos procedentes de estrato socioeconómico bajo, y los de narraciones orales producidas por alumnos de estrato socioeconómico alto; pues aunque los alumnos de estrato socioeconómico alto produjeron narraciones más extensas que los alumnos de estrato bajo, con una diferencia de 52% a su favor, fueron las producciones de estos últimos las que mostraron promedios más altos en los tres índices primarios (véase tabla 23). Sin embargo, al aplicar la prueba de *Kruskal-Wallis*, esas diferencias no fueron estadísticamente significativas: promedio de longitud de la unidad terminal ( $\chi^2=0.6733$ ), promedio de longitud de las cláusulas ( $\chi^2=0.6733$ ) y promedio de cláusulas por unidad terminal ( $\chi^2=1.012$  valor), puesto que esos valores son mayores a  $p<0.05$ .

Tabla 23.  
Medidas descriptivas de índices primarios según estrato socioeconómico

Índice	Estrato bajo	Estrato alto
1) PAL/UT	8,7	8,6
2) PAL/CL	7,85	6,24
3) CL/UT	1,53	1,37

Fuente: Elaboración propia con base en Balboa et al. (2012).

En cuanto a los índices secundarios, una vez más, fueron los alumnos de estrato socioeconómico bajo, quienes obtuvieron los promedios más altos en el promedio de cláusulas subordinadas adjetivas por unidad terminal (0.2, en contraste con 0.1); pues en el promedio de cláusulas subordinadas sustantivas por unidad terminal (0.2 en ambos grupos) y en el promedio de cláusulas subordinadas adverbiales por unidad terminal (estrato bajo, 0.1 y estrato alto, 0.09), la diferencia fue mínima entre ambos grupos. No obstante, estas diferencias tampoco fueron estadísticamente significativas al aplicar la prueba de medias, pues los resultados fueron mayor a  $p<0.05$ : cláusulas subordinadas adjetivas  $\chi^2=0.1778$ , cláusulas subordinadas sustantivas  $\chi^2=0.1778$  y cláusulas subordinadas adverbiales  $\chi^2=1.6837$ . Por otra parte, el género femenino del estrato socioeconómico bajo obtuvo índices más altos de complejidad sintáctica, tanto en los índices primarios como en los secundarios; a diferencia de los alumnos de estrato alto, cuyos resultados fueron homogéneos entre hombres y mujeres.

De acuerdo con los resultados anteriores, Balboa et al. (2012) establecieron que no se corroboró la teoría de Bernstein seguida en este estudio, la cual sostiene: “que los sujetos de estrato bajo, al utilizar código restringido, manifiestan un habla cuya organización sintáctica será limitada

y que, por lo tanto, sería inferior a la de los estratos altos” (como se citó en Balboa et al., 2012). Por el contrario, los resultados de este estudio se inclinaron más por la propuesta variacionista, puesto que coincidieron con dos investigaciones: 1) la de Labov (como se citó en Balboa et al., 2012), quien establece que “los sujetos de estratos altos y medios utilizan más vocabulario y producen en general textos más extensos y ampulosos, pero, ... no son más complejos que los producidos por sus pares de estratos menos favorecidos” (p. 161). Y 2) la de Véliz et al. (como se citó en Balboa et al., 2012), quienes “observan mayores valores en los textos escritos elaborados por sujetos de estrato socioeconómico bajo, por sobre los valores obtenidos de los textos elaborados por sujetos del estrato socioeconómico alto... señalan que las diferencias encontradas no son significativas estadísticamente” (p. 161).

Para concluir, los autores propusieron corroborar sus resultados con la elaboración de futuros estudios sobre la misma línea de investigación; pero modificando las variables ya existentes, ya sea variando el nivel académico de los hablantes y el modo de discurso, o en su caso, añadiendo nuevas variables, para verificar si estos cambios influirían en los índices de complejidad sintáctica producidos por los alumnos de diferentes clases sociales. También propusieron que en los corpus se incluyan grabaciones de habla espontánea e informal, puesto que su investigación optó por el uso de un lenguaje formal y académico.

Dos años después, en su tesis de maestría, Hugo (2014) midió la complejidad sintáctica de composiciones narrativas y explicativas producidas por estudiantes de la Región Metropolitana de Santiago de Chile, con el fin de analizar la relación entre los resultados obtenidos y las variables edad/nivel escolar, género discursivo y grupo socioeconómico. Este estudio analizó el desarrollo de la complejidad sintáctica desde dos perspectivas: la funcional-discursiva (paquetes clausulares) y la generativista (índices de Hunt); por lo cual, la autora también se propuso identificar las similitudes y las diferencias entre los resultados obtenidos con ambos métodos.

La muestra de esta investigación incluyó 120 alumnos, con un rango de edad entre diez y dieciséis años, procedentes de un colegio público y un colegio particular. Se seleccionaron estudiantes de séptimo básico, primero de secundaria y tercero de secundaria, y con base en la clasificación establecida por el Ministerio de Educación para los colegios participantes (como se citó en Hugo, 2014), se distribuyó a los alumnos en dos grupos socioeconómicos: 60 en el grupo medio-alto y 60 en el grupo alto; 10 hombres y 10 mujeres por cada nivel escolar y por cada grupo socioeconómico (véase tabla 24).

Tabla 24.  
Cantidad de participantes según nivel, grupo social y sexo

Nivel	Grupo Medio		Grupo Alto	
	Hombres	Mujeres	Hombres	Mujeres
SEP	10	10	10	10
PRI	10	10	10	10
TER	10	10	10	10

Fuente: Hugo (2014, p. 26).

Se obtuvo el corpus, previo diseño de un video de 3 minutos y 40 segundos (sin voces de fondo) sobre “Internet y las redes sociales”<sup>12</sup>, con el cual se estimuló la redacción de discursos narrativos y explicativos, puesto que este estudio buscó la producción de discursos mono gestionados. El estímulo visual creado mostraba imágenes de las redes sociales y otros recursos de internet, como por ejemplo: facebook, twitter, youtube, skype, entre otros. También presentaba frases sobre internet y las redes sociales, por ejemplo: “Hoy un 44.5% de la población tiene acceso a internet”, con la finalidad de que los alumnos tuvieran más información sobre el tema seleccionado.

Por consiguiente, la prueba consistió en que los alumnos observaran el video, y posteriormente escribieran sus composiciones de acuerdo con las instrucciones establecidas para cada género discursivo:

... La tarea narrativa se titulaba “Historias en la Red”:

A continuación te pedimos relatar una historia o anécdota que te haya ocurrido a ti o a algún conocido con el uso de Internet (por ejemplo, un encuentro en una red social o comunidad virtual, un conflicto con alguien, alguna actividad realizada en el colegio, la historia de cómo aprendiste a usar estos medios, la historia de una confusión musical o artística gracias a estos medios, etc.).

... La tarea explicativa se titulaba “Recursos de la Red”:

Elige uno de esos sitios recursos de Internet que te guste y conozcas bien (no importa si no aparece en el video). Una vez elegido, te pedimos escribir un artículo en que lo presentes y lo expliques (por ejemplo, qué es, para qué sirve, cómo se usa, etc.), para personas que no lo conocen y quieren aprender a usarlo. (Hugo, 2014, p. 27)

La recolección del corpus se llevó a cabo en tres sesiones de 90 minutos, cada una, durante las horas de la asignatura de Lenguaje, con diferencias de dos o tres semanas por reunión. En la primera sesión, se explicó a los alumnos el objetivo del proyecto, y se les informó sobre las

<sup>12</sup> Después de realizar una encuesta de intereses a unos 700 estudiantes, se seleccionó este tema porque resultó ser el que cumplía con el requisito de ser atractivo y accesible para los estudiantes de los tres niveles.

actividades que les serían solicitadas. Durante la segunda sesión, se les proyectó el video creado y se les solicitó que escribieran una composición narrativa; finalmente, en la tercera sesión, los alumnos visualizaron nuevamente el video, y posteriormente, se les solicitó una composición explicativa. Como cada alumno escribió dos textos, el corpus total constó de 240 composiciones, de las cuales, 120 pertenecían al modo narrativo y 120, al modo explicativo.

Para medir el grado de complejidad sintáctica de los textos obtenidos, Hugo (2014) analizó los datos con base en la perspectiva funcional-discursiva, la cual incluyó criterios cualitativos y cuantitativos para medir la sintaxis, y con base en la perspectiva generativista, basada en los índices de Hunt, previamente explicados en el capítulo 1 de este trabajo monográfico (véase tabla 9, p. 22). Desde la perspectiva funcionalista, se midió la complejidad oracional mediante la segmentación de paquetes clausulares y de cláusulas (véase tabla 25). Según Katzenberger, los paquetes clausulares se han definido como “una unidad textual mínima que típicamente consiste en varias cláusulas vinculadas por criterios sintácticos, temáticos y/o criterios discursivos, que no alcanza a abarcar las estructuras globales del discurso, pero se relaciona con estas” (como se citó en Hugo, 2014). Las cláusulas, según Berman y Slobin, se definen como “una unidad que contiene un predicado autónomo y unificado que expresa una situación única (actividad, evento o estado) (como se citó en Hugo, 2014)”. Para clasificar las relaciones interclausulares, se codificó cada cláusula con base en las cinco categorías propuestas por Nir-Sagiv y Berman (2010), y en la adaptación realizada por Meneses, Ow y Benítez (2012):

1. **Isotaxis:** cláusulas aisladas que no utilizan ningún nexo para establecer relaciones con otras; utilizan el asíndeton para hacer progresar la información. En estos tipos de construcciones se distinguen cláusulas únicas y cláusulas principales.
2. **Parataxis simétrica:** cláusulas que establecen una relación de equivalencia informativa; se utiliza el encadenamiento de cláusulas a una cláusula principal mediante asíndeton (cláusulas yuxtapuestas) o nexos de coordinación (cláusulas coordinadas). En esta categoría se consideran, asimismo, mecanismos de cohesión textual, específicamente, de mantención del referente y procedimientos gramaticales de cohesión (sujeto implícito o explícito).
3. **Parataxis asimétrica:** cláusulas que dependen parcialmente de una cláusula principal; la cláusula dependiente sólo puede ser interpretada con relación a la cláusula precedente. Se consideran dentro de este grupo cláusulas coordinadas y yuxtapuestas con verbo elíptico, cláusulas subordinadas sustantivas en función de objeto, atributo o complemento en verbos con régimen preposicional.
4. **Hipotaxis:** cláusulas subordinadas que establecen una relación de dependencia de una cláusula principal. En estas construcciones se consideran cláusulas adjetivas especificativas, las cláusulas sustantivas en función sujeto, las cláusulas adverbiales

propias (tiempo, modo y lugar), las cláusulas concesivas, consecutivas, causales, finales, condicionales y comparativas.

5. **Endotaxis:** cláusulas incrustadas y anidadas dentro de otras funcionando al interior de un grupo. En estas construcciones se consideran cláusulas adjetivas explicativas, adverbiales en aposición y cláusulas reparadoras en función parentética. (Como se citó en Hugo, 2014, p. 11)

Tabla 25.

Fórmulas para medir la complejidad sintáctica según la perspectiva funcionalista

Índice	Fórmula
Longitud de paquetes clausulares	(PAL/PC)
Longitud de la cláusula	(PAL/CL)
Índice de empaquetamiento	(CL/PC)

Nota: PAL= total de palabras, PC= total de paquetes clausulares, CL= total de cláusulas.

Fuente: Elaboración propia con base en Hugo (2014, pp. 54-59).

Por consiguiente, cuatro personas con formación lingüística se encargaron de transcribir y segmentar en cláusulas las composiciones. En primera instancia, siguieron la definición de cláusula propuesta por Berman y Slobin, correspondiente a la perspectiva funcional-discursiva, y posteriormente, segmentaron los textos con la definición de cláusula establecida por Hunt. Después, según la metodología de cada perspectiva, codificaron los textos en unidades terminales y en paquetes clausulares; en este caso, además de los criterios sintácticos, también tuvieron en cuenta los criterios temáticos y discursivos. El análisis de los datos se realizó mediante un programa denominado CLAN (*Computerized Language Analysis*), el cual permitió medir el número de palabras por cláusula para la perspectiva funcionalista y el número total de palabras de todas las composiciones, para obtener los índices primarios en el modo generativista. Se aplicó un análisis de varianza (con un nivel de significación de 0.05), para determinar si los resultados obtenidos arrojaban diferencias estadísticamente significativas. Y para comparar los resultados obtenidos entre ambos métodos, se aplicó una prueba de correlación entre las variables establecidas en este estudio y el coeficiente de *Pearson*, lo cual permitió clasificar las correlaciones como bajas, moderadas y altas. La autora consideró una correlación como baja, cuando el resultado obtenido era menor o igual a 0.30; moderada, cuando el resultado era cercano a 0.50; y alta, cuando era superior o igual a .80.

Según la perspectiva generativista, Hugo (2014) estableció los siguientes resultados. La variable edad/nivel escolar no presentó resultados significativos en los textos narrativos escritos por alumnos del grupo socioeconómico medio, con ninguno de los índices analizados. Por el

contrario, sí presentó resultados significativos en los textos narrativos escritos por alumnos del grupo socioeconómico alto: la productividad verbal (séptimo, 26.20; primero, 41.65; tercero, 57.15) y el índice longitud de la cláusula (séptimo, 6.81; primero, 7.46; tercero, 8.06) incrementaron conforme los alumnos avanzaron de curso. En el índice longitud de la unidad terminal (séptimo, 13.81; primero, 12.37; tercero, 18.73) y en el índice de subordinación (séptimo, 2.06; primero, 1.65; tercero, 2.29), los resultados mostraron un promedio alto en los alumnos de séptimo básico, este promedio disminuyó en primero de secundaria e incrementó en tercero de secundaria; aun así, estos resultados fueron significativos. En los textos explicativos, los resultados fueron significativos para ambos grupos socioeconómicos en todos los índices analizados, los alumnos de tercero escribieron composiciones más complejas que los alumnos de séptimo básico y los de primero de secundaria (véase tabla 26).

Respecto a la variable grupo socioeconómico, los textos narrativos escritos por alumnos de séptimo básico del grupo socioeconómico medio, presentaron resultados más altos que los textos escritos por alumnos del grupo socioeconómico alto en el índice longitud de cláusula y en su productividad verbal; sin embargo, estos resultados no fueron significativos. No obstante, esta variable sí presentó resultados significativos en los textos narrativos escritos por alumnos del grupo socioeconómico alto en los siguientes casos: los alumnos de primero y tercero de secundaria escribieron textos con mayor productividad verbal respecto al otro grupo socioeconómico (primero, 41.65 y tercero, 57.15, en contraste con primero, 28.55 y tercero, 40.50), los alumnos de tercero obtuvieron mayor promedio de longitud de unidad terminal (18.73, en contraste con 13.17) y el índice de subordinación fue mayor en los textos escritos por alumnos de séptimo básico y tercero de secundaria (séptimo, 2.06 y tercero, 2.29, en contraste con séptimo, 1.61 y tercero, 1.69). En los textos explicativos, los alumnos del grupo socioeconómico alto obtuvieron promedios más altos con respecto al otro grupo socioeconómico en los siguientes casos: productividad verbal (primero, 32.50 y tercero, 42.05 en contraste con primero, 19.40 y tercero, 27.05), longitud de la cláusula (tercero, 8.78, en contraste con tercero, 7.90), longitud de la unidad terminal (primero, 14.58 y tercero, 19.58, en contraste con primero, 12.57 y tercero, 14.46) y en el índice de subordinación (séptimo, 1.94 y tercero, 2.23, en contraste con séptimo, 1.52 y tercero, 1.86). El único promedio significativo para el grupo socioeconómico medio en este tipo de textos fue para los alumnos del séptimo básico en el índice longitud de la cláusula (7.70, en contraste con 6.86).

La variable género discursivo presentó resultados significativos con la productividad verbal en ambos grupos sociales, los textos narrativos fueron más extensos que los explicativos (véase tabla 26). También se observó mayor longitud de cláusula en los textos de modo explicativo, sin embargo, este índice sólo fue significativo en los textos escritos por alumnos del grupo socioeconómico alto.

Tabla 26.  
Resultados de los índices, según la perspectiva generativista

		Grupo Medio				Grupo Alto			
		Séptimo	Primero	Tercero	Total	Séptimo	Primero	Tercero	Total
PV	Narrativo	33.00	28.55	40.50	34.01	26.20	41.65	57.15	41.66
	Explicativo	18.25	19.40	27.05	21.56	25.95	32.50	42.05	33.50
L/UT	Narrativo	11.72	12.12	13.17	12.34	13.81	12.37	18.37	14.97
	Explicativo	11.65	12.57	14.46	12.90	13.28	14.58	19.58	15.82
L/CL	Narrativo	7.29	7.45	7.81	7.52	6.81	7.46	8.06	7.44
	Explicativo	7.70	7.51	7.90	7.70	6.86	7.90	8.78	7.85
C/UT	Narrativo	1.61	1.63	1.69	1.64	2.06	1.65	2.29	2.00
	Explicativo	1.52	1.69	1.86	1.69	1.94	1.87	2.23	2.02

Nota: PV= productividad verbal, L/CL= longitud de cláusula, L/UT=longitud de unidad terminal, C/UT= índice de subordinación.

Fuente: Elaboración propia con base en Hugo (2014, pp. 41-48).

De acuerdo con Hugo (2014), la perspectiva funcional-discursiva presentó los siguientes resultados. La variable edad/nivel escolar mostró resultados significativos en los textos narrativos escritos por alumnos del grupo socioeconómico medio, sólo con los índices: longitud del paquete clausular (séptimo, 30.87; primero, 23.72; tercero, 32.10) e índice de empaquetamiento (séptimo, 5.58; primero, 4.25; tercero, 5.68), en ambos casos, los alumnos de séptimo básico obtuvieron un promedio elevado, este promedio disminuyó en primero de secundaria e incrementó nuevamente en tercero de secundaria. En cambio, los textos narrativos escritos por alumnos del grupo socioeconómico alto mostraron resultados significativos con todos los índices analizados: la productividad verbal y la longitud de cláusula incrementaron conforme los alumnos avanzaron de curso, y la longitud de los paquetes clausulares y el índice de empaquetamiento mostraron mayor promedio en los textos escritos por alumnos de séptimo básico que en los escritos por alumnos de primero de secundaria, estos a su vez fueron más bajos que los escritos por alumnos de tercero de secundaria, los cuales mostraron mayor promedio que los escritos por alumnos de primero de secundaria (véase tabla 27). En los textos explicativos, los resultados fueron significativos para ambos grupos socioeconómicos en todos los índices analizados, los alumnos de tercero de

secundaria escribieron composiciones más complejas que los alumnos de séptimo básico y los de primero de secundaria; excepto en el índice longitud de cláusula de textos escritos por alumnos del grupo socioeconómico medio, pues aunque se observó un incremento conforme los alumnos cambiaron de curso, ese incremento no fue significativo (véase tabla 27).

La variable grupo socioeconómico no mostró resultados significativos en los textos narrativos escritos por alumnos del grupo socioeconómico medio, excepto en la productividad verbal, donde los alumnos de séptimo básico produjeron textos más extensos que los del otro grupo socioeconómico (43.10, en contraste con 33.50). Los textos narrativos escritos por alumnos del grupo socioeconómico alto sí presentaron resultados significativos en los siguientes índices: productividad verbal (primero, 54.35 y tercero, 75.30, en contraste con primero, 36.10 y tercero, 54.35), longitud del paquete clausular (primero, 25.60 y tercero, 43.47, en contraste con primero, 23.72 y tercero, 32.10) e índice de empaquetamiento (primero, 4.58 y tercero, 7.32, en contraste con primero, 4.25 y tercero, 5.68), en los tres casos, los alumnos de primero y de tercero de secundaria obtuvieron promedios más elevados que los alumnos de séptimo básico respecto del otro grupo. En los textos explicativos, fueron los alumnos del grupo socioeconómico alto quienes obtuvieron promedios más altos con respecto al otro grupo socioeconómico en todos los niveles e índices analizados, excepto la longitud de la cláusula en textos escritos por alumnos de séptimo básico (véase tabla 27).

La variable género discursivo presentó resultados significativos con todos los índices analizados en ambos grupos sociales, los textos narrativos fueron más extensos que los explicativos; excepto en la longitud de la cláusula y en el índice de empaquetamiento en textos escritos por alumnos del grupo socioeconómico alto, en ambos casos, hubo incrementos en favor de uno u otro modo discursivo, pero los resultados no fueron significativos (véase tabla 27).

Tabla 27.  
Resultados de los índices, según la perspectiva funcionalista

		Grupo Medio				Grupo Alto			
		Séptimo	Primero	Tercero	Total	Séptimo	Primero	Tercero	Tota
PV	Narrativo	43.10	36.10	54.35	44.52	33.50	54.35	75.30	54.38
	Explicativo	27.30	26.70	38.85	30.95	35.80	44.80	60.10	46.90
L/PC	Narrativo	30.87	23.72	32.10	28.90	27.71	25.60	43.47	32.26
	Explicativo	21.72	18.80	28.48	23.00	23.22	24.11	41.58	29.64
L/CL	Narrativo	5.64	5.65	5.75	5.68	5.31	5.62	6.00	5.64
	Explicativo	5.08	5.41	5.46	5.32	4.98	5.65	6.13	5.59
CL/PC	Narrativo	5.58	4.25	5.68	5.17	5.26	4.58	7.32	5.72
	Explicativo	4.32	3.48	5.39	4.40	4.71	4.37	6.83	5.30

Nota: PV= productividad verbal, L/PC= longitud de paquetes clausulares, L/CL=longitud de cláusulas, CL/PC= índice de empaquetamiento.

Fuente: Elaboración propia con base en Hugo (2014, pp. 41-48).

Según los resultados obtenidos con las dos perspectivas aplicadas en este estudio, Hugo (2014) concluyó que las variables grupo socioeconómico y género discursivo sí influyeron en el desarrollo de la complejidad sintáctica de textos narrativos y textos explicativos producidos por alumnos de entre doce y dieciséis años. En cuanto al grupo socioeconómico, los alumnos pertenecientes a cursos superiores del grupo socioeconómico alto obtuvieron promedios más altos que los alumnos del grupo socioeconómico medio; respecto al género discursivo, las composiciones narrativas fueron más extensas en ambos métodos aplicados, excepto las realizadas por alumnos de séptimo básico del grupo socioeconómico alto. Por el contrario, la variable edad no se consideró significativa para este estudio, pues los resultados no fueron uniformes. Por ejemplo, en la perspectiva generativista, los textos narrativos realizados por alumnos del grupo socioeconómico medio no mostraron resultados significativos en ninguno de los índices de complejidad sintáctica analizados; en la perspectiva funcionalista, los resultados mostraron un desarrollo en forma de U, es decir, los alumnos incrementaban y disminuían su complejidad sintáctica de tal manera, que los alumnos de séptimo básico obtuvieron promedios más altos que los alumnos de primero de secundaria, los de primero de secundaria obtuvieron promedios más bajos que los de séptimo, y los de tercero de secundaria obtuvieron promedios más altos que los otros dos cursos escolares. Esta última tendencia también se observó en los textos explicativos realizados por los alumnos del grupo socioeconómico alto, en ambos métodos aplicados.

Este estudio mostró que los resultados de cada investigación pueden variar según el método utilizado para el análisis de datos. En este caso, hubo mayor correlación con ambos métodos, cuando el índice de complejidad sintáctica de los alumnos se incrementó conforme ellos cambiaron de curso, lo cual se observó en los textos escritos por estudiantes del nivel socioeconómico alto. En cuanto a los índices que tuvieron menor correlación con los métodos aplicados, se encuentran: la longitud de las unidades terminales y los paquetes clausulares, y los índices de subordinación y empaquetamiento; por lo cual, Hugo (2014) estableció la unidad terminal como una unidad de subordinación y los paquetes clausulares, como indicadores relacionales. Para finalizar, la autora destacó la necesidad de desarrollar en los alumnos el dominio de la sintaxis y su vínculo con el discurso.

## 2.3 Estudios sobre complejidad sintáctica realizados en Argentina

En la Argentina, uno de los primeros trabajos realizados sobre complejidad sintáctica fue el de Delicia (2011), quien estableció el índice de madurez sintáctica de textos narrativos y textos argumentativos, escritos por jóvenes cordobeses que cursaban el cuarto año de educación secundaria. Su objetivo fue analizar la relación existente entre los índices obtenidos y las variables sexo y modo de discurso. Para tal efecto, su estudio estuvo conformado por veinte estudiantes de la especialidad de turismo procedentes de una secundaria privada de la ciudad de Córdoba, todos con un rango de entre quince y dieciséis años de edad. Para la obtención del corpus, se les solicitó a los alumnos, que en el salón de clase y con un tiempo máximo de 40 minutos, realizaran dos composiciones, siguiendo las temáticas previamente establecidas para cada modo discursivo:

### Prueba de redacción del texto narrativo

Relaten el día más feliz que vivieron en el colegio en este 2010: una anécdota en el recreo, una aventura en un paseo, alguna travesura, etcétera.

### Prueba de redacción del texto argumentativo

Imaginen que son los encargados de organizar, junto con los preceptores, el festejo del “día del maestro” en la escuela. Uno de ustedes es amigo de un conocido cantante popular de Córdoba, a quien deciden invitar el día de la fiesta para agasajar a los profesores. Cuando piden permiso a la directora del Instituto para traer al artista, ella se niega. Disconformes con la respuesta redactan una carta dirigida a la directora en la que exponen las razones por las cuales no están de acuerdo con su decisión. (Delicia, 2011, p.186)

Del corpus total, únicamente se seleccionaron las redacciones que presentaron un mínimo de doscientas palabras, y las que cumplieron con los tipos de texto solicitados. El corpus final constó de 20 composiciones, de las cuales, 10 pertenecían al modo narrativo, y 10, al modo argumentativo. Todas fueron analizadas con base en los índices primarios de Hunt (como se citó en Delicia, 2011): 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas, y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22). Para el análisis estadístico, el autor aplicó el análisis de varianza, con el fin de determinar si las diferencias de medias obtenidas fueron significativas.

Los resultados de esta investigación arrojaron índices más altos para los textos escritos por hombres en ambos modos discursivos, excepto en el promedio de longitud de las cláusulas de los

textos argumentativos, en el que los varones obtuvieron un índice de 7.21, en contraste con las mujeres, cuyo promedio de longitud de las cláusulas en este tipo de texto fue 7.86 (véase tabla 28). Sin embargo, estas diferencias no fueron estadísticamente significativas, lo que condujo a Delicia (2011) a hablar de una ausencia asociativa entre los índices de madurez sintáctica y la variable sexo. Con respecto al modo discursivo, la autora estableció que los índices más altos, tanto para hombres como para mujeres, se observaron en las composiciones argumentativas en los tres índices primarios; por lo que concluyó que el modo discursivo sí es una variable significativa. Los alumnos realizaron composiciones más complejas en la modalidad argumentativa, que en la narrativa (véase tabla 28).

Tabla 28.  
Índices primarios, en la escritura de hombres y mujeres, según el modo discursivo

Índice	Sexo masculino		Sexo femenino	
	Narrativo	Argumentativo	Narrativo	Argumentativo
	$\bar{X}$	$\bar{X}$	$\bar{X}$	$\bar{X}$
1) PAL/UT	11.83	16.41	11.60	15.95
2) PAL/CL	6.82	7.21	6.62	7.86
3) CL/UT	1.73	2.31	1.72	2.04

PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas, CL/UT= promedio de cláusulas por unidad terminal.

Fuente: Elaboración propia con base en Delicia (2011).

Por último, Delicia (2011) impulsó el desarrollo de futuros proyectos bajo esta línea de investigación, incluyendo otros tipos de discurso, o aplicando los índices secundarios de Hunt; puesto que su estudio se centró solamente en los índices primarios. También sugirió añadir análisis cualitativo de los datos, como complemento de los índices de madurez sintáctica que se obtuvieron con jóvenes cordobeses.

Por su parte, Delicia, Fernández, Muse y Porporato (2012) establecieron el índice de madurez sintáctica de alumnos procedentes del segundo año de la Facultad de Lenguas de la Universidad Nacional de Córdoba, e investigaron si existía alguna correlación entre los índices obtenidos y la variable modalidad del lenguaje (oral y escrito). Por ello, se les solicitaron exposiciones orales e informes escritos sobre un mismo tema: Exponga su experiencia como alumno investigador dentro de la asignatura Teoría y Práctica de la Investigación. Previamente, los autores informaron a los jóvenes que formarían parte de esa investigación lingüística.

Para el análisis de los datos, los autores tuvieron que transcribir las producciones grabadas, y del total de la muestra, constituida por 72 fragmentos; seleccionaron 28 textos de cada modalidad, teniendo en cuenta que los textos incluyeran doscientas palabras como mínimo, la estructura general del texto fuese expositiva, y se respetara la temática previamente establecida. Una vez obtenidos los textos, el análisis se centró en la evaluación de los índices primarios establecidos por Hunt y la adaptación de Véliz (como se citó en Delicia et al., 2012): 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas, 3) promedio de cláusulas por unidad terminal, y 4) intensidad de incrustación por unidad terminal o de recursividad oracional. En este apartado, se explicará únicamente el cuarto índice por tratarse de la primera vez que en una investigación sobre complejidad sintáctica, se incluía este índice, el cual determina el grado en el que las cláusulas se incluyen en otras, mediante la siguiente fórmula: “dividir el sumatorio del nivel de incrustación clausal por el total de unidades terminales (II/U-T)”. Los índices primarios establecidos por Hunt y Véliz se refirieron en la tabla 9 del capítulo 1 (véase p. 22).

Los resultados de esta investigación mostraron que los índices primarios en la producción de discursos escritos fueron más altos en el promedio de longitud de la unidad terminal (18.54, en contraste con 16.55), y en el promedio de longitud de las cláusulas (8.70, en contraste con 6.88); sin embargo, solamente este último resultó significativo al aplicar el análisis de varianza (véase tabla 29). En cuanto a la producción de discursos orales, aunque esta modalidad de lenguaje presentó índices más altos que la modalidad escrita en el promedio de cláusulas por unidad terminal (2.39, en contraste con 2.17), esta diferencia de 0.22 no fue estadísticamente significativa, por lo que los autores coincidieron con la afirmación de Castella (como se citó en Delicia et al., 2012) de que “La lengua oral y la escrita emplean un número similar de construcciones subordinadas, incluso con una cierta ventaja para algunos géneros orales y, sobre todo, para la lengua oral forma expositiva”.

Tabla 29.  
Los índices primarios en la oralidad y en la escritura

Índice	Modalidad oral $\bar{X}$	Modalidad escrita $\bar{X}$	ANOVA
1) PAL/UT	16.55	18.54	0.077
2) PAL/CL	6.88	8.70	0.000**
3) CL/UT	2.39	2.17	0.103
4) II/U-T	0.84	0.84	NaN

\*\* Significatividad < 0.05

PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas, CL/UT= promedio de cláusulas por unidad terminal y II/U-T= intensidad de incrustación por unidad terminal.

Fuente: Elaboración propia con base en Delicia et al. (2012).

De acuerdo con los datos anteriores, Delicia et al. (2012) concluyeron que en su estudio, solamente hubo correlación entre el índice promedio de longitud de las cláusulas y la variable modalidad del lenguaje escrito, pues este tipo de composición tuvo mayor número de palabras que las producciones orales. Respecto a la intensidad de incrustación por unidad terminal, los autores establecieron que este índice fue homogéneo en los dos tipos de discurso; pues como se observa en la tabla 29, el promedio fue de .84 en ambos casos, lo cual indica la no significatividad de ese índice novedoso en una investigación sobre complejidad sintáctica. Al finalizar, los autores analizaron la importancia de seguir evaluando la madurez sintáctica de los jóvenes universitarios, con la finalidad de mejorar su competencia lingüística.

Delicia y Muse (2013) analizaron posteriormente la relación entre los índices de complejidad sintáctica de discursos académicos escritos por alumnos de la Facultad de Lenguas de la Universidad Nacional de Córdoba y la variable nivel de escolarización. La muestra de este estudio quedó conformada por veinte universitarios, pertenecientes a segundo y quinto año, inscritos en dos asignaturas introductorias a tareas de investigación. Como cada alumno escribió una composición, el corpus total fue de veinte textos, los cuales pertenecían al género expositivo-argumentativo, y a su vez, incluían un mínimo de doscientos vocablos, o un máximo de doscientos cuarenta. Los autores seleccionaron fragmentos de un “proyecto de investigación” que se les solicitó al concluir el curso, y esos fragmentos constituyeron el corpus de este estudio sobre complejidad sintáctica.

Para la evaluación de las composiciones, los autores tuvieron como referencia la teoría de Kellogg Hunt (como se citó en Delicia y Muse, 2013), por lo cual, se propuso aplicar los índices sintácticos primarios: 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas, y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22); más el índice de incrustación de cláusulas: 4) intensidad de incrustación por unidad terminal. El análisis estadístico incluyó la obtención de la varianza, y para complementar sus resultados, los autores hicieron observaciones cualitativas.

Según los resultados de este estudio, Delicia y Muse (2013) concluyeron que la variable nivel de escolarización no presentó una relación asociativa con todos los índices sintácticos

establecidos, pues aunque los alumnos de quinto año obtuvieron índices más altos en tres de los índices aplicados (véase tabla 30), solamente en el índice promedio de cláusulas por unidad terminal (0.006) y en la intensidad de incrustación por unidad terminal (0.36), los resultados fueron significativos al obtener la varianza. Por el contrario, la diferencia entre el índice promedio de longitud de la unidad terminal de segundo de secundaria y el de quinto de secundaria, así como la diferencia entre el promedio de longitud de las cláusulas de ambos grados no presentó resultados significativos; pues esa diferencia fue mínima y los promedios se mantuvieron estables en ambos cursos.

Tabla 30.  
Índices primarios de los discursos de 2º y 5º año, según la variable nivel de escolarización

Grupo	1) PAL/U-T	2) PAL/CL	3) CL/U-T	4) II/U-T
2º año Media	24.9080	10.7540	2.1800	.6480
5º año Media	28.3380	10.7500	2.8380	.9240
Total	26.6230	10.7520	2.5090	.7860
ANOVA	.344	.998	.006**	.036**

\*\* Significatividad < 0.05

PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas, CL/UT= promedio de cláusulas por unidad terminal y II/U-T= intensidad de incrustación por unidad terminal.

Fuente: Elaboración propia con base en Delicia y Muse (2013).

Entre las observaciones cualitativas sobre el índice intensidad de incrustación por unidad terminal, se estableció que los textos escritos por los alumnos del segundo curso mostraron un nivel dos de incrustación, pero los textos escritos por los alumnos de quinto mostraron nivel tres o nivel cuatro, grado máximo. Esta asignación se realizó con base en los niveles de Véliz (como se citó en Delicia y Muse, 2013):

Nivel 0: describe unidades-t simples, sin subordinadas.

Nivel 1: describe unidades-t que incluyen cláusulas subordinadas de estructura simple, esto es, que no incluyen a su vez otra cláusula.

Nivel 2: describe unidades-t a las cuales se incrusta una cláusula compuesta, es decir, que a su vez posee una cláusula subordinada y así sucesivamente (nivel 3, nivel 4...). (p.11)

Los autores concluyeron que a mayor nivel de escolarización de los alumnos, mayor complejidad gramatical en sus textos; por esa razón, fueron los alumnos de quinto grado quienes escribieron textos más complejos. Delicia y Muse (2013) sostuvieron que el estudio de la

complejidad sintáctica puede influir en la enseñanza de la lengua para desarrollar estrategias que mejoren el desempeño lingüístico de jóvenes universitarios.

## **2.4 Estudios sobre complejidad sintáctica realizados en otros países de Sudamérica**

En Colombia, Duarte, Espejo y Hernández (2013) investigaron el índice de complejidad sintáctica de textos escritos por alumnos de la Universidad de La Salle, con el fin de analizar la relación entre los índices obtenidos y las variables género, semestre, jornada (de día y nocturna), y tipo de discurso (narrativo y argumentativo). La muestra de este estudio quedó conformada por 64 alumnos procedentes de los dos primeros y los dos últimos semestres (I, II, IX, X) de la Licenciatura en Lengua Castellana, en Inglés y en Francés; 25 pertenecían al género masculino y 39, al género femenino; quienes pertenecían al grupo inicial tenían un promedio de diecisiete a veintiún años, y los demás tenían una edad en un rango de veintitrés a veinticinco años. Cabe aclarar que al principio de la investigación, la edad se consideró como una variable; sin embargo, como muchos alumnos no mencionaron su edad, se tuvo que descartar este dato.

Para obtener el corpus, los investigadores impartieron dos talleres en diferentes sesiones, uno para explicar la narración y otro para explicar la argumentación. De este modo, al finalizar, les solicitaron a los alumnos una composición de cada modo discursivo, ya sea con los temas sugeridos durante el curso, o con el de su preferencia; por tanto, el corpus final constó de 64 textos narrativos y 64 descriptivos, los cuales se transcribieron sin cambiar la puntuación, ni la ortografía. Se analizó el corpus con base en los índices primarios de Hunt (como se citó en Duarte et al., 2013). Primero se contó el total de palabras producidas en cada composición, seguidamente se segmentaron las producciones en unidades terminales, posteriormente, se identificaron el total de cláusulas de cada texto; y, por último, se analizaron los índices utilizados en cada texto, los cuales se dividieron en dos grupos, principales y secundarios. Los índices primarios fueron: 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22). En cuanto a los índices secundarios, los autores incluyeron los más utilizados en las composiciones, por lo que analizaron las siguientes cláusulas: 11) subordinadas de relativo, 12) subordinadas adverbiales de finalidad, 13) subordinadas adverbiales de causa, 14)

subordinadas sustantivas de complemento directo y 15) subordinadas adjetivas explicativas. Este estudio fue el primero en el que se analizaron estos índices secundarios, los cuales "... se obtienen dividiendo el total de cláusulas subordinadas por el total de unidades -t" (Duarte et al., 2013, p. 20).

De acuerdo con los resultados obtenidos con el análisis estadístico, tanto en los índices primarios como en los secundarios y en ambos tipos de texto, Duarte et al. (2013) establecieron que las variables género, semestre y jornada no fueron significativas en esta investigación. Respecto a la variable género, aunque el índice de complejidad sintáctica de las mujeres fue superior al de los hombres, estas diferencias no fueron significativas. En cuanto a las variables semestre y jornada, los resultados arrojaron que no hubo diferencias significativas entre los alumnos de I y II respecto de los alumnos de IX y X; por lo que no se validó la hipótesis inicial en la que se planteaba "... que los estudiantes de los dos últimos semestres presentaban un desarrollo sintáctico y una calidad textual superior a la de los dos primeros semestres..." (Duarte et al., 2013, p. 30). Pese a estos resultados, las variables anteriores (semestre y jornada) presentaron las siguientes excepciones en textos narrativos: sí hubo diferencias significativas entre los alumnos de IX diurno y los de IX nocturno en el índice primario promedio de cláusulas por unidad terminal; y entre los alumnos de X diurno y los de X nocturno, en el índice primario promedio de longitud de la unidad terminal.

Por el contrario, la variable tipo de discurso sí fue significativa en este estudio, por lo que Duarte et al. (2013) afirmaron que el modo discursivo "... es el que más claramente demuestra tener una incidencia marcada en la producción de textos" (p. 29). Por ejemplo, como se observa en la tabla 31, los textos narrativos presentaron un incremento de madurez sintáctica en los tres índices primarios en las composiciones realizadas por alumnos de los primeros semestres del grupo nocturno; sin embargo, esta tendencia cambió en las composiciones de los últimos semestres, pues el promedio más alto de estos escritos fue para los alumnos del grupo diurno, excepto en el promedio de cláusulas por unidad terminal: 5.5 con respecto a 5.0. En los textos argumentativos, escritos por alumnos de I y II, ya no fueron los del grupo nocturno los que presentaron promedios más altos, sino los del grupo diurno, a excepción del índice promedio de cláusulas por unidad terminal: 4.0 con respecto a 3.9. En las composiciones realizadas por alumnos de IX y X, los promedios más altos correspondieron a los alumnos del grupo diurno.

Tabla 31.

## Índices primarios de textos narrativos y argumentativos

Índices	Texto	Grupo inicial y final		Semestre y jornada					Género	
		I y II	IX y X	I y II diurno	I y II nocturno	IX y X diurno	IX y X nocturno	Promedio general	Fem.	Mas.
PAL/UT	Narrativo	12.8	12.4	12.2	13.4	14.0	10.8	12.6	12.1	13.5
	Argumentativo	13.2	12.8	13.4	12.9	12.4	13.2	13.0	13.2	12.7
PAL/CL	Narrativo	47.3	58.0	40.2	54.5	62.2	53.7	52.6	53.1	51.8
	Argumentativo	43.5	47.8	44.1	43.0	52.4	43.1	45.7	47.9	41.6
CL/UT	Narrativo	4.5	5.3	3.9	5.2	5.0	5.5	4.9	4.9	4.9
	Argumentativo	3.9	4.4	3.9	4.0	4.4	4.3	4.1	4.3	3.9

PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas, CL/UT= promedio de cláusulas por unidad terminal.

Fuente: Elaboración propia con base en Duarte et al. (2013).

Continuando con la variable tipo de discurso, Duarte et al. (2013) destacaron que los textos narrativos fueron más extensos que los argumentativos; no obstante, los alumnos utilizaron más cláusulas coordinadas copulativas, coordinadas adversativas y yuxtapuestas. Por el contrario, los textos argumentativos presentaron composiciones más elaboradas, puesto que este modo discursivo implica un uso de estructuras sintácticas más complejas en la escritura; cabe mencionar que este modo solamente fue producido por alumnos de los últimos semestres. Ahora bien, como los resultados de este estudio mostraron deficiencias en las composiciones de los estudiantes, los autores propusieron reformular las estrategias metodológicas de la enseñanza de la escritura, para que los jóvenes desarrollen las habilidades que les permitan usar la lengua escrita en diferentes situaciones comunicativas.

Otra aportación para Sudamérica fue el estudio realizado por Trelles (2015), quien por medio de una prueba piloto, estableció el índice de madurez sintáctica de composiciones escritas por alumnos de nuevo ingreso a la facultad de Ciencias Económicas y Empresariales de la Universidad de Piura, en el Perú. La muestra quedó conformada por 38 alumnos, hombres y mujeres, procedentes de secundarias nacionales y particulares de la Educación Básica Regular del Perú. Para la obtención del corpus, la autora diseñó una prueba que constó de cuatro secciones: 1) evaluación sobre ortografía, 2) conocimientos gramaticales, 3) comprensión lectora y 4) redacción. Este último apartado, que fue el que se analizó sintácticamente, consistió en solicitar a los alumnos que escribieran un párrafo de ocho líneas a partir del siguiente vocabulario: miedo, esperanza, empresario, empresa, los trabajadores, la orden, oficina, préstamo, generosidad y rigidez.

Una vez obtenido el corpus, Trelles (2015) segmentó los textos en cláusulas y unidades terminales, siguiendo los criterios de Véliz (como se citó en Trelles, 2015); asimismo, clasificó los textos en oraciones simples, coordinadas y subordinadas sustantivas, subordinadas adjetivas y

subordinadas adverbiales. Se aplicaron los índices primarios y los secundarios. Índices primarios: 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22); índices secundarios: 4) promedio de cláusulas adjetivas por unidad terminal, 5) promedio de cláusulas sustantivas por unidad terminal, y 6) promedio de cláusulas adverbiales por unidad terminal (véase tabla 10, p. 22).

Según los resultados de este estudio, el corpus estuvo conformado por 2, 618 palabras y 182 unidades terminales. En los índices principales, se observó que los alumnos produjeron más oraciones coordinadas que subordinadas, 42.24% respecto a 35.34%, asimismo, cada unidad terminal tuvo al menos una cláusula, puesto que el índice promedio de longitud de las cláusulas (14.21) fue muy aproximado al promedio de longitud de la unidad terminal (16.71). Respecto a los índices secundarios, fue significativo el mayor empleo de oraciones subordinadas sustantivas utilizadas como complementos directos y complementos preposicionales; los alumnos produjeron más cláusulas sustantivas por unidad terminal (0.35), cláusulas adverbiales por unidad terminal en menor porcentaje, 0.19 y con un promedio similar de 0.18, cláusulas adjetivas por unidad terminal. El modo discursivo predominante en este estudio fue el modo narrativo, con un 78.95% con respecto a una combinación de modo expositivo-narrativo-instructivo, con un 21.05%.

De acuerdo con los resultados anteriores, Trelles (2015) determinó que los alumnos que egresan de la secundaria son capaces de crear composiciones escritas (con predominio del modo discursivo), partiendo de un vocabulario previamente establecido, en las cuales incluyen oraciones coordinadas y oraciones subordinadas; no obstante, en las narraciones del corpus también se encuentran otras limitaciones en cuanto a aspectos gramaticales y ortográficos. Ahora bien, puesto que este estudio se trató de una prueba piloto, la autora no estableció estas conclusiones como definitivas, pues para ello, se considera necesario aplicar una segunda prueba a los alumnos al finalizar el ciclo escolar.

Recientemente, Ibarra (2018) estableció en su tesis de licenciatura el índice de madurez sintáctica de composiciones narrativas escritas por estudiantes de primero, segundo y tercer grado de secundaria, procedentes de centros educativos del departamento de Lima, Perú. Su objetivo fue determinar si se producía algún incremento de madurez sintáctica según las variables nivel escolar y procedencia (cuatro zonas de Lima); asimismo, si existía alguna correlación entre los resultados obtenidos y la influencia de la variable uso de las tecnologías de la información y la comunicación.

En cuanto a esta última variable establecida, Ibarra (2018) diseñó un cuestionario de opciones múltiples para obtener los datos sociológicos de cada participante, el cual incluía dos preguntas principales sobre qué medios de internet eran los más utilizados por los alumnos, y cuántas horas diarias le dedicaban los alumnos a esos medios; puesto que esa información era relevante para su objetivo de investigación (véase figura 2). De acuerdo con los resultados del cuestionario, esta investigadora asignó tres distinciones al número de horas que los alumnos dedicaban a navegar en internet y en las redes sociales, ya fuera en un dispositivo móvil, en una tableta o en su computadora: P, R y M, abreviaturas de poco, regular y mucho. Si los estudiantes se exponían de 0 a 3 horas por día a las tecnologías de la información y la comunicación, lo consideraba poco; si el tiempo era entre 4 a 6 horas por día, lo consideraba regular; y lo consideraba mucho, cuando el tiempo era de 7 a 9 horas por día. Ibarra (2018) estableció como negativa, una correlación entre mayores horas de exposición a internet y menores índices de madurez sintáctica; y como positiva, una correlación entre menor horas de exposición a internet y mayores índices de madurez sintáctica.

Figura 2.  
Preguntas principales del cuestionario diseñado

2. ¿Cuál de los siguientes medios de Internet utilizas? (puedes seleccionar más de una respuesta)						
e-mail	<input type="checkbox"/>	Facebook	<input type="checkbox"/>	Twitter	<input type="checkbox"/>	
Buscadores web (google..)	<input type="checkbox"/>	Messenger	<input type="checkbox"/>	WhatsApp	<input type="checkbox"/>	
Line	<input type="checkbox"/>	Foros	<input type="checkbox"/>	Blogs	<input type="checkbox"/>	
4. ¿Cuántas horas le dedicas diarias a estos medios de Internet?						
e-mail	<input type="checkbox"/>	Facebook	<input type="checkbox"/>	Twitter	<input type="checkbox"/>	
Buscadores web (google..)	<input type="checkbox"/>	Messenger	<input type="checkbox"/>	WhatsApp	<input type="checkbox"/>	
Line	<input type="checkbox"/>	Foros	<input type="checkbox"/>	Blogs	<input type="checkbox"/>	

Fuente: Ibarra (2018, p. 73).

La muestra quedó conformada por 117 alumnas de nivel secundaria, 41 de primer grado, 31 de segundo grado, y 45 de tercer grado. La procedencia de las alumnas seleccionadas se estableció con base en el nivel socioeconómico del distrito al cual pertenecían: Lima Centro, Lima Norte, Lima Sur y Provincia Constitucional del Callao; no obstante, debido a que ninguna de las alumnas pertenecía a Lima Sur, no se presentan datos correspondientes a esa zona. Para obtener el corpus, se les solicitó a las participantes, que en el ámbito escolar y con un tiempo máximo de 40 minutos,

realizaran la composición de un discurso narrativo, siguiendo alguna de las temáticas previamente establecidas:

Prueba de redacción de texto narrativo:

Elige uno de los siguientes temas y redacta un texto de una cara (30 líneas)

¿Cuál fue la última película que viste, de qué trató, qué te gustó más?

¿Qué hiciste en tu último viaje o paseo?

Inventa un cuento, puede ser de terror, ficción, aventura, etc. (Ibarra, 2018, p. 45)

Del corpus total, únicamente se seleccionaron las redacciones que presentaron un mínimo de doscientas palabras; por tanto, el corpus final constó de 75 composiciones (primer grado, 23 textos; segundo grado, 21 textos y tercer grado, 31 textos). Para el análisis del corpus, la autora partió de las fórmulas propuestas por Hunt (como se citó en Ibarra, 2018), no obstante, en este estudio solamente se aplicaron los índices primarios: 1) promedio de longitud de la unidad terminal, 2) promedio de longitud de las cláusulas y 3) promedio de cláusulas por unidad terminal (véase tabla 9, p. 22). Por consiguiente, primero se dividió cada texto en unidades terminales y en cláusulas, después se contó el total de palabras de cada texto; y posteriormente, se contó la cantidad de cláusulas principales y cláusulas subordinadas.

El análisis estadístico de este estudio se realizó mediante el programa informático *Statistical Package for Social Sciences versión 20.0* en español, versión para el sistema operativo Macintosh. Se realizó el análisis de varianza para determinar las diferencias entre las medias de los grupos, en algunos casos, dichos resultados se contrastaron con la prueba de *Kruskal Wallis*; y para medir el nivel de significación de los resultados obtenidos, se realizó un análisis *post-hoc* mediante la prueba de *Bonferroni* y la prueba de *U de Mann-Whitney* (como se citó en Ibarra, 2018). La autora también aplicó el análisis de regresión logística, “el cual permite determinar si hay o no influencia entre dos variables” (Ibarra, 2018, p. 60), para determinar si la variable uso de las tecnologías de la información y la comunicación influía en el grado de madurez sintáctica de los alumnos.

Los resultados de este estudio mostraron que en los tres índices analizados, hubo un incremento de madurez sintáctica entre los textos escritos por alumnas de primero a segundo grado, no obstante, la diferencia de medias sólo fue significativa en el promedio de longitud de la unidad terminal (primero, 8.99 en contraste con segundo, 10.31) y en el promedio de longitud de la cláusula (primero, 6.68 en contraste con segundo, 7.51). Contario a lo anterior, no se observó ningún aumento de madurez sintáctica, en ninguno de los tres índices, entre los textos escritos por

alumnos de segundo a tercer grado (véase tabla 32). En cuanto a la diferencia de medias entre los textos de primero a tercer grado, esta tampoco fue significativa en ningún caso.

Tabla 32.  
Resultados generales de los índices primarios, según la variable nivel escolar

Grados	N	Media		
		PAL/UT	PAL/CL	CL/UT
Primero	23	8.9983	6.6896	1.3465
Segundo	21	10.3157	7.5129	1.3724
Tercero	31	9.3655	7.0410	1.3265
Total	75	9.5189	7.0653	1.3455

PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas, CL/UT= promedio de cláusulas por unidad terminal, N= número.

Fuente: Elaboración propia con base en Ibarra (2018, pp. 49-54).

Para la variable procedencia, los resultados arrojaron que en los tres índices analizados, los textos que presentaron mayor complejidad sintáctica fueron los de las alumnas pertenecientes a la zona de Lima Norte: promedio de longitud de la unidad terminal (Lima Norte, 10.34 en contraste con el Callao, 9.54 y Lima Centro, 9.47), promedio de longitud de las cláusulas (Lima Norte, 7.76 en contraste con el Callao, 7.11 y Lima Centro, 7.01) y promedio de cláusulas por unidad terminal (Lima Norte, 1.37 en contraste con el Callao, 1.34 y Lima Centro, 1.34) véase tabla (33) . Sin embargo, estos resultados no fueron significativos en ningún caso al realizar el análisis de varianza.

Tabla 33.  
Resultados generales de los índices primarios, según la variable procedencia

Distritos por zonas	Media		
	PAL/UT	PAL/CL	CL/UT
Lima Centro	9.4735	7.0163	1.3445
Provincia Constitucional del Callao	9.5495	7.1159	1.3455
Lima Norte	10.3400	7.7600	1.3700
Total	9.5189	7.0653	1.3455

PAL/UT= promedio de longitud de la unidad terminal, PAL/CL= promedio de longitud de las cláusulas, CL/UT= promedio de cláusulas por unidad terminal.

Fuente: Elaboración propia con base en Ibarra (2018, pp. 55-58).

En cuanto al uso de las tecnologías de la información y la comunicación, los resultados mostraron que esta variable sólo influyó en el índice promedio de cláusulas por unidad terminal. El análisis de regresión logística demostró que esta variable presentó una correlación negativa con este índice, pues tal como se estableció inicialmente, entre más horas de exposición a internet, las

alumnas presentaron menores índices de madurez sintáctica en sus composiciones. Cabe mencionar que fueron las alumnas de tercer grado, quienes pasaban más tiempo (7 a 9 horas por día) expuestas a las tecnologías de la información y la comunicación, a diferencia de las alumnas de primero y segundo grado (tercero, 41.9 % en contraste con segundo 28.6 % y primero, 31.8 %), (véase tabla 34).

Tabla 34.  
Número de horas que las alumnas reportan hacen uso de las TIC

Grado		Horas que dedica al internet			Total
		Poco	Regular	Mucho	
Primero	Recuento	11	4	7	22
	Porcentaje	50 %	18.2 %	31.8 %	100.0 %
Segundo	Recuento	11	4	6	21
	Porcentaje	52.4 %	19.0 %	28.6 %	100.0 %
Tercero	Recuento	12	6	13	31
	Porcentaje	38.7 %	19.4 %	41.9 %	100.0 %
Total	Recuento	34	14	26	74
	Porcentaje	45.9 %	18.9 %	35.1 %	100.0 %

Fuente: Ibarra (2018, p. 60).

De acuerdo con los resultados anteriores, Ibarra (2018) afirmó “que el uso de las TIC es un factor que influye para que no se evidencie un crecimiento en la complejidad sintáctica de las alumnas de tercero respecto de las de segundo y más aún de las de primero” (p. 65), lo cual confirma su hipótesis inicial en la que estableció que el grado de complejidad sintáctica de las alumnas tendría una correlación negativa con el uso de las tecnologías de la información y la comunicación. Respecto al incremento de madurez sintáctica de las alumnas según su nivel escolar y su procedencia, la autora estableció, de acuerdo con los resultados obtenidos, que no se observó tal desarrollo en ninguno de los dos casos; pues los alumnos de segundo podían realizar estructuras sintácticas complejas, en el mismo nivel que los de tercer grado. Asimismo, las alumnas de las cuatro zonas de Lima tendrían el mismo nivel de madurez, para formar composiciones estructuralmente complejas.

Para finalizar, la autora menciona que este estudio mostró de manera general cómo se encuentra el nivel de composición de alumnas de los tres cursos de secundaria de la ciudad de Lima. Y recomienda seguir esta línea de investigación con muestras más amplias en diversos colegios del mismo distrito estudiado, así como ofrecer talleres para mejorar la ortografía y la redacción de los alumnos, con el fin de estimular en ellos el hábito de la escritura.

### **CAPÍTULO 3. REFLEXIONES CRÍTICAS SOBRE LOS ESTUDIOS DE COMPLEJIDAD SINTÁCTICA REALIZADOS FUERA DE MÉXICO**

La revisión bibliográfica sobre los estudios de complejidad sintáctica realizados en España y en Hispanoamérica (excepto en México) indicó que las primeras investigaciones datan a partir de la década de los ochenta, y que Hispanoamérica cuenta con una amplia investigación respecto a este tema, pues de los 23 estudios analizados en este estudio monográfico, 17 pertenecen a esta región, y las demás pertenecen a España. De acuerdo con el total de trabajos consultados sobre este tema, los primeros estudios que se llevaron a cabo sobre complejidad sintáctica fueron el de Salvador (1985) en la ciudad de Granada, España y el de Véliz, Muñoz y Echeverría (1985) en la ciudad de Concepción, Chile. En cuanto a los estudios más recientes, se consultaron el de Bartolomé (2016) realizado en Madrid, España y el de Ibarra (2018) realizado en Lima, Perú. Si bien la complejidad sintáctica es un tema cuyo estudio no ha alcanzado un gran desarrollo, se observó que desde los primeros trabajos realizados en los años ochenta hasta la actualidad, se ha evolucionado en algunos aspectos, como por ejemplo: la metodología, las variables seleccionadas, la recolección del corpus, entre otros, los cuales se describirán de manera más específica en los siguientes párrafos.

Los diversos autores realizaron sus investigaciones con base en una perspectiva teórica que les permitiera conocer el desarrollo de las estructuras sintácticas en composiciones escritas u orales. De este modo, el 80% de los estudios presentados tuvieron como referencia los índices propuestos por Hunt; sin embargo, en algunos casos, además de seguir la teoría de Hunt, los investigadores aplicaron otro método con el fin de obtener resultados más completos. Por tanto, en los siguientes estudios, además de aplicar esos índices, algunos autores complementaron sus investigaciones con un segundo método: tal es el caso de Hugo (2014), el cual incluyó la perspectiva funcionalista (paquetes clausulares); M. Torres (2001), quien también contabilizó la oración simple y la oración compleja, dos unidades básicas de la sintaxis; y Véliz, Muñoz y Echeverría (1985), quienes incluyeron el método de combinación de oraciones. Sólo hubo tres investigaciones que aplicaron teorías distintas de las ya mencionadas, de las cuales dos fueron realizadas por el mismo autor, Salvador (1985, 1995), quien basó su primer estudio en el índice diseñado por Simon y diez años más tarde, realizó otro estudio para el cual tuvo como referencia el índice de subordinación

diseñado por Loban; Balboa, Crespo y Rivadeneira (2012) tuvieron como referencia la teoría del déficit de Bernstein, así como la sociolingüística variacionista. Por su parte, Coloma, Fernández y Peñaloza (2007) fueron la excepción en cuanto a la aplicación de algún marco teórico para su estudio; no obstante, sí establecieron los términos que tuvieron en cuenta para el análisis de sus composiciones.

Las variables utilizadas con mayor frecuencia fueron: “curso”, “clase social”, “sexo”, “modo de discurso” y “edad”. En varios de los estudios analizados, se observaron dos indicadores en común para determinar la clase social de los participantes, con base en el nivel académico y la profesión de sus padres (Balboa, Crespo y Rivadeneira, 2012; Salvador, 1985,1995; A. Torres, 1993; M. Torres, 2001); y según el tipo de institución a la que acudían, escuela pública y escuela privada (Espinet, 1996; Hugo, 2014; Montenegro, 1988, 1991). Es importante mencionar que pocos fueron los estudios en los que se procuró contar con una cantidad equitativa de participantes por variable.

En el caso de la variable clase social, fue más común encontrar un porcentaje elevado de participantes que pertenecieran a la clase baja y no, a la clase alta, tal es el caso de Salvador (1985), quien conformó su muestra con un 55.3% perteneciente a la clase baja en contraste con un 13% perteneciente a la clase alta, A. Torres (1993) conformó su muestra con un 57.2% perteneciente a la clase baja, en contraste con un 17.1% perteneciente a la clase alta, por mencionar algunos. El único estudio que contó con una muestra equitativa con respecto a la variable clase social fue el de Balboa, Crespo y Rivadeneira (2012), pues del total de su muestra, el 50% de los informantes pertenecía a la clase baja y el otro 50%, a la clase alta.

En cuanto a la variable curso y la variable edad, se marcaron diferencias muy dispersas entre un grupo u otro; respecto a la variable sexo, se observó más equidad entre el número de participantes de ambos géneros, no obstante, también se observaron algunos casos donde había mayor número de participantes del sexo masculino que del sexo femenino, o viceversa. En cuanto al modo de discurso, la mayoría de los investigadores que utilizó esta variable optó por textos narrativos, descriptivos y argumentativos, tanto en la modalidad oral como en la escrita.

Se encontraron cuatro variables a las que califico como innovadoras porque pocos autores las incluyeron en su análisis: M. Torres (2001) incluyó la variable “opción” y la variable “relación con los medios de comunicación”, en el primer caso, la autora midió el índice sintáctico del alumno según el área de estudio al que pertenecía (Ciencias o Letras), y en el segundo caso, midió la

complejidad según la relación entre el alumno y los medios de comunicación a los que este tenía acceso (televisión, radio o prensa); Duarte, Espejo y Hernández (2013) establecieron la variable “jornada” para medir el nivel de complejidad sintáctica de los alumnos según el turno en el que estos acudían a la escuela, ya fuera de día o de noche; y Bartolomé (2016) e Ibarra (2018) incluyeron la variable “horas de exposición a las tecnologías de la información y la comunicación”. Esta última variable se encontró en dos de los estudios más recientes, y según los resultados obtenidos, sería interesante incluirla con mayor frecuencia en los trabajos futuros, debido al fuerte impacto que se tiene actualmente con las tecnologías de la información, tanto en los jóvenes como en los adultos, además de que sería de gran apoyo para los docentes y población en general, el conocer si la influencia de las redes sociales podría llegar a repercutir en el desarrollo sintáctico de los estudiantes, ya sea de manera positiva o negativa.

En cuanto al tipo de lenguaje solicitado para la recolección del corpus en los estudios previos, la mayoría de los autores optó por la lengua escrita, pues del total de los trabajos analizados, se observó que sólo tres estudios elicitaban narraciones de manera oral: Balboa, Crespo y Rivadeneira (2012); Coloma, Fernández y Peñaloza (2007); y Sánchez (1998). Si bien estos autores obtuvieron el corpus a través de conversaciones orales, es importante mencionar que en los tres casos, las narraciones de los alumnos fueron grabadas, y posteriormente, transcritas para su análisis. Respecto a la lengua escrita, se observó que los alumnos mostraban mejores producciones textuales cuando se les solicitaban redacciones de tema libre; sin embargo, los temas controlados fueron los que más predominaron. Entre las técnicas más comunes para obtener el corpus de manera controlada, destacan el recontado de cuentos (Coloma, Fernández y Peñaloza, 2007), la proyección de videos (Balboa, Crespo y Rivadeneira, 2012; Hugo, 2014) y la proposición previa de una lista de temas para que los alumnos elijan el de su preferencia (Bartolomé, 2016; Delicia, 2011; Ibarra, 2018; Montenegro, 1991).

Después del análisis y de la revisión de la bibliografía expuesta en este estado del arte sobre los estudios de complejidad sintáctica en español, he llegado a la conclusión de que la recolección de datos de lengua oral es una técnica poco aplicada por los investigadores con informantes adolescentes o adultos, pues en el caso de los trabajos pertenecientes a España, todos los autores consultados optaron por lengua escrita. Los únicos trabajos en los que se elicitaban datos de manera oral fueron realizados en Hispanoamérica, específicamente, en Chile y en Costa Rica, lo cual permite sugerir a los futuros investigadores inclinarse por esta técnica poco aplicada, pues la lengua

oral permitiría obtener un corpus más espontáneo y un panorama más amplio sobre la complejidad sintáctica; o en su caso, aplicar ambos métodos, tal como lo realizaron Delicia, Fernández, Muse y Porporato (2012), quienes solicitaron exposiciones orales e informes escritos sobre un mismo tema a sus participantes. De este modo, esta fue la única investigación que contrastó los resultados de las producciones escritas respecto de las producciones orales.

En cuanto al análisis de los datos, para poder medir el índice de complejidad sintáctica, los autores primero clasificaron las oraciones del corpus obtenido, y posteriormente, puesto que estas investigaciones son corte cuantitativo, posteriormente sometieron los datos a un análisis estadístico. En los estudios realizados en la década de los noventa, se encontró que para el análisis de las composiciones, Salvador (1985) optó por crear sus propias fórmulas con base en otros autores, por lo cual, creó cuatro índices de complejidad sintáctica, basándose en el índice diseñado por Simon (como se citó en Salvador, 1985); para su estudio de 1995, Salvador consideró la oración y la proposición como unidades sintácticas, y estableció las clases de oraciones subordinadas que se considerarían en su estudio con base en la terminología gramatical propuesta por la Administración Educativa (como se citó en Salvador, 1995).

Respecto los demás estudios, tanto en España como en Hispanoamérica, la teoría que más predominó en el análisis de los textos fue la de Kellogg Hunt (como se citó en A. Torres, 1993), quien basó su terminología sintáctica en la unidad terminal y en la cláusula, habiendo propuesto dos tipos de índices sintácticos para medir la complejidad oracional: índices primarios e índices secundarios. Recientemente, los investigadores han optado por la perspectiva funcionalista, la cual mide la complejidad oracional mediante la segmentación de paquetes clausulares y mediante la segmentación de cláusulas (véase tabla 25, p. 51). Sin embargo, también hay casos en los que se contrastan los resultados de ambos métodos, tal como se observó en la investigación realizada por Hugo (2014), quien analizó el desarrollo de la complejidad sintáctica desde dos perspectivas: la funcionalista (paquetes clausulares) y la generativista (índices de Hunt), para identificar las similitudes y las diferencias entre los resultados obtenidos con ambos métodos.

Un gran porcentaje de los estudios analizados procesó sus datos estadísticos, aplicando el análisis de varianza y la prueba de *Kruskal -Wallis*, los cuales permiten determinar las diferencias de medias entre los resultados obtenidos. Actualmente también se aplica la técnica de regresión logística, “la cual permite determinar si hay o no influencia entre dos variables” (Ibarra, 2018, p.

60). Es importante mencionar que existen programas de computación que facilitan el procesamiento de datos, y por consiguiente, brindan una mayor precisión en los resultados.

En general, puede observarse que los estudios realizados analizaron diferentes variables, no aplicaron la misma metodología y la recolección del corpus varío según su objetivo y la zona geográfica en la que se obtuvo. Los trabajos contemporáneos, por ejemplo, incluyeron nuevas variables y nuevas técnicas de recolección de datos, buscando satisfacer necesidades del presente; sin embargo, el hecho de innovar tuvo cierta desventaja en el momento de presentar los resultados, pues no se contaba con un sustento teórico, o con investigaciones análogas previas con las cuales se pudieran establecer una comparación. Un claro ejemplo sería el caso de la variable “uso de las tecnologías de la información y la comunicación” incluida en los trabajos más recientes (Bartolomé, 2016; Ibarra, 2018). En cuanto a las técnicas que se consideraron novedosas para la recolección de datos, se encuentra la proyección de videos, en algunos casos sin audio y en otros casos, videos diseñados especialmente para incitar a algún tipo de producción textual en específico.

De acuerdo con el material consultado para esta investigación documental, el estudio de la complejidad sintáctica ha ido evolucionando con el paso del tiempo, de tal manera que España e Hispanoamérica cuentan ya con bibliografía sobre este tema; sin embargo, es necesario incrementar el número de estudios, aplicando metodologías apropiadas, con muestras más amplias y equitativas, cuando así se requiera (por ejemplo, con la variable sexo, clase social, entre otros), e incluyendo recolección de corpus oral. Este trabajo monográfico también muestra que realizar este tipo de estudios en colegios públicos y en particulares permite conocer el nivel en el que se encuentran sus alumnos respecto al desarrollo sintáctico de la lengua, y que con base en esos resultados, deberían mejorarse las técnicas de enseñanza de producción de estructuras sintácticas complejas orales o escritas, para ayudar a los alumnos a desarrollar todos los componentes lingüísticos que les permitan producir textos, cuyo índice de complejidad sintáctica corresponda a su nivel académico.

## CONCLUSIONES

Si bien el estudio de la complejidad sintáctica fuera de México ha presentado grandes aportes para la investigación lingüística, aún no se ha alcanzado un gran desarrollo sobre el estudio de este tema con respecto a la lengua española. España e Hispanoamérica cuentan ya con diversas investigaciones respecto a esta temática, y se destaca que la región Hispanoamericana muestra una producción más amplia y generalizada. Después de un riguroso análisis de la teoría lingüística, la metodología aplicada, los resultados y las conclusiones en los veintitrés estudios consultados, se permitió cubrir el objetivo principal de esta investigación; describir minuciosamente los estudios previos, para plantear una reflexión crítica sobre los estudios de complejidad sintáctica realizados en España y en Hispanoamérica.

Como consecuencia, se estableció que las primeras investigaciones respecto a este tema se realizaron en la década de los ochenta y en la década de los noventa. En España, destacaron Salvador (1985, 1995) y A. Torres (1993, 1997); en Hispanoamérica, los investigadores pioneros fueron Espinet (1996), Montenegro (1988, 1991), Sánchez (1998), y Véliz (1985, 1988, 1999). Por su parte, Bartolomé (2016) e Ibarra (2018) han aportado en sus investigaciones consultadas la influencia de una variable innovadora “horas de exposición a las tecnologías de la información y la comunicación” en el desarrollo sintáctico de jóvenes estudiantes.

Un gran porcentaje de los trabajos analizados se basó en una teoría lingüística, por tanto, se concluyó que hubo tres enfoques predominantes: el tradicionalista, el generativista y el funcional-discursivo. La teoría generativista fue la más aplicada por los investigadores, está representada fundamentalmente por la investigación de Hunt, quien utilizó un método de medición cuantitativo, sustentado en la aplicación de sus ya conocidos índices primarios e índices secundarios. El enfoque tradicionalista se utilizó principalmente en trabajos que buscaron implementar nuevos índices de medición sintáctica, como en el caso de Salvador (1985, 1995), y cuando se buscó medir la complejidad sintáctica mediante dos teorías, como en el caso de Véliz, Muñoz y Echeverría (1985). El enfoque funcional-discursivo se encontró en trabajos recientes, específicamente en el de Hugo (2014), esta perspectiva incluyó criterios cualitativos y cuantitativos por lo que se midió la complejidad oracional mediante la segmentación de paquetes clausulares y de cláusulas.

Examinar la metodología de los estudios previos también permitió observar que de los 23 estudios encontrados, 19 analizaron lengua escrita, 3 se enfocaron en la lengua oral, y solamente el estudio de Delicia, Fernández, Muse y Porporato, (2012) se interesó en el análisis de las dos modalidades de lengua. Respecto a las variables incluidas, se observó que “curso”, “clase social” y “sexo” se encontraron en el 90% de los estudios analizados, sin embargo, no en todos los casos, estas variables resultaron pertinentes y significativas; la variable “curso” fue la que más presentó resultados significativos, por ejemplo, en el estudio de Bartolomé (2016), Montenegro (1988), A. Torres (1993), entre otros; pues de los 13 estudios en los cuales se estudió dicha variable, sólo 4 no presentaron relación asociativa con los índices establecidos. Aunque las variables “clase social” y “sexo” se presentaron en un gran porcentaje de estudios, los resultados no fueron significativos, como se planteó inicialmente.

Dadas las consideraciones anteriores, se espera que este estado del arte sea útil para estudiantes, profesores y próximos investigadores que se interesen por el estudio de la complejidad sintáctica. El proceso de revisión y análisis de los estudios aquí presentados también puede considerarse una primera etapa de investigación sobre estudios de complejidad sintáctica, y el cimiento de un trabajo sólido en el futuro en el que se mejore la metodología de medición y análisis de los datos, con base en una relación de variables pertinentes en el estado de Quintana Roo. Por otra parte, debido a que el estudio de la complejidad sintáctica se encuentra en pleno desarrollo, también sería conveniente considerar este trabajo como una fuente de consulta que permita discernir entre la teoría lingüística y la metodología más conveniente por seguir, para elaborar futuras investigaciones. Con el fin de obtener resultados más objetivos, se recomienda combinar los métodos cuantitativos con los cualitativos, establecer muestras más amplias, considerar muestras equitativas con las variables que así lo requieran, incluir el análisis sintáctico de la lengua oral o combinarla con el de la escrita, así como incluir otras modalidades del discurso.

## REFERENCIAS

- Alarcón, L., y Auza, A. (2015). Uso y función de nexos en la subordinación y coordinación. Evidencia de dos tareas narrativas de niños mexicanos de primero de primaria. En I. Rodríguez, y E. Velázquez (Eds.), *Estudios de Lingüística Funcional* (pp. 223-251). Querétaro, Querétaro: Editorial Universitaria.
- Alarcón, L., y García, K. (2015). ¿Cómo conectan los jóvenes escolares sus ideas para argumentar? Análisis del uso de conectores en su discurso. *Entreciencias: Diálogos en la Sociedad del Conocimiento*, 3(7), 253-262.
- Balboa, C., Crespo, N., y Rivadeneira, M. (2012). El desarrollo de la sintaxis en la adolescencia: posibles influencias de naturaleza social. *Literatura y Lingüística*, (25), 145-168.
- Bartolomé, R. (2016). *Madurez sintáctica: influencia de las TIC en sus índices y estudio comparativo entre las generaciones pre y post-internet* (Tesis doctoral). Universidad Complutense de Madrid, Madrid, España.
- Becerra, Y. (2014). *Coordinación y subordinación en la narración oral de niños y jóvenes del primer y último año de la educación básica mexicana* (Tesis de pregrado). Universidad Autónoma de Querétaro, Querétaro, Querétaro.
- Cohuo, A. (2015). Syntactic Complexity and the Effects of Transfer in EFL Students' Writing at Universidad de Quintana Roo (Tesis de maestría). Universidad de Quintana Roo, Chetumal, Quintana Roo.
- Coloma, C., Fernández, R., y Peñaloza, C. (2007). Producción de oraciones complejas en niños de 8 y 10 años. *RLA: Revista de Lingüística Teórica y Aplicada*, 45(1), 33-41.
- Delicia, D. (2011). Madurez sintáctica y modos de organización del discurso: un estudio sobre la competencia gramatical adolescente en producciones narrativas y argumentativas. *Onomázein: Revista de Lingüística, Filología y Traducción*, 24(2), 173-198.
- Delicia, D., y Muse, C. (2013). Complejidad sintáctica y discurso académico: parámetros gramaticales en la producción escrita de estudiantes universitarios. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, (14), 4-14.
- Delicia, D., Fernández, M., Muse, C., y Porporato, G. (2012). Madurez sintáctica en estudiantes universitarios: un estudio comparativo sobre la producción del discurso académico oral y escrito. En I. Bosio, V. Caltel, G. Ciapusio, L. Cubo, y G. Müller (Eds.), *Discurso especializado: estudios teóricos y aplicados* (pp. 209-221). Mendoza: Editorial Facultad de Filosofía y Letras, Universidad Nacional de Cuyo y Sociedad Argentina de Lingüística.

- Duarte, G., Espejo, M., y Hernández, G. (2013). La madurez sintáctica en español de los dos primeros y los dos últimos semestres de la Licenciatura en Lengua Castellana, Inglés y Francés de la Universidad de La Salle. *Lenguaje*, 41(1), 17-34.
- Espinet, L. (1996). Índices primarios de madurez sintáctica en escritores profesionales puertorriqueños: variables sexo y modo del discurso. *Revista de Estudios de Adquisición de la Lengua Española*, (6), 37-52.
- Gómez, M., Galeano, C., y Jaramillo, D. (2015). El estado del arte: una metodología de investigación. *Revista Colombiana de Ciencias Sociales*, 6(2), 423-442.
- Herrera, M. (1991). Madurez sintáctica en escolares de Ciudad de México. Análisis preliminar. En H. López (Ed.), *La enseñanza del español como lengua materna: Actas del II seminario internacional sobre aportes de la lingüística a la enseñanza del español como lengua materna* (pp. 155-169). Río Piedras: Universidad de Puerto Rico.
- Hugo, E. (2014). *La medición de la complejidad sintáctica y su relación con la edad, el grupo socioeconómico y el género discursivo en dos perspectivas de estudio* (Tesis de maestría). Pontificia Universidad Católica de Chile, Santiago, Chile.
- Ibarra, M. (2018). *Madurez sintáctica en textos narrativos de alumnas de educación secundaria y la influencia de las TIC* (Tesis de pregrado). Universidad Nacional Mayor de San Marcos, Lima, Perú.
- López, A. (2013). *Complejidad sintáctica observada en el discurso argumentativo de preadolescentes, adolescentes y adultos jóvenes de Querétaro* (Tesis de pregrado). Universidad Autónoma de Querétaro, Querétaro, Querétaro.
- Maya, S. (2014). *Conectores en la complejidad sintáctica de producciones narrativas de niños en etapa escolar* (Tesis de maestría). Universidad Autónoma de Querétaro, Querétaro, Querétaro.
- Montenegro de, L. (1988). El discurso escrito de los escolares dominicanos: diagnóstico y bases para la planificación. *Ciencia y Sociedad*, 13(4), 540-557.
- Montenegro de, L. (1991). La investigación de la madurez sintáctica y la enseñanza de la lengua materna. En H. López (Ed.), *La enseñanza del español como lengua materna: Actas del II seminario internacional sobre aportes de la lingüística a la enseñanza del español como lengua materna* (pp. 113-131). Río Piedras: Universidad de Puerto Rico.
- Rodríguez, L. (1991). Índices de madurez sintáctica en escolares puertorriqueños de escuela primaria. En H. López (Ed.), *La enseñanza del español como lengua materna: Actas del II seminario internacional sobre aportes de la lingüística a la enseñanza del español como lengua materna* (pp. 133-143). Río Piedras: Universidad de Puerto Rico.
- Salvador, F. (1985). Los índices de complejidad sintáctica, instrumentos de evaluación de la expresión escrita: estudio experimental en el ciclo medio de E. G. B. *Enseñanza & Teaching: Revista Interuniversitaria de Didáctica*, 3, 59-81.

- Salvador, F. (1995). Estructuras sintácticas de subordinación en la composición escrita: análisis evolutivo y diferencial. *Enseñanza & Teaching: Revista Interuniversitaria de Didáctica*, 181-197.
- Sánchez, V. (1998). Madurez sintáctica en niños preescolares costarricenses. *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 24(1), 113-120.
- Torres, A. (1993). *Madurez sintáctica en estudiantes no universitarios de la zona metropolitana de Tenerife* (Tesis doctoral). Universidad de la Laguna, Santa Cruz de Tenerife.
- Torres, A. (1997). Complejidad sintáctica en escritores de las Islas Canarias. *Revista de Estudios de Adquisición de la Lengua Española*, (8), 105-125.
- Torres, M. (2001). Índices primarios de madurez sintáctica: análisis comparativo en alumnos de COU. *Revista de Estudios de Adquisición de la Lengua Española*, (12), 93-118.
- Trelles, N. (2015). Diagnóstico sobre el desarrollo sintáctico de los alumnos de quinto de secundaria que ingresan a la Universidad de Piura-Perú. *SIGNOS ELE: Revista de Español como Lengua Extranjera*.
- Véliz, M. (1988). Evaluación de la madurez sintáctica en el discurso escrito. *RLA: Revista de Lingüística Teórica y Aplicada*, 26, 105-141.
- Véliz, M. (1999). Complejidad sintáctica y modo del discurso. *Estudios Filológicos*, (34), 181-192.
- Véliz, M., Muñoz, G., y Echeverría, M. (1985). Madurez sintáctica y combinación de oraciones en estudiantes universitarios. *RLA: Revista de Lingüística Teórica y Aplicada*, 23, 107-120.