



## ¿Cómo crece la sintaxis en los años escolares? Un estudio de la frase nominal en textos explicativos escritos por estudiantes secundarios

### How does syntax grow during school years? A study of the noun phrase in explanatory texts written by high school students

Recibido: 20-09-2022 Aceptado: 12-02-2023 Publicado: 30-06-2023

María Soledad Aravena Reyes

Pontificia Universidad Católica de Chile  
maravena@uc.cl

 0000-0001-8646-9755

Riva Quiroga Moya

Pontificia Universidad Católica de Chile  
riva.quiroga@uc.cl

 0000-0002-1147-4135

**Resumen:** Para examinar el desarrollo sintáctico durante la adolescencia, el presente trabajo se propuso estudiar las frases nominales en textos explicativos escritos por 80 estudiantes de séptimo básico y tercero medio pertenecientes a colegios de dos grupos socioeconómicos, NS alto y medio-bajo. Se analizó la frecuencia, densidad, diversidad léxica, función sintáctica y complejidad de las frases nominales y se aplicaron pruebas de varianza para determinar el efecto de las variables independientes (edad/nivel escolar y NS). Los resultados muestran que la edad / nivel escolar es el factor con mayor incidencia en los desempeños sintácticos observados, con un efecto particularmente significativo en la densidad sintáctica de las frases y su diversidad léxica. Se concluye que después de los 12 años los usos sintácticos experimentan importantes cambios y ajustes necesarios para alcanzar un dominio más especializado del lenguaje y de los géneros discursivos.

**Palabras clave:** sintaxis – desarrollo – frase nominal – adolescencia – texto explicativo

Citación: Aravena, M. y Quiroga, R. (2023). ¿Cómo crece la sintaxis en los años escolares? Un estudio de la frase nominal en textos explicativos escritos por estudiantes secundarios. *Logos: Revista de Lingüística, Filosofía y Literatura*, 33(1), 20-39. doi.org/10.15443/RL3302



**Abstract:** To examine syntactic development during adolescence, the present study aimed to study noun phrases in explanatory texts written by 80 high school students belonging to two socioeconomic groups, high and medium-low SES. The frequency, density, lexical diversity, syntactic function, and complexity of nominal phrases were analyzed, and variance tests were applied to determine the effect of the independent variables (age/school level and SES). Results show that age/school level is the factor with the highest incidence in the observed syntactic performances, with a particularly significant effect on syntactic density of sentences and their lexical diversity. It is concluded that after the age of 12, syntactic uses undergo important changes and adjustments necessary to achieve a more specialized domain of language and discourse genres.

**Keywords:** syntax – development – noun phrase – adolescence – explanatory text

## Introducción

Durante los años escolares, el lenguaje de los niños experimenta cambios sustanciales relacionados con diversos factores, tales como el desarrollo cognitivo, la inmersión en la cultura letrada y el creciente dominio de los géneros discursivos (Tolchinsky, 2004). Para responder a las demandas de las nuevas situaciones comunicativas tanto orales como escritas, las estructuras sintácticas se vuelven paulatinamente más extensas, variadas y complejas (Diessel, 2004).

El desarrollo de la sintaxis compleja se ha investigado a menudo en relación con la construcción oracional, es decir, con el modo en que los niños aprenden a estructurar y a combinar las cláusulas (Clark, 2009; Aravena & Hugo, 2016). Sin embargo, se conoce mucho menos acerca del comportamiento evolutivo de los elementos del nivel frástico, y en particular de las frases nominales (FN). Al respecto, algunos estudios en otras lenguas reportan que estos grupos de palabras proporcionan información muy completa de los crecimientos lingüísticos y sintácticos durante los años escolares (Ravid & Berman, 2010). En español, en cambio, es un tema todavía poco investigado. Además, los escasos estudios sobre desarrollo de FN con población hispanohablante se concentran en las etapas tempranas de adquisición y/o en niños con desarrollo atípico (Pavez, Coloma, Araya, Maggiolo & Peñaloza, 2015). En Chile, de hecho, no se conocen trabajos que analicen este aspecto del desarrollo sintáctico en niños y adolescentes con desarrollo típico del lenguaje y, mucho menos, que comparen perfiles evolutivos en dos tipos de población.

En este marco, el presente estudio se propone observar el desarrollo sintáctico durante la adolescencia, mediante el análisis de las frases nominales en textos explicativos escritos por estudiantes de séptimo básico y tercero medio pertenecientes a dos grupos socioeconómicos. Para ello, se analizan cinco dimensiones de las FN por nivel

escolar y socioeconómico, en particular: la frecuencia de las FN, su densidad sintáctica (extensión), su diversidad léxica, la función sintáctica que cumplen en la oración y la complejidad de su arquitectura interna.

## La frase nominal

En el presente estudio se analizan los grupos de palabras encabezados por sustantivos, es decir, las unidades léxicas de clase abierta que sirven para denotar entidades de todo tipo y naturaleza (Bosque, 1999; Picallo, 1999). Se incluyen en esta categoría también las nominalizaciones, que son sustantivos derivados de adjetivos o verbos, como [mi *búsqueda* de capítulo nuevos]; infinitivos y adjetivos sustantivados, como [el *actuar* de su hija]; y expresiones nominales que funcionan como una sola pieza léxica, por ejemplo, *red social* o *página web*.<sup>1</sup> El sustantivo es una unidad léxica central del lenguaje y su estudio evolutivo es de la mayor relevancia, no solo por su aporte en la configuración del repertorio léxico y semántico de los hablantes, sino también por su contribución al desarrollo de la sintaxis compleja que se despliega en la frase nominal (Ravid, 2006).

La FN es una unidad sintáctica argumental de todo predicado que se construye en torno a un sustantivo como núcleo, y comparte con este las mismas propiedades semánticas y posibilidades funcionales. Alternativamente, diversos tipos de pronombres también pueden ser elementos nucleares de una FN (Dixon, 2021; Rigau, 1999). Esta unidad, intermedia entre la palabra y la oración, se caracteriza por un amplio rango de posibilidades en cuanto a la diversidad de sus componentes internos (Ravid & Berman, 2010). Puede, por ejemplo, presentarse como una estructura básica conformada por un solo sustantivo, como en el ejemplo *falta [tinta]*; un núcleo nominal más su determinante, como en *compré [un repuesto]*; o un núcleo más algún elemento modificador que le atribuye propiedades al sustantivo, como en *compré [un repuesto barato]*. En el otro extremo del rango de posibles componentes, también puede presentarse una estructura como las anteriores, complementadas además con frases adjetivas o preposicionales como en *compré [un repuesto [de tinta]]*; cláusulas relativas, como en *compré [un repuesto [de tinta] [que instalaré hoy]]*; y la adición y combinación de varias de estas estructuras como en *compré [un repuesto [de tinta]] [muy barato] [para la impresora] [que instalaré hoy]]*. Además, como será explicado en la metodología, se analizan también los grupos nominales encabezados por ciertos pronombres como *este*, *algunos*, *aquello*; elementos cuantificadores y numerales como *uno*, *ciertos*, *todos*; determinantes neutros como *lo* en [*lo más divertido*] *vino después* y núcleos elípticos.

## Desarrollo sintáctico y escritura de texto expositivo

En su estructura básica, los componentes de las frases nominales son definidos y su disposición responde a un orden fijo, de tal modo que en un principio el hablante

no tiene muchas opciones para alterar el orden de los componentes o insertar estructuras entre ellos (Ravid & Berman, 2010). Esta cierta obligatoriedad de base y por consiguiente, su relativa baja demanda cognitiva, hace que las frases nominales comiencen a manifestarse muy pronto en la ontogenia, hacia los 3 – 4 años (Radford, 1995). Luego, durante la primera infancia, las unidades sintácticas crecen en frecuencia, extensión y complejidad, impulsadas por la experiencia acumulada de los niños con nuevos y diversos contextos discursivos que van más allá de la conversación cotidiana. Durante los años escolares, y particularmente después de los 10 – 12 años, la sintaxis constituye una herramienta fundamental para las crecientes demandas de las prácticas de lectura y escritura propias del lenguaje de la escuela (Scott, 2004).

En la oralidad, los hablantes acceden con mayor facilidad a las entidades del discurso en el contexto extralingüístico, pero a la vez enfrentan la dificultad de tener que articular *en línea* un discurso conectado (Halliday, 1989). En cambio, los textos escritos otorgan el tiempo necesario para la planificación y la revisión, pero exigen la selección de información precisa y la construcción de estructuras sintácticas y textuales que garanticen la introducción y el acceso a la referencia intralingüística (Quiroz & Martín, 2021). Esta tarea impone a los sujetos demandas cognitivas sobre la memoria, las funciones ejecutivas y el procesamiento de tipo *top-down*, que no están listas antes de la adolescencia (Ravid, 2013). En esta etapa y hasta la edad adulta, el uso de diversas estructuras sintácticas extensas y complejas hace posible mantener el flujo de la información en el discurso, controlando la construcción de niveles oracionales y frásticos, y a la vez considerando el texto como una unidad integrada. Por ello, los textos escritos –y particularmente los expositivos– con su alta frecuencia de palabras de contenido léxico y de estructuras jerárquicas encabezadas por sustantivos (Scott, 2010; Halliday, 1989), constituyen unidades textuales idóneas para el análisis de la sintaxis compleja.

A diferencia de las narraciones, que se centran en entidades concretas y acciones, los textos expositivos y/o explicativos suelen construirse en torno a contenidos abstractos tales como conceptos, ideas y procesos (Scott, 2004; Nippold, Mansfield, Billow & Tomblin, 2008). Estos textos, típicos del lenguaje de la escuela (Christie, 2012), requieren una organización jerárquica de sus tópicos y un grado mayor de explicitación en la transmisión de los saberes, por lo que concentran elementos nominales que contribuyen a configurar su tendencia sintetizadora (Soto & Zenteno 2004). En efecto, las FN tienen, por un lado, gran capacidad de *empaquetar* información y concentrar densidad informativa y, por el otro, son construcciones sintácticas *expandibles* recursivamente. Ambas –la densidad informativa y la complejidad sintáctica de la frase nominal– han sido relacionadas con la producción típica de estudiantes mayores (Nir & Berman, 2010) y con los usos propios de la escritura y del lenguaje académico (Biber & Gray, 2016; Parkinson & Musgrave, 2014).

## Metodología

### Selección de la muestra y participantes

Los resultados que se reportan en este trabajo provienen de una investigación mayor en la que participaron alrededor de 400 estudiantes de dos colegios de Santiago, uno subvencionado y el otro particular pagado, representativos de los grupos socioeconómico alto (NSA) y medio-bajo (NSM), según datos de Ministerio de Educación. Ambos colegios son mixtos y sus estudiantes obtienen resultados que están en el promedio de su tipo de dependencia y comuna en las pruebas estandarizadas nacionales. Se les aplicaron dos tareas de escritura (una narrativa y otra explicativa, que es objeto del presente estudio) que fueron revisadas con rúbricas elaboradas y validadas en el proceso y que sirvieron para establecer puntajes de desempeño global. Considerando este puntaje, se excluyeron de la muestra los casos atípicos mediante cálculo de puntaje estándar (puntaje Z) y se redujo la muestra tras un análisis de potencia estadística calculado en G\*Power (Faul, Erdfelder, Lang & Buchner, 2007). De esta manera, se seleccionaron 60 estudiantes por colegio distribuidos en tres niveles y equiparados por sexo, aunque, para el presente estudio, se consideraron solo dos niveles escolares, por haber sido identificados como los más relevantes para los fenómenos evolutivos que interesa observar: séptimo básico y de tercero medio.

### Recolección de datos y tratamiento del corpus

Tras la aplicación de una encuesta de intereses, se seleccionó la temática de “Internet y redes sociales”, con la finalidad de ofrecer un tópico de escritura que fuera atractivo pero también viable para la escritura tanto narrativa como explicativa. Y en esta última es particularmente necesario escribir sobre un tópico del cual se tengan suficientes conocimientos (Nippold, 2010). Tras la observación de un material audiovisual sin texto sobre el tema elegido, se pidió a los estudiantes que escribieran un texto en que presentaran un sitio o recurso de Internet y explicaran su uso, pensando en un lector que no lo conoce bien y quiere aprender a usarlo.

Los 80 textos resultantes de esta tarea son de carácter *explicativo*, aunque diversos en su organización y también de muy variada extensión, quizás debido a la temática y a las grandes diferencias que mostraron los estudiantes en el dominio de este género. Los textos fueron segmentados en cláusulas para realizar diversos análisis léxicos y sintácticos. Posteriormente fueron transcritos en CHAT, formato de texto plano que sirvió para luego hacer diversos etiquetados y conteos sintácticos y léxicos con el programa CLAN (MacWhinney, 2020). La segmentación y transcripción se hizo cautelando constantemente la calibración entre transcriptores, de acuerdo con lo

que se recomienda para lograr niveles adecuados de consistencia en este tipo de procedimientos (Stemler, 2004).

## Variables y procedimientos de análisis

Las variables independientes del estudio son la edad o nivel escolar, porque este factor se asocia con distintos grados de experiencia letrada y lingüística en general; y el nivel socioeconómico (NS) porque este factor se relaciona en Chile con el grado de alfabetización familiar y el tipo de establecimiento y comuna.

Las variables dependientes son las siguientes cinco medidas relacionadas con las frases nominales (que se explicarán con detalle más adelante): frecuencia, densidad sintáctica, diversidad léxica, función sintáctica y complejidad. Como se ha explicado en la introducción, se incluyeron en el análisis todas las FN encabezadas por un sustantivo (o nominalización o expresión sustantivada) y también las de tipo pronominal con demostrativos como *este, algunos, aquello*; elementos cuantificadores y numerales como *uno, ciertos, todos* y pronombres neutros como *lo*. Las únicas FN que no se contabilizan en este análisis son los pronombres personales (*yo, tú, él/ella, nosotros, ustedes, ellos*) y los pronombres complementarios dativos o acusativos como *lo (lo escribió) la, le, les, se (se lo mandó)*. También quedan fuera del análisis, naturalmente, las locuciones verbales o adverbiales cuyo elemento nominal está gramaticalizado como: *darse cuenta, dar permiso, dar vergüenza, pasar(serme algo) por la cabeza, llamar la atención, etc.*; y entre las adverbiales, *por suerte, en el fondo, la verdad, a la perfección, al día siguiente, (quedar) en blanco, hoy en día, etc.* En esta línea, también hubo expresiones consideradas “intermedias” entre el contenido léxico pleno y un proceso de gramaticalización, que fueron analizadas caso a caso y fueron finalmente incorporadas en el análisis, tales como *con el tiempo, un día, una vez y otra vez*.

Para obtener la medida de frecuencia, se contaron todas las FN de *primer nivel jerárquico* de cada texto y se calculó su proporción en relación con el total de cláusulas del texto, medida utilizada como referencia de extensión. Estas FN cumplen la función de sujeto gramatical o forman parte del predicado, por ejemplo, como complementos del verbo o atributos predicativos, en el caso de las oraciones copulativas. Y son de primer nivel jerárquico porque no constituyen unidades subordinadas al interior de otras frases nominales. Es decir, las FN que forman parte de los componentes internos, subordinados al núcleo o a otras unidades al interior de una FN, no se contaron en el análisis de frecuencia. Así, en el siguiente ejemplo, las cinco FN delimitadas entre corchetes se contabilizaron para el cálculo de la frecuencia, pero no entraron en este cálculo las FN internas / subordinadas – complementos del nombre y cláusulas relativas, destacadas en cursiva – porque estas se cuentan en el análisis de la complejidad interna, como se explicará.

“Creo que [el uso *de Facebook*] se presta para [muchas cosas delicadas]. Lo digo porque [todas las personas] tienen acceso a [la información *de tus contactos*] y a veces incluso a [cierta información delicada *que puede perjudicar a la persona involucrada por otros en un lío sin tener idea*].”

En cuanto a las demás medidas, la densidad sintáctica da cuenta de la extensión de las FN calculada en función del número de palabras que contienen. La diversidad léxica, aunque no es una medida estrictamente sintáctica, se considera interesante en el análisis de las frases nominales, puesto que el desarrollo de la sintaxis compleja se relaciona con el desarrollo léxico (Ravid & Berman, 2010). Esta medida alude a la habilidad para utilizar palabras distintas y evitar la repetición (Read, 2000) y se mide a través de la fórmula TTR (*type-token ratio*), que es la proporción entre las palabras distintas (*types*) y el total de palabras (*tokens*) (en este caso, son las palabras de las FN, no de todo el texto). Dado que TTR solo puede aplicarse en textos de igual extensión, en este análisis se utiliza CTTR, una fórmula corregida y apta para trabajar con textos de diversa extensión (Malvern, Richards, Chipere & Durán, 2004). La función sintáctica, en tanto, distingue las FN que son sujeto gramatical de la oración, de aquellas que cumplen otras funciones al interior del predicado o fuera de la oración (por ejemplo, títulos, complementos del verbo, complementos del adverbio o atributos predicativos). Por último, la complejidad mide los tipos de arquitectura interna de las FN, considerando la naturaleza jerárquica de sus componentes.

Para el análisis de la complejidad, se utilizaron dos medidas. Con la primera se cuentan todos los sustantivos y verbos subordinados al interior de las FN del texto (ya sean subordinados directamente al núcleo o subordinados a otras estructuras a su vez subordinadas al núcleo). La segunda atañe a las FN de primer nivel, que fueron clasificadas en tres tipos o grados de complejidad interna, siendo las del Tipo 1 las más simples y las de Tipo 2 y 3 más complejas por sus elementos frásticos u oracionales. Estos tres tipos de complejidad y el modo en que se contaron los sustantivos y verbos subordinados se ilustran en los ejemplos de la Tabla 1:

**Tabla 1.** Análisis de la complejidad interna de FN

Tipo	Descripción	Ejemplos	Susts.	Vrbs.
			subs.	subs.
1	Núcleo (N) o N + determinante (artículo, numeral, posesivo, etc.) y eventualmente un adjetivo.	falta [tinta]	0	0
		[todas las personas]	0	0
		[muchas cosas delicadas]	0	0
2	La estructura anterior o N + un componente frástico que modifica al núcleo (como frase adjetiva o preposicional).	[Un repuesto de tinta]	1	0
		[Un repuesto de tinta para la impresora]	2	0
		[la información de tus contactos]	1	0
3	Estructuras anteriores o solo N + un componente oracional que modifica al núcleo o a otra de sus estructuras subordinadas.	[cierta información delicada que puede perjudicar a la persona involucrada por otros en un lío sin tener idea]	2	3

## Procesamiento de datos

Para el análisis de los datos se aplicó un diseño de ANOVA factorial mixto con dos factores intergrupo: edad/nivel escolar (séptimo básico y tercero medio) y grupo socioeconómico (grupo medio-bajo o NSM y grupo alto o NSA). Todas las pruebas de varianza se hicieron con R (R Core Team, 2022). El análisis de los tipos de palabras contenidas en las FN se hizo con el programa UDPipe (Straka, 2018) y la diversidad léxica (CTTR) de las FN se obtuvo con el programa Quanteda de R (Benoit et al., 2018).

## Resultados

### Resultados generales

Los resultados de esta investigación muestran que casi todas las medidas analizadas se ven impactadas de manera significativa por el factor edad / nivel escolar, y algunas se ven afectadas por el factor socioeconómico (NS). La Tabla 2 presenta los estadísticos descriptivos de todas las medidas y sirve como referencia para los análisis de varianza posteriores. Asimismo, se incluyen en la tabla datos que no serán comentados pero que sirven de base para obtener los resultados. Son los primeros tres datos de la tabla, todos de frecuencias absolutas: número total de cláusulas del texto, número de FN del texto (las FN de primer nivel jerárquico, como se explicó) y número total de palabras

contenidas en las FN por texto. Más abajo, se despliegan nueve medidas para cubrir las cinco dimensiones que sí son el objeto específico de este trabajo (frecuencia, densidad sintáctica, diversidad léxica, función y complejidad). Todas estas medidas se expresan en frecuencias relativas, es decir, se calculan en relación con el total de FN de los textos o en relación con sus cláusulas.

**Tabla 2.** Descriptivos de todas las medidas analizadas (medias grupales y DE por nivel escolar y NSE)

		NS medio-bajo (NSM)				NS alto (NSA)			
		Media	DE	Mín.	Máx.	Media	DE	Mín.	Máx.
Total cláusulas en texto	Séptimo	27,3	13,02	12	52	35,8	11,02	17	61
	Tercero	38,9	19,79	13	90	60,1	19,63	28	92
Total FN en texto	Séptimo	20,9	9,62	8	40	24,9	9,82	13	42
	Tercero	26,3	11,79	10	58	48,5	17,18	21	80
Palabras en FN	Séptimo	74,2	33,06	33	156	92,8	33,88	43	176
	Tercero	120,5	43,68	55	231	237,1	84,96	106	390
Frecuencia FN <sup>1</sup>	Séptimo	0,79	0,17	0,48	1,08	0,69	0,17	0,47	1,03
	Tercero	0,71	0,17	0,49	1,17	0,83	0,21	0,40	1,20
Densidad FN <sup>2</sup>	Séptimo	3,67	0,90	2,24	5,94	3,90	1,08	2,60	6,57
	Tercero	4,90	1,52	2,88	9,79	5,08	1,52	2,76	8,81
Diversidad léxica FN <sup>3</sup>	Séptimo	4,04	0,57	3,10	4,92	4,30	0,46	3,06	5,14
	Tercero	4,99	0,54	3,81	5,99	5,93	0,67	4,19	6,91
FN función sujeto <sup>4</sup>	Séptimo	0,17	0,10	0	0,33	0,20	0,09	0,07	0,39
	Tercero	0,24	0,11	0,10	0,50	0,31	0,08	0,11	0,43
Complejidad Tipo 1 <sup>5</sup>	Séptimo	0,76	0,11	0,50	0,94	0,70	0,08	0,50	0,85
	Tercero	0,67	0,12	0,40	0,84	0,64	0,09	0,48	0,81
Complejidad Tipo 2 <sup>5</sup>	Séptimo	0,08	0,07	0	0,23	0,11	0,08	0	0,28
	Tercero	0,11	0,08	0	0,30	0,17	0,05	0,05	0,29
Complejidad Tipo 3 <sup>5</sup>	Séptimo	0,15	0,07	0	0,33	0,19	0,08	0,07	0,38
	Tercero	0,22	0,09	0,07	0,36	0,20	0,09	0,09	0,37
Sustantivos subordinados <sup>6</sup>	Séptimo	0,43	0,33	0,06	1,38	0,40	0,26	0,13	0,19
	Tercero	0,53	0,41	0,12	1,86	0,60	0,29	0,19	1,31
Verbos subordinados <sup>7</sup>	Séptimo	0,25	0,16	0	0,67	0,29	0,17	0,10	0,70
	Tercero	0,37	0,18	0,11	0,79	0,35	0,20	0,12	0,80

<sup>1</sup> Proporción de FN sobre el total de cláusulas del texto.

<sup>2</sup> Proporción de palabras en FN sobre el total de FN del texto.

<sup>3</sup> Índice de diversidad léxica (calculado con CTTR) de las FN del texto.

<sup>4</sup> Proporción de las FN en función sujeto sobre el total de FN del texto.

<sup>5</sup> Proporción de las FN de complejidad 1, 2 y 3, sobre el total de FN del texto.

<sup>6</sup> Proporción de sustantivos subordinados al interior de las FN sobre el total de FN del texto.

<sup>7</sup> Proporción de verbos subordinados al interior de la FN sobre el total de FN del texto.

## Frecuencia y densidad sintáctica de frases nominales

La medida de frecuencia alude al número de FN de primer nivel jerárquico en relación con el total de cláusulas del texto, unidad que se ha utilizado como medida de extensión. Esta medida arroja un resultado, en cierto sentido, inesperado, y se comporta de manera distinta en los dos establecimientos. Tal como se reporta en la Tabla 2, en el colegio de NSM, la frecuencia es un poco más alta en séptimo (0,79) y luego baja levemente en tercero medio (0,71), mientras que en el colegio de NSA es algo más baja en séptimo (0,69) y sube un poco en tercero medio (0,83), aunque ninguna de estas diferencias por nivel escolar resulta estadísticamente significativa. En otras palabras, la frecuencia relativa de FN de primer nivel jerárquico es similar en todos los estudiantes de la muestra, de ambos niveles escolares y establecimientos. Una explicación de este resultado se relaciona con la extensión de los textos. Si bien lo esperable sería un aumento en el uso de FN entre séptimo y tercero medio, cuando los textos son muy breves, como ocurre en algunos casos en el grupo menor del NSM, entonces la cantidad de FN resulta alta en proporción a la extensión del texto. Y la escasa extensión de los textos en este grupo se relaciona presumiblemente, con un menor dominio del género discursivo, lo que resulta en textos muy básicos, a veces casi de estilo enumerativo (situación que no se observa en los textos narrativos de estos mismos estudiantes, pero que no son objeto del presente análisis). En la Tabla 3, más abajo, se presenta un ejemplo del fenómeno descrito.

La densidad, en tanto, alude a la extensión de las FN de primer nivel jerárquico en cantidad de palabras. En esta medida, tal como se observa al comparar las medias de la Tabla 2, la densidad aumenta de séptimo a tercero medio: de 3,67 a 4,9 palabras por FN en el NSM y de 3,90 a 5,08 palabras por FN en el NSA. La prueba ANOVA para el análisis de varianza muestra que se trata de diferencias estadísticamente significativas ( $F(1,76) = 17,135$ ;  $p = ,0000$ ;  $\eta^2_G = 0,184$ ) y las comparaciones *post-hoc* hechas con el test de Tukey señalan que la diferencia por nivel escolar es significativa tanto en el NSM ( $p = ,0183$ ) como en el NSA ( $p = ,0285$ ). Este resultado se traduce en que, a los 16 años, los estudiantes escriben textos explicativos con FN mucho más extensas, lo que les permite *empaquetar* una mayor cantidad de información en una sola frase; sin embargo, también hay casos en los que la extensión es excesiva porque se pierde el control de los elementos subordinados, como ocurre a veces en el Ejemplo 2 de la Tabla 3, comentado más abajo. A modo de ilustración, considérese una de las FN de este texto:

[diferentes modos de juego en el cual el jugador controla a un personaje, el cual se puede ver tanto en tercera como primera persona, y el cual es situado en un terreno creado al azar]

En los ejemplos siguientes (Tabla 3) se incluyen las FN de dos textos que representan muy bien las diferencias respecto de la frecuencia y la densidad de las FN entre séptimo básico y tercero medio. El primero (Ejemplo 1) es breve (tiene solo 19 cláusulas en

total) y contiene 15 frases nominales de primer nivel jerárquico. Esto resulta en una alta frecuencia de FN (0,79, idéntica a la media de su grupo). Si se considera la extensión de las FN, sin embargo, se puede observar a simple vista que la mayoría tiene entre 1 y 4 palabras, lo que resulta en una densidad baja (3,4, un poco por debajo de la media de su grupo, que es 3,67). El segundo ejemplo, en cambio, tiene un total de 44 FN de primer nivel en un texto de 74 cláusulas, lo que da una proporción de 0,59 FN por cláusula (inferior al texto de séptimo). Sin embargo, si se considera la extensión de las FN, se constata que muchas de ellas tienen más de 4 palabras, y algunas llegan a tener más de 10. Esto resulta en una densidad muy alta (7,52, superior a la media de su grupo, que es de 5,08).

**Tabla 3.** Ejemplos de análisis (incluyen todas las FN de cada texto)

Ejemplo 1. Séptimo, NSM	Ejemplo 2. Tercero, NSA
<p>GoogleGoogle  una página donde se puede buscar  información de todo tipo  esta página  lo que quieres buscar  imágenes  la parte superior izquierda  imagen  imágenes de lo que estás buscando  esta página  lo que quieres encontrar  Google  Google  información sobre tareas, trabajos, etc.  páginas de juegos o páginas de redes sociales  (1102115)</p>	<p>unos años atrás  los avances en la tecnología  una revolución en los juegos de video  estos  la realidad  una excepción de esto  un nuevo juego creado hace unos años atrás llamado Minecraft.  este juego  diferentes modos de juego en el cual el jugador controla a un personaje  el cual se puede ver tanto en tercera como primera persona, y el cual es  situado en un terreno creado al azar.  este terreno  el personaje  recursos  herramientas y estructuras las cuales son fruto de la imaginación del  jugador.  el aspecto más interesante de este juego  el hecho de que todo está formado por cubos que a escala representan  1m3  estos  distintos lugares  el jugador  las estructuras que estos deseen  estos mundos o terrenos  fin  el terreno ya investigado  los creadores de este juego  propiedades físicas a ciertos bloques como la arena el cual si es colocado  a cierta altura se cae hacia el suelo,  el agua  sus propiedades  terrenos más bajos.  bloques como la madera que pueden ser suspendidos en el aire.  dentro del juego  animales como gallinas, vacas y ovejas las cuales pueden ser criadas y  proveen al jugador de recursos  monstruos los cuales salen de noche o están presentes en partes oscuras  como cuevas los cuales atacarán al personaje el cual deberá defenderse  o huir.  estas cualidades del juego en el cual se llama a los jugadores a crear y  utilizar su imaginación en conjunto con aventurarse por terrenos que  nunca son iguales a los ya conocidos  un modo multijugador el cual con el uso de internet reúne a distintos  jugadores en servidores creados por estos mismos en los cuales  dependiendo de qué es lo que se quiera hacer, se juntan jugadores para  crear gigantes estructuras o competir entre estos.  este juego  cada 3 o 4 meses  la monotonía y rutina de los juegos  este increíble juego  factores de diversos juegos  el jugador  otros  el único factor que limita a las personas  su imaginación  un videojuego  buenas gráficas y violencia  (2314202)</p>

## Diversidad léxica

Tal como se ha explicado en el apartado metodológico, para calcular la diversidad léxica se usó la fórmula CTTR que permite calcular el TTR corrigiendo la distorsión derivada de las diversas extensiones de los textos (Malvern, et al., 2004). En este caso, el CTTR calculado es la proporción de palabras nuevas (*types*) de la FN por el total de palabras de la FN (*tokens*). En la Tabla 2 se observa que las medias grupales para esta variable aumentan (de 4,03 a 4,99 en el NSM y de 4,30 a 5,93 en el NSA), medias que, analizadas mediante pruebas de varianza resultan en diferencias estadísticamente significativas tanto por edad / nivel escolar ( $F(1,76) = 103,981$ ;  $p = ,0000$ ;  $\eta^2_G = 0,578$ ) como por NS ( $F(1,76) = 22,790$ ;  $p = ,0000$ ;  $\eta^2_G = 0,231$ ). Como puede constatarse, los tamaños de efecto son altos y, consecuentemente, las comparaciones *post-hoc* muestran que las diferencias por edad son significativas en ambos colegios ( $p = ,0000$  y  $p = ,0000$  en NSM y NSA respectivamente); además, la diferencia por colegio es significativa también en tercero medio ( $p = ,0000$ ); es decir, los dos terceros medios son significativamente diferentes, pero no lo son los séptimos ( $p = ,4610$ ). Aunque es difícil constatar el fenómeno de la diversidad léxica a simple vista, los ejemplos de la Tabla 3 contribuyen a ilustrarlo. El ejemplo 1 tiene una diversidad léxica de 3,27, un poco por debajo de la media grupal (4,04) y el ejemplo 2 tiene una diversidad de 6,49, un poco superior a la media de su grupo (5,93). Esto quiere decir que en el primer texto las FN contienen aproximadamente 3 palabras nuevas en relación con el total del palabras y en el segundo texto esta cifra se duplica.

## Función sintáctica: sujeto gramatical

El sujeto gramatical de una oración se formula nominalmente, pero la gran mayoría de las FN de todos los textos y en todos los grupos se ubican en el predicado, puesto que el sujeto de una oración está conformado por un argumento, mientras que las posibilidades gramaticales de las FN en el predicado son muchas más (varios tipos de complementos del verbo, atributos predicativos, complementos del adverbio, entre otras). Por esto, y porque el sujeto gramatical es la unidad que permite mantener los tópicos del texto (Ravid, van Hell, Rosado, & Zamora, 2002), es interesante observar el comportamiento evolutivo de las FN en función de sujeto. Tal como se informa en la Tabla 2, las FN en función sujeto aumentan de manera clara con la edad / nivel escolar tanto en el NSM (de 0,17 a 0,24) como en el NSA (de 0,20 a 0,31) y también aumentan por efecto del grupo socioeconómico. Consecuentemente, el análisis de varianza muestra significatividad estadística para el factor edad / nivel escolar  $F(1,76) = 18,114$ ;  $p = ,0000$ ;  $\eta^2_G = 0,192$ , y para el factor NS  $F(1,76) = 4,754$ ;  $p = ,0320$ ;  $\eta^2_G = 0,059$ . Las pruebas *post-hoc* indican que la diferencia mayor se da entre séptimo y tercero del NSA ( $p = ,0018$ ) lo que determina el resultado general de todos los grupos. No es posible “observar” la presencia de los sujetos gramaticales en los ejemplos de la Tabla 3 (porque se requeriría

comprobar la función sintáctica en el texto completo) pero los datos respecto de esta variable son los siguientes: el texto del ejemplo 1 tiene una proporción de 0,26 FN en función sujeto, mientras que el texto del ejemplo 2 tiene casi el doble, 0,40. Esto significa que los estudiantes mayores hacen más frecuentes y explícitos los sujetos, porque los necesitan para mantener el tema en textos donde se elabora una mayor cantidad de información.

## Complejidad sintáctica

### Sustantivos y verbos subordinados

La primera medida de complejidad consiste en cuantificar los sustantivos y verbos subordinados. Esta medida es importante porque entrega información complementaria de la complejidad interna de las FN, no ya exclusivamente de las de primer nivel jerárquico sino sobre todas las demás, donde también se constatan diferencias entre los textos de los menores y los mayores. En efecto, la proporción de sustantivos subordinados (SS) aumenta de 0,43 a 0,54 en el NSM y de 0,40 a 0,60 en el GSA y la prueba ANOVA indica que se trata de diferencias significativas por edad / nivel escolar ( $F(1,76) = 4,336$ ;  $p = ,041$ ;  $\eta^2_G = 0,054$ ) pero no por tipo de colegio. En cuanto a los verbos subordinados (VS), se trata de una medida que informa acerca de las cláusulas que se encadenan al interior de las FN, no solamente de las FN de primer nivel, como se ha explicado, sino de otras unidades tales como frases preposicionales o adjetivas. Como se constata en la Tabla 2, los VS son menos numerosos que los SS, pero igualmente muestran una tendencia evolutiva clara (de 0,24 a 0,37 en el NSM y de 0,29 a 0,36 en el NSA) y estadísticamente significativa ( $F(1,76) = 5,514$ ;  $p = ,021$ ;  $\eta^2_G = 0,068$ ). La manera en que los estudiantes de séptimo y tercero medio utilizan los sustantivos y verbos subordinados al interior de las FN se ilustra muy claramente en los ejemplos 1 y 2 de la Tabla 3.

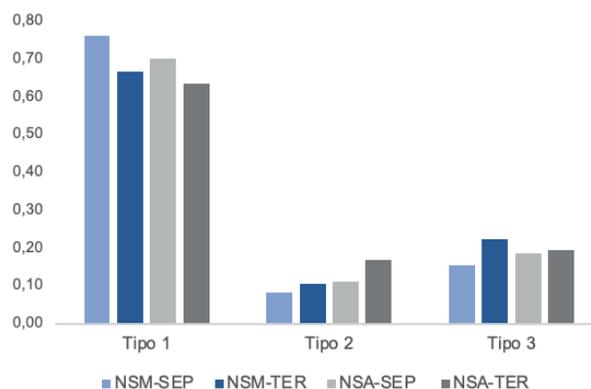
### Arquitectura interna

El segundo análisis de la arquitectura interna se hizo, tal como se explicó, clasificando las FN en tres tipos de acuerdo con su grado de complejidad. Los resultados de estos análisis muestran que las FN más simples (Tipo 1) son las más usadas por todos los grupos, tal como se ilustra en el Gráfico 1, más abajo. También se constata que estas FN son más frecuentes en los dos grupos de séptimo básico (0,76 y 0,70 en el NSM y el NSA respectivamente) y que su uso disminuye en ambos grupos de tercero medio (0,67 en el NSM y 0,64 en el alto). La comparación estadística de las medias revela que son significativas las diferencias por edad/nivel escolar ( $F(1,76) = 12,622$ ;  $p = ,0007$ ;  $\eta^2_G = 0,142$ ), y que también lo son –aunque en menor grado– las diferencias por nivel socioeconómico ( $F(1,76) = 4,176$ ;  $p = ,0440$ ;  $\eta^2_G = 0,052$ ). Las comparaciones *post-hoc* revelan que la diferencia se da entre los menores y los mayores del NSM ( $p = ,0229$ ), mientras que los dos niveles escolares del NSA no se diferencian estadísticamente ( $p = ,1630$ ).

Las FN de Tipo 2 (estructuras frásticas subordinadas), en tanto, son las menos utilizadas por todos los grupos, como se muestra en el Gráfico 1, y se comportan de modo inverso a las del Tipo 1. Es decir, su uso es menor en séptimo y mayor en tercero medio en ambos colegios (sube de 0,08 a 0,011 en el NSM y de 0,11 a 0,17 en el NSA). Se repite la tendencia en las pruebas de varianza: el factor edad / nivel escolar se revela significativo ( $F(1,76) = 6,729$ ;  $p = ,011$ ;  $\eta^2_G = 0,081$ ) al igual que el NSE ( $F(1,76) = 8,227$ ;  $p = ,005$ ;  $\eta^2_G = 0,098$ ).

Las FN de Tipo 3 (estructuras oracionales subordinadas) son un poco más frecuentes que las de Tipo 2 en todos los grupos, aunque su presencia es muy inferior a las de Tipo 1. Esta variable también aumenta de séptimo a tercero medio, pero lo hace de manera muy leve en el NSA (0,19 a 0,20) y de manera importante en el NSM (0,15 a 0,22). Consecuentemente, el efecto de la edad / nivel escolar vuelve a resultar significativo ( $F(1,76) = 4,807$ ;  $p = ,031$ ;  $\eta^2_G = 0,059$ ), mientras que el factor grupo socioeconómico no interviene. Al igual que ocurría con las FN de Tipo 1, las mayores diferencias se dan entre los estudiantes menores y los mayores del NSM, como muestran las comparaciones post-hoc ( $p = ,0380$ ), mientras que los estudiantes del NSA se comportan de modo similar (aun así, la gran diferencia del NSM hace que la diferencia por edad / nivel escolar sea significativa considerando los datos analizados conjuntamente).

En su conjunto, este análisis de la complejidad de las FN indica que los estudiantes de 16 años utilizan más FN complejas que los de 12 años, tanto las de tipo frástico como las de tipo oracional, situación que también se muestra claramente en los ejemplos de la Tabla 3.



**Gráfico 1.** Complejidad de las FN, comparación por grupos.

## Conclusiones

Con la finalidad de observar y comprender aspectos centrales del desarrollo sintáctico en edades escolares, el presente estudio se propuso conocer el modo en que los estudiantes de séptimo básico y tercero medio de dos grupos socioeconómicos utilizan las frases nominales en sus textos explicativos. Los resultados obtenidos permiten formular diversas conclusiones en relación con las nueve medidas y los dos factores analizados (edad/nivel escolar y NS). La conclusión primera y más evidente es la relevancia del factor edad / nivel escolar en los cambios sintácticos de esta etapa. En efecto, de las nueve medidas analizadas (para las cinco dimensiones de frecuencia, densidad sintáctica, diversidad léxica, función y complejidad) ocho se ven afectadas significativamente por este factor. Es decir, a los 16 años, debido a los cambios cognitivos y a la mayor experiencia letrada y discursiva, los estudiantes construyen frases nominales: más extensas, (con capacidad de condensar más información), más diversas léxicamente, con mayor preferencia por la función de sujeto y más complejas en su arquitectura interna (es decir, con más sustantivos y verbos subordinados, con menos estructuras simples –Tipo 1– y con más estructuras frásticas y oracionales –Tipo 2 y 3, respectivamente–). De estas ocho medidas hay dos en las que el efecto de la edad / nivel escolar es particularmente robusto: la densidad y la diversidad léxica. Por lo tanto, dado que en ambos colegios los estudiantes de tercero medio construyen FN más extensas y diversas léxicamente, estas resultan ser las medidas que mejor reflejan el desarrollo sintáctico en estas edades, al menos en lo que respecta al uso de FN en textos explicativos escritos.

La segunda conclusión es que el grupo socioeconómico tiene una incidencia menor que la edad en las construcciones sintácticas analizadas. Este factor interviene parcialmente en las medidas de diversidad léxica, función sujeto y complejidad (Tipos 1 y 2) y afecta principalmente a los estudiantes mayores. Así, los estudiantes de tercero medio del NS medio-bajo, comparados con sus pares del NS alto, presentan en sus FN menor diversidad léxica y utilizan menos FN en función sujeto y menos FN de complejidad de Tipo 2. Por otro lado, los estudiantes de séptimo básico del NSM presentan una proporción mayor de FN de Tipo 1 que sus pares del NSA. De manera general, por lo tanto, se puede afirmar que los dos grupos socioeconómicos son similares en sus desempeños sintácticos a los 12 años. Sin embargo, a los 16 años, aparecen diferencias que revelan un mayor dominio de algunos recursos sintácticos en la escritura de textos explicativos en el NSA. Tales diferencias se expresan en dos maneras de gestionar la referencia y la complejidad de las FN: los estudiantes del NSM usan menos sujetos explícitos para la mantención de los temas y prefieren el uso de cláusulas relativas en las FN complejas; en cambio los del NSA hacen más explícita la mantención del tema en el sujeto gramatical y prefieren el uso de frases adjetivas y preposicionales, ambas, estrategias que favorecen el carácter sintetizador de las frases nominales en la transmisión de conocimientos. Dadas las características de los colegios, es probable que esta ventaja del NSA se relacione con

mayores prácticas letradas en el contexto personal y escolar y una mayor familiaridad con el género y el lenguaje de la escuela (Christie, 2012). Por otro lado, se requieren más estudios acerca de las cláusulas relativas al interior de las FN, para indagar los límites de esta estrategia de subordinación a veces excesiva. En cualquier caso, estos resultados corroboran que la práctica con estos géneros discursivos favorece el uso y desarrollo de determinados recursos lingüísticos, sintácticos y también léxicos, que no se movilizan en la comunicación cotidiana ni en el uso exclusivo de textos narrativos, como se ha encontrado (Nippold, 2010; Scott, 2004).

Respecto de la frecuencia de FN, única variable que no se vio afectada por ninguno de los factores, es interesante constatar que parece haber una cantidad básica de FN de primer nivel jerárquico que se mantiene relativamente estable en los textos de las distintas edades y grupos. Es, posiblemente, la cantidad de FN necesaria para la construcción de un texto de estas características, cualquiera sea la edad o las estrategias desplegadas. Sería interesante comprobar si esta cierta estabilidad de FN se manifiesta del mismo modo en la escritura de otros géneros (textos narrativos, por ejemplo). Por otro lado, y vistos los resultados respecto de sustantivos subordinados, quizás se encontrarían importantes diferencias intergrupales si se contabilizaran todas las FN (no solo las de primer nivel, sino también todas las que están contenidas al interior de otras FN). Esto sería consistente con las conclusiones formuladas más arriba, en cuanto a que la densidad sintáctica de las frases y también sus tipos de composición interna se manifiestan con un claro perfil evolutivo.

Tras los resultados de este estudio, se abren nuevas interrogantes que sería interesante abordar en una investigación posterior. Debido a la relevancia del sujeto gramatical en el desarrollo de la gestión referencial de los textos (Hickmann, 2003; Ravid et al., 2002) y, dado que en este estudio se comprobó una tendencia evolutiva de las FN con esta función, sería conveniente estudiar los sujetos con más detenimiento. Por ejemplo, estudiarlos comparando el uso de formas pronominales y formas léxicas; y analizando también el uso de formas genéricas (como el pronombre “uno” o el sintagma “la gente”) y las formas de referencia más específica, particularmente en relación con los rasgos del género discursivo. En segundo lugar, en el análisis del corpus se encontraron fenómenos que, aunque se contabilizaron, no fueron analizados de forma particular por razones de espacio, pero que sin duda requieren una nueva mirada. Este es el caso, por ejemplo, de las habituales FN encabezadas por un pronombre neutro y con cláusulas subordinadas del tipo “*lo que me molestó*”, estructuras que cumplen funciones sintácticas pero también referenciales y de mantención del tópico. Por último, y dada la contundencia del resultado en diversidad léxica, sería conveniente explorar de manera más detenida la relación entre algunos rasgos semánticos de los sustantivos (como el grado de abstracción o de sofisticación, por ejemplo) y la complejidad de las estructuras sintácticas, porque sin duda, ambas dimensiones del desarrollo lingüístico están estrechamente vinculadas.

El presente estudio permitió corroborar que al término de la enseñanza básica y durante los años de enseñanza media, se observan importantes crecimientos sintácticos relacionados con el uso de las frases nominales. Estos cambios son necesarios para los nuevos desafíos comunicativos de esta etapa, en que los estudiantes deben enfrentar una gran cantidad de nuevos conocimientos que requieren a su vez un mayor dominio de los géneros escolares y un manejo específico de sus recursos lingüísticos. Una práctica frecuente con estos géneros vinculada a una enseñanza más explícita de sus recursos sintácticos puede contribuir a mejorar notablemente los desempeños en lectura y escritura en el contexto escolar.

## Referencias bibliográficas

- Aravena, S., & Hugo, E. (2016). Desarrollo de la complejidad sintáctica en textos narrativos y explicativos escritos por estudiantes secundarios. *Lenguas Modernas*, 47, 9-40.
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S. & Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open-Source Software*, 3(30), 774. DOI: 10.21105/joss.00774
- Biber, D. & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.
- Bosque, I. (1999). El nombre común. En *Gramática descriptiva de la lengua española* (pp. 3-76). Espasa Calpe.
- Clark, E. (2009). *First Language Acquisition*. Cambridge University Press
- Christie, F. (2012). *Language education throughout the school years: A functional perspective*. Wiley-Blackwell.
- Diessel, H. (2004). *The acquisition of complex sentences* (Vol. 105). Cambridge University Press.
- Dixon, R. (2021). *The Essence of Linguistic Analysis. An Integrated Approach*. Brill.
- Halliday, M. A. K. (1989). *Spoken and written language*. Oxford University Press.
- Hickmann, M. (2003). *Children's discourse. Person, Space and Time across Languages*. Cambridge University Press.
- Faul, F., Erdfelder, E., Lang, A. & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39, 175-191. DOI: 10.3758/bf03193146
- MacWhinney, B. (2020). *The CHILDES project: Tools for analyzing talk*. <https://childes.talkbank.org/>

- Malvern, D., Richards, B., Chipere, N. & Durán, P. (2004) *Lexical diversity and language development. Quantification and Assessment*. Palgrave Macmillan.
- Nippold, M. A. (2010). Explaining Complex Matters: How Knowledge of a Domain Drives Language. In *Expository discourse in children, adolescents, and adults* (pp. 41 – 62) Psychology Press.
- Nippold, M., Mansfield, T., Billow, J. & Tomblin, B. (2008). Expository Discourse in Adolescents with Language Impairments: Examining Syntactic Development. *American Journal of Speech-Language Pathology*, 17, 4, 356 – 633. DOI: 10.1044/1058-0360(2008/07-0049)
- Nir, B. & Berman, R. (2010). Complex syntax as a window on contrastive rhetoric. *Journal of Pragmatics* 42.3, 744-765. DOI:10.1016/J.PRAGMA.2009.07.006
- Parkinson, J. & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for Academic Purposes students. *Journal of English for Academic Purposes*, 14, 48-59. DOI: 10.1016/j.jeap.2013.12.001
- Pavez, M. M., Coloma, C. J., Araya, C., Maggiolo, M., & Peñaloza, C. (2015). Gramaticalidad y complejidad en narración y conversación en niños con trastorno específico del lenguaje. *Revista de logopedia, foniatría y audiolgía*, 35(4), 150-158. DOI: 10.1016/j.rlfa.2015.07.004
- Picallo, C. (1999). La estructura del sintagma nominal. En *Gramática descriptiva de la lengua española* (pp. 363-393), Espasa Calpe.
- Quiroz, B. & Martin, J. (2021). Perfil sistémico-funcional del grupo nominal en español: estructura, funciones discursivas básicas y organización sistémica. *Estudios Filológicos*, (68), 123-151 doi.org/10.4067/S0071-17132021000200123
- R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Radford, A. (1995). Phrase Structure and Functional Categories. In *The Handbook of Child Language* (pp. 483 – 506), Blackwell.
- Ravid, D. (2006). Semantic development in textual contexts during the school years: Noun scale analyses. *Journal of Child Language*, 33(4), 791-821 DOI: <https://doi.org/10.1017/S0305000906007586>
- Ravid, D. (2013). Syntactic complexity in discourse production across different text types. *Across the Line of Speech and Writing Variation. Corpora and Language in Use-Proceedings*, 2, 51-66.
- Ravid, D. & Berman, R. (2010). Developing noun phrase complexity at school age: A text-embedded cross-linguistic analysis. *First Language*, 30(1), 3-26. <https://doi.org/10.1177/0142723709350531>

- Ravid, D., van Hell, J. G., Rosado, E., & Zamora, A. (2002). Subject NP patterning in the development of text production: Speech and writing. *Written Language & Literacy*, 5(1), 69-93. DOI: 10.1075/wll.5.1.04rav
- Read, J. (2000). *Assessing Vocabulary*. Cambridge University Press.
- Rigau, G. (1999) La estructura del sintagma nominal: los modificadores del nombre. En *Gramática descriptiva de la lengua española* (pp. 311 – 362), Espasa Calpe.
- Scott, C. (2004). Syntactic Contributions to Literacy Learning. In *Handbook of Language and Literacy*, (pp. 340 – 362), Guilford Press.
- Scott, C. M. (2010). Assessing Expository texts produced by school-age children and adolescents. In *Expository discourse in children, adolescents, and adults* (pp. 201-224). Psychology Press.
- Soto, G. & Zenteno, C. (2004). Los sintagmas nominales en textos científicos escritos en español. *ELUA. Estudios de Lingüística*, 18, 275-292. DOI: <https://doi.org/10.14198/ELUA2004.18.14>
- Stemler, S. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation* 9.1, 4, DOI:10.7275/96JP-XZ07
- Straka, M. (2018, October). UDPipe 2.0 prototype at CoNLL 2018 UD shared task. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, 197-207. 10.18653/v1/K18-2020
- Tolchinsky, L. (2004). The nature and scope of later language development. *Language development across childhood and adolescence*, (pp. 233-247), John Benjamins Publishing.