

Desarrollo de la complejidad léxica en dos géneros escritos por estudiantes de distintos grupos socioeconómicos

Development of Lexic Complexity in Two Genres Written by Students of Different Socioeconomic Groups

Soledad Aravena

Pontificia Universidad Católica de Chile
Chile

Riva Quiroga

Pontificia Universidad Católica de Chile
Chile

ONOMÁZEIN 42 (diciembre de 2018): 197-224

DOI: 10.7764/onomazein.42.03

ISSN: 0718-5758



Soledad Aravena: Departamento de Ciencias del Lenguaje, Pontificia Universidad Católica de Chile, Chile.

| Correo electrónico: maravena@uc.cl

Riva Quiroga: Departamento de Ciencias del Lenguaje, Pontificia Universidad Católica de Chile, Chile.

| Correo electrónico: rivaquiroga@gmail.com

Fecha de recepción: abril de 2017

Fecha de aceptación: septiembre de 2017

Resumen

El presente estudio se propuso conocer el modo en que algunas medidas de complejidad léxica se manifiestan en textos escritos por estudiantes de tres niveles de Enseñanza Media y dos grupos socioeconómicos. 240 textos narrativos y explicativos fueron elicitados mediante un breve video sobre los usos de Internet y las redes sociales y luego etiquetados con CLAN (MacWhinney, 2000) en tres tipos de medidas léxicas: diversidad, densidad, abstracción nominal y tipos de procesos. Para el análisis de datos se aplicaron Pruebas T y un diseño de ANOVA factorial mixto con dos factores intergrupales (edad/nivel escolar y grupo social). Los resultados muestran que los desempeños léxicos cambian significativamente por efecto de dos factores principales: la edad/nivel escolar y el género discursivo. Por otro lado, también hacen evidente que la complejidad léxica debe ser estudiada de acuerdo con diversas medidas y en variadas situaciones comunicativas. Las conclusiones del estudio reafirman, además, el importante papel del léxico en el desarrollo de las habilidades de escritura y el dominio de los géneros.

Palabras clave: complejidad léxica; escritura; desarrollo; géneros.

Abstract

The study aimed to know the way in which some measures of lexical complexity are used in texts written by students of three levels of High School and two socioeconomic groups. 240 narrative and explicative texts were elicited through a brief video about Internet and social networks and then labeled with CLAN (MacWhinney, 2000) in three types of lexical measures: diversity, density, nominal abstraction and types of processes. For the data analysis, T-tests and a mixed factorial ANOVA design were applied with two intergroup factors (age/school level and SES group). The results show that lexical performance changes significantly as a result of two main factors: age/school level and discursive genre. On the other hand, they also make clear that lexical complexity must be studied according to various measures and in various communicative situations. The conclusions of the study also reaffirm the important role of lexicon in the development of writing skills and the command of genres.

Keywords: lexical complexity; writing; development; genres.

1. Introducción

Hacia el final de la infancia y durante toda la adolescencia se presentan cambios neurológicos y cognitivos que, combinados con la creciente experiencia social, comunicativa y letrada de los hablantes, resultan en importantes crecimientos tanto lingüísticos como discursivos (Tolchinsky, 2004; Nippold, 2007). Uno de estos crecimientos está dado por el conocimiento de vocabulario, cuyo dominio es crucial en los procesos de desarrollo no solo semántico, sino que también sintáctico (Serrat, Sanz y Bel, 2004; Devescovi, Caselli, Marchione, Pasqualletti, Reilly y Bates, 2005), y constituye el cimiento lingüístico del niño.

Aprender una nueva palabra implica formar nuevas representaciones semánticas y establecer vínculos con dominios conceptuales, algunos preexistentes en la memoria del sujeto y otros en proceso de elaboración. Por ello, el crecimiento del léxico potencia exponencialmente la construcción de nuevas relaciones semánticas y conceptuales, lo que configura una creciente “cartografía entre el mundo, el significado y la forma” (Dockrell y Messer, 2004: 36). De esta manera, podemos afirmar que el léxico define muy claramente un campo de interacción entre el contexto y la cognición (Tolchinsky, Martí y Llauradó, 2010) y, por lo tanto, que una mayor complejidad léxica exige y, a la vez, refleja una mayor complejidad cognitiva en distintos procesos y niveles.

Junto con estas importantes cuestiones psicolingüísticas, el estudio ontogenético del léxico es relevante, porque su crecimiento se relaciona también con múltiples habilidades comunicativas complejas, tales como la comprensión lectora (Perfetti, 2010; McKoon y Ratcliff, 2016), la producción escrita (Galloway y Uccelli, 2015; Kyle y Crossley, 2016), el rendimiento escolar y académico (Milton y Treffers-Daller, 2013; Roche y Harrington, 2013) y la adquisición de segundas lenguas (Bogaards y Laufer, 2004), entre otras.

En este contexto, la presente investigación tiene por objetivo explorar algunas medidas vinculadas al desarrollo del conocimiento léxico de estudiantes chilenos. Se busca determinar el efecto que tienen factores como la edad (medida como nivel de escolaridad), el grupo socioeconómico y el género discursivo en el uso que hacen los estudiantes de los recursos léxicos. Al respecto, el estudio se pregunta: qué medidas permiten dar cuenta de la complejidad léxica, de qué manera estas medidas expresan la complejidad léxica y cómo tales medidas se ven modificadas por efecto de la edad/nivel escolar, el grupo socioeconómico y el género discursivo.

1.1. Medición de la complejidad léxica

Debido a su innegable centralidad en múltiples procesos cognitivos y lingüísticos, la búsqueda de medidas léxicas eficientes resulta tan necesaria como compleja, razón por la cual

el diagnóstico de la *calidad* o *complejidad* léxica no puede reducirse a un fenómeno de definición única. La mayoría de los autores coincide en señalar que la calidad léxica es un constructo multidimensional que puede ser observado de acuerdo con un conjunto de medidas (Malvern y otros, 2004; Jarvis, 2013a, 2013b). Algunas de las medidas más habituales son las de disponibilidad léxica, densidad, diversidad (variación), sofisticación o rareza, y el análisis de diversos rasgos semánticos (tales como el grado de abstracción/concreción o el carácter metacognitivo) en sustantivos y verbos, entre otras. Todas informan de aspectos diversos y su pertinencia, por lo tanto, varía también de acuerdo con distintos factores. Así, se ha encontrado que algunas se relacionan más significativamente con la edad/nivel de los hablantes, otras con la modalidad (oral o escrita) y otras con el género discursivo, como se comentará a continuación.

Los estudios sobre disponibilidad léxica, de manera general, se proponen contar el número de palabras que los sujetos pueden mencionar o escribir en torno a un determinado campo semántico y han demostrado de modo bastante consistente que la cantidad disponible de vocabulario se ve significativamente afectada por la edad/nivel (Valencia, 1997; Valencia y Echeverría, 1999). Si bien es un fenómeno interesante, por tratarse de una medición de vocabulario sin contexto discursivo, no ha sido contemplada en la presente investigación.

La noción de densidad léxica, en tanto, alude al contenido informativo de los textos. Algunos autores la miden calculando la proporción de las palabras de contenido (sustantivos, verbos, adjetivos y, en algunos casos, adverbios) sobre el total de palabras, y otros calculan una proporción entre las palabras de contenido y las de función o palabras de lista cerrada (preposiciones, conjunciones, artículos, etc.). En cualquiera de los dos casos el índice da cuenta de la densidad informativa del mensaje (aunque la mayor o menor densidad léxica de un texto no es claramente perceptible para el oyente o lector). Numerosas investigaciones señalan la pertinencia de este indicador para distinguir entre las modalidades oral y escrita del discurso (Castellà, 2001; Ravid, 2005; Berman y Nir, 2010) y entre géneros (Johansson, 2008; Berman y Nir, 2010; Berman, Nayditz y Ravid, 2011). Asimismo, se ha encontrado que a medida que los estudiantes avanzan en la escolaridad escriben textos más densos léxicamente (Johansson, 2008; Berman y Nir, 2010).

La variación o diversidad léxica, por su parte, alude normalmente a la habilidad para utilizar palabras distintas y evitar la repetición (Read, 2000) y se mide a través del TTR (*type-token ratio*), que es la proporción entre las palabras distintas (*types*) y el total del palabras (*tokens*). Dado que el TTR solo puede aplicarse en textos de igual extensión, suele usarse alguna fórmula corregida (apta para trabajar con textos de diversa extensión) como CTTR o D (Malvern y otros, 2004). En diversos estudios la diversidad léxica, a veces identificada como riqueza, se ha revelado sensible al factor edad/nivel y

género discursivo (Malvern y otros, 2004; Hudson, 2009; Berman y Nir, 2010). Sin embargo, también ha recibido críticas por tratarse de una medida que no puede dar cuenta apropiadamente de la riqueza implicada en la noción de *diversidad* (Jarvis, 2013a), debido a que solo mediría la no repetición de palabras y no la complejidad composicional que supone este concepto.

La idea de *sofisticación* léxica, por otro lado, se ha relacionado con la extensión de las palabras (las palabras con un mayor número de sílabas serían más sofisticadas), pero, más habitualmente, con el uso de palabras de baja frecuencia (incluidos términos técnicos y especializados), que permiten a los hablantes/escritores expresarse de un modo más preciso (Hudson, 2009; Read, 2000). Además de ser una variable relacionada con la edad/nivel, el vocabulario sofisticado y de baja frecuencia es un rasgo típico de ciertos géneros y contextos discursivos, como la comunicación académica y profesional (Berman y Ravid, 2009; Snow y Uccelli, 2009). Una gran dificultad de las investigaciones evolutivas de la sofisticación léxica vinculada con la frecuencia de palabras es que requieren estudios previos de léxico local frecuente y no frecuente, además de un cierto conocimiento del vocabulario utilizado por la población a la que pertenecen los sujetos (conocimiento que todavía no se ha desarrollado en Chile).

Por último, una evaluación más cualitativa del desempeño léxico apunta a observar ciertos rasgos semánticos en las palabras de contenido, dado que son estas las portadoras de la densidad informativa en los textos y también son las que están más claramente relacionadas con el desarrollo conceptual de los hablantes. Diversos estudios han encontrado, por ejemplo, que la capacidad de definir y utilizar ciertos tipos de sustantivos abstractos y expresiones metafóricas es tardía y que se ve significativamente afectada por el grado de escolaridad, literacidad y por el género discursivo (Nippold, Hegel, Sohlberg y Schwarz, 1999; Ravid, 2006; Hess, 2013; Quiroga, 2015). Este desafío cognitivo implicado en el uso de los sustantivos abstractos también tiene su correlato en los verbos metacognitivos y metalingüísticos, cuyo uso eficiente solo se alcanza durante la adolescencia (Fusté-Herrmann y otros, 2006; Sun y Nippold, 2012). Para dar cuenta también de esta variable más cualitativa de la evaluación del léxico, se han incluido en este estudio (parcialmente, por razones de espacio) también algunos resultados relacionados con ciertos aspectos de sustantivos y verbos.

2. Metodología

Con el fin de alcanzar los objetivos presentados, se elaboró un diseño cuasiexperimental con grupos naturales que permitiese evaluar el conocimiento productivo de vocabulario de los participantes.

2.1. Medidas de complejidad léxica

Las variables definidas para el presente estudio fueron: diversidad léxica (a través de la medida *D*), densidad léxica (proporción de palabras de contenido por total de palabras), abstracción nominal y tipos de procesos expresados por los verbos.

En el caso de la abstracción nominal, se midió la frecuencia relativa de los sustantivos, clasificados según la escala de la tabla 1. Esta escala adapta para el español las propuestas de Ravid (2006) y Berman y Nir (2010) y agrupa las diez categorías originales en cuatro niveles: el primero incluye todos los sustantivos concretos (tangibles) y los nombres propios; el segundo agrupa varias categorías de sustantivos con menor grado de concreción (relaciones, locaciones, etc.); el tercero agrupa sustantivos abstractos imaginables y sustantivos genéricos; por último, el cuarto agrupa sustantivos abstractos genuinos (no imaginables) y expresiones metafóricas. Debido al interés que presentan para los géneros abordados en este estudio, se reportan solo los resultados de los sustantivos pertenecientes al tercer nivel, es decir, sustantivos abstractos “imaginables” y genéricos. Si bien también resultaban de interés los sustantivos del nivel cuatro —por estar asociados a usos léxicos más sofisticados—, su frecuencia de uso era demasiado baja como para poder considerarlos en los análisis estadísticos realizados.

Respecto de los tipos de procesos, se adaptó la clasificación propuesta por Halliday (1985) que distingue seis tipos: materiales, mentales, relacionales, de comportamiento, verbales, existenciales. En el caso de este estudio, el análisis se centró en tres de estas categorías, debido a la particular importancia que tienen en los géneros estudiados: los procesos existenciales y relacionales, que se analizaron agrupados, y los procesos mentales. Los primeros resultan fundamentales en las generalizaciones esperables en una explicación y los segundos están asociados a una mayor complejidad narrativa: permiten introducir la reflexión sobre eventos y personas. En la tabla 2 se describen ambas categorías y se muestran ejemplos del corpus estudiado.

2.2. Instrumentos

Para medir las variables de interés se diseñaron dos tareas de escritura, que correspondieron a elicitaciones semiespontáneas mediante material audiovisual acerca del tópico, como es habitual en estudios evolutivos y sobre escritura (Scott y Windsor, 2000; Berman y Verhoeven, 2002; Behrens, 2008; entre otros). Además, se aplicó con anterioridad una encuesta de intereses acerca de la escritura, que buscaba conocer qué tema resultaba atractivo de abordar en las dos tareas. Los resultados de la encuesta permitieron determinar el tópico más elegido por los estudiantes: Internet y las redes sociales. Tanto la encuesta como las tareas de escritura fueron evaluadas por expertos, a quienes se les pidió revisarlas en relación con aspectos tales

TABLA 1

Escala de abstracción nominal

NIVEL	TIPOS DE SUSTANTIVOS	EJEMPLOS
1	1. Nombres propios	Internet, Youtube, María, Pearl_Jam
	2. Objetos concretos y personas concretas	(Un) niño (<i>de mi curso</i>); computador, película
	3. Sustantivos colectivos y locaciones particulares	curso, plaza, metro, sala, página web, red social, cuenta
2	4. Roles y relaciones	padre, miembro, usuario, profesor, alumnos
	5. Términos temporales	semana, año, semestre, vez (<i>esa vez, la mayoría de las veces, tres veces...</i>)
	6. Eventos y situaciones particulares	discusión, sesión, conversación, trabajo, conflicto, publicación, opinión (en un medio)
3	7. Términos genéricos para objetos, actividades, lugares, tiempos y personas (cuando son categorías aplicables a muchas entidades distintas)	cosa, asunto, personas, animales, gente, servicio, contenido, género, función, forma (<i>la única forma de ...</i>), eventos (<i>en general</i>), uso, veces (<i>en general</i>), en la actualidad, días (<i>en general</i>), mundo (<i>se usa en todo el mundo</i>)
	8. Sustantivos abstractos imaginables (vinculados con los sentidos, las cualidades)	plan, tamaño, costo, aroma, meta, horario, forma (<i>tiene forma de...</i>), perfil, paso, rumor
4	9. Sustantivos abstractos genuinos, <i>no imaginables</i>	desafío, peligro, prejuicio, equivocación, error, problema, suerte
	10. Extensiones metafóricas de términos concretos y abstractos	(<i>hundido en un</i>) pozo; travesía (<i>hacia el interior</i>)

TABLA 2

Tipos de procesos expresados por los verbos

CATEGORÍA	DESCRIPCIÓN	EJEMPLOS
MENTAL	Verbos que designan actividades mentales o sentimientos, sin manifestación física / externa evidente. Siempre debe haber un agente humano involucrado en un proceso mental	desear, concluir, decidir, comparar, cumplir un objetivo, engañar, pensar, creer
EXISTENCIAL/ RELACIONAL	Verbos que designan atribuciones de cualidad, identidad o circunstancia y relaciones entre fenómenos	ser inteligente, estar helado (el día), parecer un niño, estar enfermo, haber/tener un problema, ocupar un servicio (habitualmente, como una cualidad de alguien)

como pertinencia e interés del tema, claridad de instrucciones, suficiencia de la situación retórica, entre otros. Sus sugerencias permitieron mejorar los instrumentos antes de la aplicación y establecer como versiones definitivas de ambas tareas las que se presentan en la tabla 3.

TABLA 3

Tareas de escritura

HISTORIAS EN LA RED	RECURSOS DE LA RED
<p>Como viste en el video, cada día vivimos distintas experiencias relacionadas con Internet: conocemos personas o proyectos interesantes, nos informamos sobre actualidad, compramos o vendemos, expresamos y compartimos nuestras ideas y emociones, nos divertimos con música y videos, hacemos denuncias o convocatorias, etc.</p>	<p>Como viste en el video, Internet ofrece muchos sitios y recursos para conocer nuevas cosas y aprender sobre lo que nos interesa: nos informamos sobre actualidad en la prensa y en las redes sociales, utilizamos buscadores de información y enciclopedias, aprendemos nuevos juegos, creamos y desarrollamos perfiles personales o grupales, compartimos información, música, películas y programas, entre otras posibilidades.</p>
<p>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 difusión musical o artística gracias a estos medios, etc.).</p>	<p>Elige uno de esos sitios o 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.</p>
<p>Tu texto podría aparecer en una revista digital o Blog donde niños y jóvenes cuentan sus experiencias con Internet.</p>	<p>Tu texto podría aparecer en una revista digital o Blog donde niños y jóvenes aprenden acerca de los usos de Internet.</p>
<p>Es importante contar la historia lo más completa posible. Puedes hacer un borrador si lo deseas y utilizar todas las hojas que necesites.</p>	<p>Es importante informar y explicar de la manera más completa posible. Puedes hacer un borrador si lo deseas y utilizar todas las hojas que necesites.</p>

Debido a que resultaba necesario informar a los establecimientos educacionales que participaron sobre los resultados obtenidos por sus estudiantes, se diseñaron y validaron, además, dos rúbricas para evaluar el desempeño global obtenido en las tareas de escritura. Estas rúbricas incluyeron indicadores de organización, coherencia, cohesión, léxico y construcción oracional. Para los efectos del estudio aquí presentado, los resultados de las rúbricas solo se utilizaron para establecer puntajes de desempeño global para contar con una referencia acerca de la calidad general de los textos y para seleccionar el corpus definitivo, tal como se explica en el siguiente apartado.

2.3. Participantes

La muestra estuvo conformada por 120 estudiantes pertenecientes a los niveles de séptimo básico (12 años), primero medio (14 años) y tercero medio (16 años) de dos establecimientos

mixtos de la Región Metropolitana representativos de los grupos socioeconómicos alto (GS alto) y medio-bajo (GS medio-bajo), según información proporcionada por el Ministerio de Educación. Se seleccionaron estos niveles, ya que las edades que típicamente tienen sus estudiantes se consideran relevantes para los fenómenos que interesaba observar en el estudio (Nippold, 2007).

Las dos tareas de escritura fueron aplicadas a todos los estudiantes que cursaban estos niveles (unos 720 en total). Los textos resultantes fueron revisados con las dos rúbricas, lo que permitió asignarles puntajes de desempeño global. Considerando este puntaje, fueron excluidos de la muestra todos los sujetos con rendimiento *sobre* y *bajo* las 1,96 unidades de desviación estándar (puntaje Z), medida utilizada habitualmente para la detección de los casos atípicos. Un análisis de potencia estadística calculado en G*Power (Faul, Erdfelder, Lang y Buchner, 2007) permitió establecer un N muestral mínimo de 108 sujetos. Este número garantizaba óptimas posibilidades de detectar significaciones estadísticas asociadas a un tamaño de efecto cercano al típicamente considerado como mediano (0,25) para las pruebas estadísticas que serían utilizadas. Por lo tanto, después de excluir los casos atípicos se seleccionaron de manera aleatoria 120 participantes: 20 estudiantes de cada nivel y establecimiento (10 hombres y 10 mujeres), cuyos 240 textos conformaron el corpus definitivo, tal como se ilustra en las tablas 4 y 5. La primera muestra la distribución de los 240 textos en los distintos grupos estudiados y la segunda las medias grupales de la extensión de los textos (contada en cantidad de palabras) y el puntaje de desempeño global obtenido en la rúbrica antes mencionada (sobre un total de 23 puntos).

2.4. Procedimientos de transcripción, segmentación y etiquetado del corpus

Con el fin de asegurar consenso y consistencia en el tratamiento del corpus, se siguieron las orientaciones que la bibliografía especializada sugiere para este tipo de procedimientos (Neuendorf, 2002; Stemler, 2004). Así, se elaboró un protocolo de transcripción y un manual de

TABLA 4

Distribución en los distintos subgrupos de los 240 textos producidos por los 120 participantes

	TAREA NARRATIVA			TAREA EXPLICATIVA			TOTAL
	SÉPTIMO BÁSICO	PRIMERO MEDIO	TERCERO MEDIO	SÉPTIMO BÁSICO	PRIMERO MEDIO	TERCERO MEDIO	
GS Medio-Bajo	20	20	20	20	20	20	120
GS Alto	20	20	20	20	20	20	120
Total	40	40	40	40	40	40	240

TABLA 5

Caracterización del corpus: media de extensión de los textos (y media de puntaje global obtenido en la rúbrica, de un total de 23 puntos)

	TAREA NARRATIVA			TAREA EXPLICATIVA		
	SÉPTIMO BÁSICO	PRIMERO MEDIO	TERCERO MEDIO	SÉPTIMO BÁSICO	PRIMERO MEDIO	TERCERO MEDIO
GS Medio-Bajo	232 (13)	198 (14)	295 (16)	136 (13)	146 (14)	201 (15)
GS Alto	178 (15)	302 (18)	442 (19)	177 (14)	243 (16)	366 (17)

segmentación y etiquetado, y parte del corpus fue sometido a distintas rondas de calibración con el fin de asegurar la consistencia de los investigadores en la aplicación de los criterios definidos, tal como se describe a continuación.

Los 240 textos fueron transcritos de forma literal utilizando un procesador de textos y luego convertidos a formato CHAT (*.cha) con el fin de poder realizar los análisis en el programa CLAN (MacWhinney, 2000), habitual en estudios del área. Preparar los archivos para el análisis requirió hacer algunas modificaciones, como unir locuciones y nombres propios compuestos, con el fin de no alterar el número total de palabras, decisión que también ha sido tomada en otros estudios sobre léxico (Valencia y Hermansen, 1985; Valencia, 1994, 1997; entre otros), y para evitar que el análisis morfológico realizado a través de CLAN interpretara erróneamente algunas categorías gramaticales.

Posteriormente, los textos fueron segmentados en cláusulas, según la propuesta de Ber- man y Slobin (1994). La rigurosidad en este proceso era clave para algunos análisis posteriores (como la densidad nominal, que no se presenta aquí), por lo que un 30% del corpus (72 textos) fue segmentado por dos investigadores distintos y revisado por un tercero, hasta asegurar un 100% de acuerdo en la aplicación de los criterios.

Con los archivos ya preparados, se ejecutó el programa MOR de CLAN, que genera un análisis morfológico de cada una de las cláusulas en las que se dividieron los textos. Luego, a las palabras no reconocidas se les asignó la categoría gramatical correspondiente y, posteriormente, todo el corpus fue revisado para resolver ambigüedades debidas a la interpretación de ciertas palabras como pertenecientes a más de una categoría (por ejemplo, como participio y adjetivo).

Una vez segmentado y analizado morfológicamente el corpus, se inició el proceso de etiquetado manual de fenómenos léxicos. Para asegurar un etiquetado consistente, 24 textos (10% del total) fueron etiquetados cuatro veces (por distintos investigadores) y sus resultados se cotejaron hasta lograr un 100% de acuerdo en la aplicación de los criterios. Finalmente,

dos investigadores se dividieron el etiquetado de los textos restantes, verificando constantemente la calibración entre ambos, y un tercero realizó la revisión final, corroborando la consistencia y decidiendo la notación definitiva en casos de discrepancia.

2.5. Procesamiento de datos

Para el análisis de los datos se aplicó un diseño de ANOVA factorial mixto con dos factores intergrupo: edad/nivel escolar (en el entendido de que este factor se asocia con distintos grados de experiencia letrada y lingüística en general) y grupo socioeconómico (en el entendido de que este factor se relaciona en Chile con el grado de alfabetización familiar y el tipo de establecimiento y comuna). Para las comparaciones *post hoc* se utilizó la prueba Bonferroni. También se aplicaron pruebas T para muestras relacionadas con el objetivo de comparar el desempeño entre ambos géneros discursivos de cada sujeto.

3. Resultados

3.1. Diversidad y densidad léxica

Los resultados del estudio muestran que los estudiantes mayores tienden a escribir textos narrativos y explicativos significativamente más diversos léxicamente que los menores, aunque, como veremos, los perfiles evolutivos son distintos en cada establecimiento. El factor socioeconómico y el género discursivo no modifican significativamente la diversidad, de acuerdo con estos resultados. La densidad léxica cambia significativamente producto del género en algunos grupos, preferentemente de estudiantes mayores. En las siguientes tablas se muestran las medias grupales para diversidad (tabla 6) y densidad léxica (tabla 7) y se comentan los resultados obtenidos en las pruebas ANOVA (comparación por edad/nivel y grupo socioeconómico) y en la prueba T de muestras relacionadas para observar las diferencias atribuibles al género discursivo.

En cuanto a la diversidad —que, recordemos, se está midiendo con la fórmula corregida de TTR, D —, los resultados muestran que la edad/nivel es un factor significativo, tanto en la tarea narrativa ($F(1,114) = 9,550$; $p = ,000$) como en la explicativa ($F(1,114) = 5,842$; $p = ,004$). En el caso del establecimiento GS alto, las comparaciones *post hoc* evidencian que los textos escritos por estudiantes de tercero medio son más diversos léxicamente que los de séptimo en ambos géneros (narrativo: $p = ,000$; explicativo: $p = ,013$). En el caso del establecimiento GS medio-bajo solo existen diferencias estadísticamente significativas en el género narrativo, en el que los textos de los estudiantes de tercero medio presentan una mayor diversidad léxica que los de primero medio ($p = ,040$). Como se observa en la tabla 6, en este caso resulta interesante la comparación de la dispersión de los valores de D . En tercero medio del GS medio-bajo

TABLA 6Diversidad léxica (*D*)

		NARRATIVO			EXPLICATIVO		
		MEDIA (DE)	MÍN.	MÁX.	MEDIA (DE)	MÍN.	MÁX.
GS Medio	Séptimo Básico	82,49 (22,31)	45,56	129,68	89,06 (33,01)	38,61	166,18
	Primero Medio	77,66 (18,23)	35,79	102,96	84,20 (16,87)	51,41	117,30
	Tercero Medio	92,76 (17,45)	58,34	129,99	98,66 (21,73)	61,22	144,07
GS Alto	Séptimo Básico	76,74 (16,25)	54,23	122,03	83,02 (12,80)	62,14	106,56
	Primero Medio	87,58 (19,07)	48,03	132,74	91,08 (15,22)	55,47	115,16
	Tercero Medio	101,14 (19,84)	58,12	133,66	101,79 (15,39)	63,90	126,73

no hay estudiantes que escriban textos narrativos más diversos respecto de los de séptimo (en ambos niveles el máximo es 129): la mejora en la media general se debe a que los puntajes mínimos son más altos (de 45,56 en séptimo a 58,34 en tercero).

El factor grupo socioeconómico, por su parte, no resultó tener incidencia en la diversidad. Si bien este resultado parece algo inesperado, conocemos otras investigaciones recientes en que el factor GS no incide en algunas medidas léxicas habituales (Quiroga, 2015). Al respecto, no creemos que no existan diferencias entre la diversidad del léxico que dominan los estudiantes de diversos grupos socioeconómicos, particularmente en Chile, donde la segregación del sistema escolar refleja experiencias escolares y letradas muy dispares. Postulamos, más bien, que hay importantes diferencias en el dominio léxico de estudiantes pertenecientes a estos dos grupos socioeconómicos, pero que este tipo de medidas no se muestra sensible a este factor, o lo hace de manera poco consistente. Tal como se comentará más adelante, se requieren más estudios que exploren el comportamiento de estas y otras medidas de desempeño léxico en distintos tipos de población. Tampoco se identificaron diferencias estadísticamente significativas entre géneros, es decir, la diversidad léxica pareciera ser una medida vinculada más a características propias de los sujetos que de los textos que escriben.

En el caso de la densidad léxica, no hay efecto significativo de la edad/nivel, pero se identifica una interacción entre este factor y el grupo socioeconómico en la tarea explicativa ($F(2,114) = 4,083$; $p = ,019$), es decir, ambos establecimientos presentan patrones evolutivos distintos. Como se observa en los datos de la tabla 7, la diferencia se produce ya que los es-

TABLA 7Densidad léxica (*palabras de contenido por total de palabras*)

		NARRATIVO			EXPLICATIVO		
		MEDIA (DE)	MÍN.	MÁX.	MEDIA (DE)	MÍN.	MÁX.
GS Medio-bajo	Séptimo Básico	0,38 (0,03)	,34	,47	0,43 (0,04)	,32	,55
	Primero Medio	0,39 (0,04)	,32	,48	0,42 (0,04)	,33	,49
	Tercero Medio	0,38 (0,03)	,33	,45	0,42 (0,03)	,38	,48
GS Alto	Séptimo Básico	0,39 (0,03)	,34	,45	0,39 (0,03)	,33	,44
	Primero Medio	0,41 (0,04)	,32	,46	0,42 (0,03)	,37	,47
	Tercero Medio	0,39 (0,03)	,32	,45	0,42 (0,03)	,34	,47

tudiantes de séptimo del GS alto escriben textos explicativos menos densos léxicamente ($p = ,002$). Así, en este grupo socioeconómico se produce un aumento de la densidad entre séptimo y tercero ($p = ,025$), mientras que en el GS medio-bajo se mantiene estable.

En relación con el género discursivo, se encontraron diferencias principalmente en los estudiantes mayores y en un grupo de séptimo: tercero medio del GS medio-bajo ($t(19) = -4,09$, $p = ,000$), tercero medio del GS alto ($t(19) = -3,12$, $p = ,010$) y séptimo básico del GS medio-bajo ($t(19) = -4,69$, $p = ,000$). Todos ellos escriben textos explicativos más densos léxicamente que los textos narrativos. Esto se explica probablemente porque los textos explicativos (de este corpus, por la tarea solicitada, pero tal vez también en general) hacen abundante uso de sustantivos comunes y propios necesarios para describir y explicar el tema seleccionado en la tarea. Además, en el caso de los estudiantes mayores se presume un uso mayor de nominalizaciones, complejo recurso gramatical que se ha relacionado con la escritura explicativa y madura (Snow y Uccelli, 2009).

A modo de síntesis de este apartado, se puede afirmar que los estudiantes mayores de ambos establecimientos escriben textos más diversos léxicamente y que los textos explicativos muestran mayor densidad léxica que los narrativos, rasgo significativo también en los mayores. Si bien séptimo básico del grupo socioeconómico medio presenta rasgos similares a los de los estudiantes mayores (más diversidad y densidad), creemos que es necesario ser cautos con este resultado, ya que las medidas de dispersión hacen sospechar de la existencia de casos fuera de rango en estas variables que hacen aumentar los promedios de ese grupo.

Puede ocurrir, por ejemplo, que algunos textos breves y fragmentados (con enumeración) presenten una altísima densidad léxica, porque al enumerar muchos objetos o acciones tienen una alta proporción de palabras de contenido.

Los siguientes ejemplos permiten ilustrar y comentar algunos de los resultados reportados en este apartado, particularmente en lo relativo a la diversidad léxica y el efecto que tiene en ella el factor edad/nivel escolar. Para la selección de estos ejemplos se requerían textos breves (para poder mostrarlos completos) de similar extensión (aunque esto no es lo habitual en textos de distintas edades) y bastante diferentes en cuanto a diversidad léxica, de modo de poder aislar la variable dentro de lo posible y observar más claramente las diferencias de los textos en relación con ella. Para esto se buscaron textos que hubieran obtenido en diversidad léxica un resultado similar o por *debajo* del promedio de su grupo (en el caso de séptimo básico) y similar o *superior* al promedio de su grupo (en el caso de tercero medio).

Ejemplo 1

No te confíes

Yo conocí a una persona que estaba metida, creo que en Facebook, que agregaba a personas, que no conocía. Una vez esa chica agregó a un chico de su misma edad, como de unos 19 años, y empezaron a hablar y hablar, durante varios días, hasta que un día el chico le dijo que “tenían que juntarse algún día”, y la chica le dijo, “dale, que sí estaba bien el viernes” entonces quedaron para juntarse el viernes.

El día viernes la chica salió y fue adonde le dijo que se iban a juntar, y regresó con un policía a la casa. Cuando los padres vieron que estaba con un policía, fueron a preguntar qué había pasado. El policía les contestó que “su hija estaba agregando gente que no conocía, y que con un chico quedó de verse hoy”.

También les dijo que su hija tuvo suerte, porque ese chico era la policía, que hace estas cosas, para proteger a las menores de edad que sin conocer a las personas, llegan hasta sacarse la ropa enfrente de la cámara, porque les dicen que lo hagan (2110126).

Ejemplo 2

Facebook en los negocios

Hace más o menos un año atrás, con un amigo se nos ocurrió una idea para tener nuestro propio negocio y así poder ganar plata para nuestro uso personal, era ser DJ's. Pero como siempre, había un problema y este era que no sabíamos cómo expandirnos y hacer que la gente sepa que había un nuevo disco que podían contratar para fiestas de 15, eventos en empresas y para todo lo que quisieran.

Entonces nos hicimos una página en Facebook (FB). Con eso empezamos a tener cada vez más “me gusta”, con eso nos empezaron a llamar, a pedirnos que hiciéramos cosas distintos días del fin de semana y así también se hizo una propaganda del boca a boca. Entonces así nos hicimos conocidos gracias a la red social llamada Facebook. Esto fue muy importante ya que invade casas

de Chile y se podría llamar un vicio. Es muy útil para emprendimientos pequeños como los de los adolescentes. Gracias a todo eso ahora estamos trabajando para una prestigiosa discotheque en el barrio alto de Santiago (2314218).

El ejemplo 1 (séptimo básico) tiene 187 palabras y una diversidad baja ($D = 61,31$ frente a la media grupal de $D = 76,74$); en la rúbrica obtuvo un puntaje de desempeño global de 15 (lo que equivale a 63% de logro). El ejemplo 2 (tercero medio) tiene 179 palabras y una diversidad alta, $D = 112,41$ (la media grupal es $D = 101,14$); su puntaje de desempeño global es 14 (58% de logro). La densidad léxica de ambos textos es 0,39. Tal como se puede observar, en (1) se repiten con frecuencia palabras que suelen usarse de manera genérica (persona, chico, chica, decir, etc.), mientras que en (2) hay menos repetición y algunos términos más precisos (expandir, contratar, invadir, propaganda, vicio, etc.). Así, se observa en estos ejemplos que la medida de diversidad léxica se vincula principalmente al fenómeno de no repetición, como señalan algunos autores (Jarvis, 2013b). Por otro lado, es importante destacar que el texto (2), no obstante su altísimo índice de diversidad léxica, es un texto de calidad mediana, similar a (1), si se consideran globalmente otros factores, como el puntaje de desempeño global o la complejidad sintáctica y los componentes de la estructura narrativa (que no son objeto del presente análisis). En este sentido, es interesante corroborar que un determinado índice de diversidad léxica no se relaciona necesariamente con la calidad global de los textos (Jarvis, 2013b).

Para ilustrar este punto, se ha incluido un tercer texto. El ejemplo 3 (tercero medio) tiene 254 palabras y una diversidad media, $D = 80,16$, muy por debajo de la diversidad de (2), que era de 112; su densidad léxica es de 0,45 y su puntaje de desempeño global es 20 (83% de logro).

Ejemplo 3

Camila era una niña que disfrutaba enormemente de su celular. Lo había adquirido recientemente, por lo que pasaba largas horas con su nuevo teléfono inteligente. “No solo me lo compré para entretenerme”, decía Camila, pero a decir verdad el aparato era su vida, hablaba con amigos todo el día por *Whatsapp* y se preocupaba de tomarle fotos a todo lo que pudiese para luego retocarlas y subirlas a la web.

Alberto y Roxana, sus padres, comenzaron a preocuparse, contemplaron el actuar de su hija y no comprendían en qué momento Camila se había transformado en una persona totalmente obsesionada con la tecnología y las apariencias.

Su actuar dio como resultado una baja en sus calificaciones, distraída por los mensajes y por las fotos de las redes sociales, aplazaba las horas de estudio y no prestaba atención en clase. Un día el profesor de ciencias naturales, preocupado por Camila, cita a su madre para hablar del comportamiento de su pupila. Tras un par de minutos de charla, la madre decide tomar una drástica decisión.

Al día siguiente Camila se despertó con una gran sorpresa, su celular ya no estaba pero en su lugar un antiquísimo aparato se encontraba en su velador. Fue así como Camila comprendió que su

obsesión había llegado a tal punto, que el moderno aparato, no era ya un teléfono para comunicarse con sus familiares y amigos, sino que se había transformado en un ente frívolo y enajenado, comprendiendo que había olvidado por completo el verdadero objetivo y función de un teléfono (2315116).

Como se puede observar a simple vista, el texto (3) es superior a los anteriores, lo que se debe a múltiples factores, algunos de los cuales no son objeto de esta exposición (estructura narrativa, complejidad sintáctica, dominio de la referencialidad, etc.), pero otros sí lo son: el léxico también parece tener un papel importante. Respecto de la densidad léxica, es bastante superior a la media grupal, lo que se manifiesta en la alta frecuencia de sustantivos, verbos, adjetivos y adverbios. Esto contribuye al *espesor* que adquiere la información en este texto. Sin embargo, también es muy importante destacar que las palabras de contenido empleadas no solo poseen una alta frecuencia, sino que también son *sofisticadas* y de *gran pertinencia* y *riqueza*: retocar, “a decir verdad”, “era su vida”, antiquísimo aparato, obsesionada, apariencias, “aplazaba las horas de estudio”, drástica decisión, frívolo, enajenado, etc. Cabe destacar que dentro de esta categoría los adjetivos son un elemento importante (Cutillas y Tolchinsky, 2017) y ameritan futuros estudios que contribuyan a esclarecer su rol como factor interviniente en la complejidad y la riqueza léxica. En el presente estudio no se analizaron los adjetivos, pero sí algunos rasgos de sustantivos y verbos, cuyos principales resultados se comentarán en el siguiente apartado.

3.2. Abstracción nominal y tipos de procesos

Los resultados para las variables de abstracción nominal y tipos de procesos que serán expuestos en este apartado recordemos son los de sustantivos abstractos/genéricos, verbos existenciales y verbos mentales. De manera general, se encontró que los estudiantes mayores de ambos colegios utilizan más sustantivos abstractos y más verbos existenciales que los menores. También, como era de esperar, se observó que muchos estudiantes de ambos establecimientos utilizan un número mayor de sustantivos abstractos y verbos existenciales en sus textos explicativos, y un número mayor de verbos mentales en sus textos narrativos. Esto da cuenta de un buen dominio de los recursos léxicos propios de estos géneros. En las siguientes tablas se muestran las medias grupales para ambos géneros de los sustantivos abstractos/genéricos (tabla 8), de los verbos mentales (tabla 9) y los verbos existenciales (tabla 10), y se reportan y comentan los resultados de las pruebas aplicadas para observar las diferencias por nivel, grupo socioeconómico y género discursivo.

La edad/nivel escolar es un factor significativo en la proporción de sustantivos genéricos y abstractos imaginables, tanto en la tarea narrativa ($F(1,114) = 21,385; p = ,000$) como en la explicativa ($F(1,114) = 6,117; p = ,003$). En la tarea narrativa, las comparaciones *post hoc* eviden-

TABLA 8

Proporción de sustantivos abstractos y genéricos (nivel 3)

		NARRATIVO			EXPLICATIVO		
		MEDIA (DE)	MÍN.	MÁX.	MEDIA (DE)	MÍN.	MÁX.
GS Medio-bajo	Séptimo Básico	0,15 (0,09)	,00	,35	0,23 (0,09)	,09	,45
	Primero Medio	0,21 (0,09)	,05	,39	0,29 (0,16)	,11	,80
	Tercero Medio	0,26 (0,11)	,10	,56	0,25 (0,07)	,13	,37
GS Alto	Séptimo Básico	0,13 (0,10)	,00	,39	0,21 (0,11)	,03	,41
	Primero Medio	0,16 (0,05)	,08	,33	0,27 (0,08)	,11	,51
	Tercero Medio	0,30 (0,11)	,15	,52	0,36 (0,11)	,15	,57

cion que existe mayor proporción de este tipo de sustantivos en los textos de los estudiantes mayores. En el GS alto, los estudiantes de tercero se diferencian tanto de los de séptimo básico (narrativo: $p = ,000$; explicativo: $p = ,000$) como de los de primero medio (narrativo: $p = ,000$; explicativo: $p = ,041$). En el establecimiento de GS medio-bajo, las diferencias se encuentran entre tercero medio y séptimo básico ($p = ,001$).

En el caso del género explicativo, es posible identificar, además, una interacción entre el factor edad/nivel escolar y el grupo socioeconómico ($F(2,114) = 4,489$; $p = ,013$), que evidencia la existencia de patrones evolutivos distintos. Esta interacción se explica por el hecho de que los estudiantes del GS medio-bajo de todas las edades utilizan una proporción similar de este tipo de sustantivos, mientras que en los del GS alto se evidencia un aumento producto de la edad.

Resulta interesante destacar que en ambos establecimientos es en el género narrativo donde este tipo de sustantivos aumenta en mayor proporción, resultado que en principio podría parecer sorprendente, ya que es el género explicativo el que exige un uso mayor de sustantivos abstractos y genéricos. Sin embargo, una posible interpretación es que es precisamente por esa razón que aumentan más en la tarea narrativa: se trata de un rasgo de sofisticación cuyo dominio se adquiere con la edad y no una exigencia de ese tipo de texto.

Las comparaciones entre géneros evidencian que en algunos niveles existe una tendencia a utilizar más sustantivos abstractos y genéricos en los textos explicativos. Es el caso, por ejemplo, de los séptimos básicos de ambos establecimientos (GS medio-bajo: $t(19) = -2,666$, p

= ,015; GS alto: $t(19) = -2,705, p = ,014$). En el caso del GS alto esta diferencia se observa también en primero medio ($t(19) = -5,404, p = ,000$), pero no así en tercero. Una posible explicación se encuentra en lo comentado en el párrafo anterior: el uso de este tipo de sustantivos es un rasgo de sofisticación en géneros en los que su utilización no es una exigencia y daría cuenta de mayor madurez en el desarrollo del léxico.

TABLA 9

Proporción de procesos mentales

		NARRATIVO			EXPLICATIVO		
		MEDIA (DE)	MÍN.	MÁX.	MEDIA (DE)	MÍN.	MÁX.
GS Medio-bajo	Séptimo Básico	0,21 (0,09)	,00	,37	0,18 (0,12)	,05	,50
	Primero Medio	0,25 (0,14)	,00	,57	0,22 (0,12)	,04	,56
	Tercero Medio	0,28 (0,10)	,14	,54	0,20 (0,07)	,11	,32
GS Alto	Séptimo Básico	0,29 (0,12)	,15	,61	0,22 (0,06)	,14	,35
	Primero Medio	0,25 (0,08)	,13	,47	0,18 (0,10)	,00	,38
	Tercero Medio	0,26 (0,08)	,14	,40	0,15 (0,06)	,07	,29

Con relación a los procesos mentales, no se identificó un efecto del nivel de escolaridad. Como se observa en la tabla 9, las proporciones se mantienen relativamente homogéneas en ambos géneros discursivos. Este resultado es inesperado, puesto que se presume que este tipo de verbos es un rasgo típico de los textos narrativos, por lo que era esperable que su dominio aumentara con la edad/nivel. Una explicación posible es que este rasgo se adquiere más temprano y que, por lo tanto, todos los estudiantes de esta muestra ya dominan este recurso al escribir sus historias.

En esta variable, las principales diferencias se encuentran en las comparaciones por género discursivo. En todos los casos se da un resultado que es esperable: existe mayor proporción de procesos mentales en el género narrativo respecto del explicativo. En el caso del GS alto, se registran diferencias estadísticamente significativas en los tres niveles (séptimo básico: $t(19) = 2,38, p = ,03$; primero medio: $t(19) = 2,49, p = ,02$; tercero medio: $t(19) = 4,29, p = ,00$). En el GS medio-bajo, por su parte, las diferencias se encuentran solo en tercero medio ($t(19) = 3,12, p = ,01$). Una interpretación posible de este resultado es que en el caso del GS alto los estudiantes dominan antes este rasgo propio del género.

TABLA 10

Proporción de procesos existenciales

		NARRATIVO			EXPLICATIVO		
		MEDIA (DE)	MÍN.	MÁX.	MEDIA (DE)	MÍN.	MÁX.
GS Medio-bajo	Séptimo Básico	0,27 (0,11)	,08	,52	0,37 (0,11)	,21	,60
	Primero Medio	0,37 (0,17)	,15	,80	0,40 (0,14)	,06	,73
	Tercero Medio	0,30 (0,10)	,10	,47	0,37 (0,11)	,09	,57
GS Alto	Séptimo Básico	0,25 (0,12)	,07	,50	0,34 (0,08)	,14	,50
	Primero Medio	0,36 (0,10)	,21	,55	0,44 (0,11)	,19	,65
	Tercero Medio	0,29 (0,11)	,10	,48	0,43 (0,08)	,25	,59

Respecto de los procesos existenciales, existe un efecto estadísticamente significativo del factor nivel escolar en ambos géneros (narrativos: $F(1,114) = 8,670$, $p = ,000$; explicativos: $F(1,114) = 3,462$, $p = ,035$). Para el género narrativo, en ambos establecimientos aumentan estos procesos en primero medio respecto de séptimo básico (GS medio-bajo: $p = ,023$; GS alto: $p = ,007$) y luego vuelven a disminuir en tercero medio. Se establece, así, un perfil evolutivo en forma de U invertida para los verbos existenciales. Como se explicará en la discusión, una posible explicación de este fenómeno —consistente con hallazgos anteriores relacionados con medidas no abordadas en este estudio pero de los mismos estudiantes (*autores, año*) y también con la bibliografía sobre este tipo de perfiles (Karmiloff-Smith, 1979, 1992; Rakison y Yermolayeva, 2011)— es que estos estudiantes utilizan estrategias similares entre ambos géneros (por ejemplo, estrategias que implican enumerar, describir para hacer avanzar el texto, incluso si este es narrativo). Y son más cautos que los de séptimo pero aún no logran el dominio de los de tercero y por eso sus textos presentan rasgos a veces híbridos o inesperados para el género. En el caso del género explicativo, por su parte, en el GS alto existen diferencias estadísticamente significativas entre séptimo básico y los otros dos niveles: los textos producidos por estudiantes de primero ($p = ,011$) y tercero medio ($p = ,022$) presentan mayor proporción de procesos existenciales que los de los estudiantes menores. Así, es posible identificar un patrón evolutivo claro: después de los 12 y antes de los 14 años se alcanzaría dominio de este rasgo típico del género.

Con relación a la comparación entre géneros, en todos los grupos, salvo en primero medio del GS medio-bajo, los textos explicativos presentan mayor proporción de procesos existenciales, un resultado esperado, debido a que se trata de un rasgo prototípico del

género (GS alto: séptimo básico, $t(19) = -3,72, p = ,000$; primero medio: $t(19) = -2,22, p = ,040$; tercero medio: $t(19) = -4,83, p = ,000$. GS medio-bajo: séptimo, $t(19) = -3,03, p = ,010$; tercero medio: $t(19) = -3,02, p = ,010$).

Como síntesis de los resultados de este apartado, observamos que dos factores inciden significativamente en la elección de los sustantivos y verbos para la construcción de los textos narrativos y explicativos escritos: la edad/nivel escolar y el género discursivo. En concreto, los estudiantes mayores de ambos colegios usan más sustantivos abstractos y verbos existenciales en ambos géneros. En cuanto al género, muchos estudiantes usan más sustantivos abstractos y más verbos existenciales en los textos explicativos, como era previsible, aunque solo en el GS alto lo hacen con diferencias estadísticamente significativas respecto de los usos en los textos narrativos.

Para ilustrar algunos de los resultados anteriores se han seleccionado dos textos, uno narrativo y el otro explicativo. Se intentó buscar textos que, junto con dar cuenta claramente de los fenómenos expuestos, pertenecieran al mismo estudiante y fueran relativamente breves. Sin embargo, no fue posible que los dos textos de un mismo estudiante cumplieran con todos los requisitos (por lo general, ocurría que el texto narrativo era muy bien logrado pero el explicativo no presentaba los rasgos requeridos, o viceversa). Por ello se seleccionaron textos de distintos estudiantes (pero del mismo grupo etario) priorizando que contuvieran los rasgos léxicos que se quieren destacar.

Ejemplo 4

Hace dos años mi profesor de música, mientras estudiábamos audición musical, me hizo escuchar un tema del artista estadounidense John Petrucci. El tema me interesó mucho, al ser un arreglo instrumental de guitarra eléctrica. Quería saber más sobre Petrucci y el grupo de G3 al que pertenecía. En mi casa hice lo que muchos harían en un caso como este, me dirigí hacia mi computador y busqué en Youtube el tema que había escuchado. Luego me fijé en las sugerencias y encontré a los otros dos guitarristas que formaban G3, Joe Satriani y Steve Vai. Después de escuchar algunos de sus temas me di cuenta que G3 era un concierto anual que reunía guitarristas famosos, organizado por Joe Satriani y no un grupo como yo pensaba. En Facebook me hice fan de John Petrucci y G3. Me sugirió otros artistas parecido como Foo Fighters. En Wikipedia busqué toda la historia de estos artistas y descubrí que Dave Grohl, el líder de Foo Fighters, había sido el baterista principal del grupo Nirvana antes de su separación. Al volver a Youtube a escuchar todos estos grupos se me ocurrió buscar arreglos de guitarra acústica. Así me encontré con los guitarristas Justin King, Ben Lapps y Eric Mongran. Mi amor por la guitarra no parecía ceder. De vuelta en Facebook me reencontré con la página de G3, y ahí supe que hacían giras por todo el mundo. Mediante Blogs y Google supe que no habían venido a Chile desde 2004 más o menos y me puse a pensar si lo harían de nuevo. Busqué por Internet si tenían planeada una gira sudamericana y supe que el 2012 vendrían Joe Satriani, Steve Morse y el que me inició en esta locura, John Petrucci. Hoy yo toco guitarra eléctrica, mi

grupo favorito es Foofighters y me mantengo alerta en Facebook para cuando empiecen a vender entradas para el concierto de G3 en Chile. Escucho caso todos los días algún pieza complicada de guitarra acústica en Youtube e intento aprendérmela mediante tablaturas en Internet. Mi estilo de música preferido es el rock y el instrumental de guitarra. Nunca podría haber descubierto estos grupos de no ser por la Internet. (2212230)

Ejemplo 5

Facebook, una vida virtual.

Facebook es una red social creada el 2004 por Mark Zuckerberg, uno de los hackers más grandes del mundo. Esta red social es una herramienta que sirve para distintas cosas, como por ejemplo:

- Conocer gente
- Hablar con gente
- Contactar gente
- Enterarse de noticias
- Organizarse
- Entre otros.

En el mundo más de 1.000.000 de usuarios se van registrando en Facebook y actualmente casi la mitad del mundo lo utiliza diariamente. Facebook es una herramienta muy útil en la sociedad, para el ámbito de organizarse es sumamente fácil de usar, por ejemplo para organizarte puedes hacer un grupo, un evento, etcétera. Además sirve para la búsqueda de personas, por ejemplo, cuando no ves hace mucho a un pariente, lo buscas en Facebook. Yo tengo parientes míos en Facebook, que en realidad nunca he visto en la vida real. Es interesante y útil porque tengo parientes que viven en: Estados Unidos, Argentina, Italia, Perú y unos en China, entonces Facebook me sirve para contactarlos cada vez que quiero. En conclusión, Facebook es una herramienta global que nos ayuda en distintos ámbitos y hace más sencilla la comunicación a larga distancia (2213102).

El texto en (4) es narrativo, con una extensión de 343 palabras, un puntaje de desempeño global de 20 (80%) y presenta algunos rasgos que permiten considerarlo como muy bien logrado desde el punto de vista de los usos léxicos en relación con el género. En particular, tiene una proporción muy alta de verbos relacionados con procesos mentales = 0,47 (media grupal 0,25), tales como “darme cuenta”, “buscar”, “descubrir”, “interesarme”, “mantenerme alerta” y una proporción normal-baja de sustantivos abstractos/genéricos de 0,13 (media grupal 0,16). Esta baja proporción de sustantivos genéricos (y por lo tanto escasa abstracción nominal) implica el uso preferente de referencia específica (personas, situaciones, lugares concretos), lo que constituye un desempeño léxico competente para este tipo de textos. La alta proporción de verbos mentales, a su vez, también imprime al texto un carácter más específico y privado (esperable, considerando que la tarea solicitada es una historia personal). Es interesante destacar, además, en relación con los resultados del apartado anterior, que este texto tiene una alta diversidad léxica $D = 91,36$ (media grupal, $D = 86$) y una densidad también alta 0,46 (media grupal, 0,41 con baja dispersión). Por último, también hay que destacar otra característica

léxica del texto (aunque no ha sido objeto del estudio): el uso de términos poco frecuentes y especializados, tales como “arreglo instrumental” o “tablatura”, lo que también refuerza su calidad léxica general. Por lo tanto, puede decirse que en este ejemplo las medidas de diversidad y densidad léxicas complementan los rasgos semánticos dados por el buen uso de sustantivos y verbos propios del género, contribuyendo así a la calidad del texto en su conjunto.

El texto en (5) es explicativo, con una extensión de 181 palabras, un puntaje de desempeño global de 17 (70%) y, si bien es de menor calidad general, como se comentará, reúne algunos de los rasgos de verbos y sustantivos que interesa destacar: tiene una proporción muy alta de sustantivos abstractos y genéricos, 0,51 (media grupal, 0,27), con términos como “grupo”, “evento”, “mundo”, “gente”, y también de verbos existenciales, 0,57 (media grupal, 0,44), con preferencia por el verbo ser, que se repite bastante. Estos tipos de verbos y sustantivos, al contrario de lo que pasaba con el texto anterior, le dan a este texto un carácter distanciado, inespecífico y cuyas aseveraciones parecen cobrar cierta validez debida a la generalización. La diversidad léxica es más baja que en el caso del texto anterior, $D = 81,6$ (lo que se puede intuir a simple vista, producto de la abundante repetición), y se ubica por debajo de la media grupal, que es $D = 91,08$. Cabe recordar que las medias en diversidad léxica (tabla 6) suelen ser un poco más altas en los textos explicativos (aunque sin diferencias estadísticamente significativas reportadas en este estudio) por las características del género, pero también, a modo de hipótesis, porque estos textos son más breves en todos los grupos (la brevedad podría influir porque la proporción se obtiene sobre el total de palabras del texto). Y si bien la fórmula empleada para medir la diversidad corrige, en teoría, los problemas derivados de la extensión de los textos para este tipo de cálculos, tal vez este factor sigue interfiriendo en la medida. La densidad, en tanto, similar a la del texto anterior, es de 0,44 y se ubica por sobre la media grupal, que es 0,42. Este texto no parece especialmente denso informativamente (si bien, como se ha explicado antes, una alta densidad léxica no es algo que se manifieste de modo evidente para el oyente o lector), pero utiliza muchos sustantivos y verbos en una corta extensión, lo que hace que la proporción resulte más alta.

4. Conclusiones y discusión

La presente investigación se propuso indagar en el modo en que estudiantes de tres niveles escolares y dos grupos socioeconómicos utilizan variados recursos léxicos en sus textos narrativos y explicativos escritos. En particular, se analizaron medidas de diversidad y densidad léxica, abstracción nominal y tipos de procesos, con la finalidad de observar el efecto de tres factores en estas variables: la edad/nivel escolar de los sujetos, el grupo socioeconómico de pertenencia y el género discursivo. Si bien todos estos factores parecen incidir de alguna manera en las variables analizadas, los resultados muestran que son la edad/nivel y el género discursivo los factores más relevantes.

Respecto de las *medidas para examinar y cuantificar la complejidad léxica*, es interesante comprobar, como se adelantaba en la introducción, que ninguna de ellas por separado puede dar cuenta adecuadamente de este fenómeno. Como se observó, aunque tenga un alto índice de diversidad y/o de densidad léxica, un texto puede ser deficitario en términos de su despliegue léxico general (porque, por ejemplo, utiliza una estrategia enumerativa, lo que le confiere gran densidad pero debilidad en otros aspectos) y a la inversa, bajos indicadores de diversidad o densidad pueden registrarse en textos de gran calidad léxica. Por eso estos resultados resaltan la necesidad de utilizar diversas medidas y observar con cautela el modo en que interactúan con los diversos factores implicados en la medición.

En cuanto a la *función que pueden tener las medidas* utilizadas, concluimos que la diversidad y ciertos usos que se hace de sustantivos y verbos en relación con las características de los géneros son medidas pertinentes para observar y comprender el desarrollo léxico. Además de los sustantivos y verbos, probablemente sea también interesante analizar todas las palabras de contenido y dar especial énfasis a los adjetivos, que al parecer pueden ofrecer información valiosa acerca del desarrollo léxico tardío (Cutillas y Tolchinsky, 2017). Por otro lado, la densidad léxica y el uso de verbos y sustantivos, particularmente en cuanto a sus grados de abstracción, resultaron también útiles para comprender y comparar los distintos géneros discursivos. Por último, estos resultados permiten afirmar que ninguna de las medidas contempladas en el estudio se ven afectadas significativamente por el origen socioeconómico, pero que, si observamos el modo en que este factor interactúa con la edad/nivel, se obtienen diferencias en la densidad léxica y el uso de los sustantivos. Es decir, estas parecen ser variables que no dan cuenta de diferencias en grupos de una misma edad, pero que muestran el efecto acumulativo de un determinado tipo de educación.

En cuanto al *modo en que los tres factores* implicados (edad/nivel, grupo socioeconómico y género discursivo) inciden en las medidas léxicas, se concluye, en primer lugar, que la edad/nivel incide en la diversidad léxica, en el uso de sustantivos abstractos y genéricos y en el de verbos de procesos existenciales. En particular, los estudiantes mayores (y más claramente en el establecimiento del GS alto) escriben textos más diversos léxicamente y utilizan más sustantivos abstractos/genéricos y más verbos de procesos existenciales en sus textos narrativos y explicativos. Aunque no contamos con datos de este tipo provenientes de investigaciones realizadas en nuestro país, estos hallazgos son, además de previsibles, consistentes con resultados de otras investigaciones en sujetos de habla hispana (Hess, 2013).

El factor grupo socioeconómico, como se ha explicado, solo interactúa significativamente con la edad/nivel respecto de la densidad léxica. Es decir, los estudiantes de ambos establecimientos, como se ha dicho, manifiestan un perfil evolutivo distinto respecto de la densidad: en el GS medio-bajo todos los grupos de edad tienen una densidad léxica similar, mientras que en el GS alto los estudiantes de séptimo tienen un desempeño más bajo que luego se ho-

mologa al nivel de los demás. Más allá de este resultado particular, es importante comentar dos cosas respecto de las diferencias entre los dos contextos socioeconómicos.

Lo primero es que las medidas aquí analizadas (diversidad, densidad, uso pertinente de verbos y sustantivos en los géneros) no dan cuenta de diferencias en los usos léxicos entre ambos tipos de población, pero ello no quiere decir que no las haya. Puede ser que existan diferencias y es necesario explorar y operacionalizar nuevas medidas, tal vez relacionadas con la sofisticación, con el uso de lenguaje especializado o de palabras de baja frecuencia para poder verificarlas empíricamente.

Lo segundo es destacar que, si bien los dos tipos de establecimientos se comportan de manera similar en cuanto a las variables analizadas, las diferencias entre ellos se acrecientan en los cursos superiores. En otras palabras, los estudiantes de séptimo del GS medio-bajo tienen desempeños léxicos similares y superiores a los de sus pares del GS alto, pero en los cursos superiores los estudiantes del GS medio tienden a mantener el desempeño, mientras que los del GS alto tienden a mejorarlo. Sería interesante, por lo tanto, indagar en los factores (tales como el tipo de escolaridad, la competencia lectora o la alfabetización familiar) que podrían estar incidiendo en que el léxico tardío siga manifestando cambios evolutivos significativos en un contexto social y no en el otro.

En tercer lugar, el género discursivo determina cambios significativos en los usos que hacen los estudiantes de ambos establecimientos en cuanto a densidad léxica, sustantivos abstractos/genéricos y verbos mentales y existenciales. Específicamente, los estudiantes mayores de ambos establecimientos (y también el grupo de séptimo de GS medio-bajo) escriben sus textos explicativos con más densidad que sus textos narrativos. También, todos los grupos del GS alto utilizan más sustantivos abstractos/genéricos, más verbos existenciales y menos verbos mentales en sus textos explicativos. Acerca de la incidencia de los géneros, estos resultados no coinciden con los de otro estudio realizado en Chile (Gómez-Vera y otros, 2016), en que se encontró que, a diferencia de lo observado aquí, la diversidad léxica sí varía significativamente en los distintos géneros. Ambos estudios no son comparables, porque las edades de los estudiantes son distintas y también lo es la naturaleza de las tareas de escritura solicitadas (en el otro estudio cada uno de los géneros solicitados tenía un tema distinto, lo cual hace más complejo comparar desempeños léxicos entre los géneros). Son necesarios, por lo tanto, más estudios en nuestro país que permitan seguir comprendiendo el valor de las medidas léxicas utilizadas y su relación con los factores evolutivos, socioculturales y discursivos.

A partir de los resultados aquí expuestos, hemos observado que los recursos léxicos como la diversidad, la densidad, la abstracción nominal y los tipos de procesos (entre otras medidas que cabría explorar, de manera conjunta y situada discursivamente) cumplen un importante papel en la construcción de los textos escritos y en el progresivo dominio de

los géneros. Y, puesto que este dominio es una pieza clave en los desarrollos epistémicos y lingüísticos, postulamos la necesidad de encontrar más y mejores estrategias de enseñanza que favorezcan en los estudiantes desempeños léxicos competentes y funcionales a las necesidades de los textos.

5. Bibliografía citada

- BEHRENS, Heike, 2008: *Corpora in language acquisition research*, Amsterdam: John Benjamins.
- BERMAN, Ruth A., y Bracha NIR, 2010: "The lexicon in writing-speech-differentiation. Developmental perspectives", *Written Language & Literacy* 13 (2), 183-205.
- BERMAN, Ruth, y Dorit RAVID, 2009: "Becoming a literate language user. Oral and written text construction across adolescence" en David OLSON y Nancy TORRANCE (eds.): *The Cambridge Handbook of Literacy*, Cambridge: Cambridge University Press, 92-111.
- BERMAN, Ruth, y Dan SLOBIN, 1994: *Relating Events in Narrative. A Crosslinguistic Developmental Study*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- BERMAN, Ruth, y Ludo VERHOEVEN, 2002: "Cross-linguistic perspectives on the development of text production abilities: speech and writing", *Written Language and Literacy* 5, 1-44.
- BERMAN, Ruth, Ronit NAYDITZ y Dorit RAVID, 2011: "Linguistic diagnostics of written texts in two school-age populations", *Written Language & Literacy* 14 (2), 161-187.
- BOGAARDS, Paul, y Batia LAUFER, 2004: *Vocabulary in a second language: selection, acquisition, and testing*, Amsterdam: John Benjamins.
- CASTELLA, Josep M., 2001: *La Complexitat lingüística en el discurs oral i escrit: densitat lèxica, composició oracional i connexió textual*. Tesis de doctorado, Universidad Pompeu Fabra.
- CUTILLAS, Laia, y Liliana TOLCHINSKY, 2017: "Use of adjectives in Catalan: A morphological characterization in different genres and modes of production through school-age development", *First Language* 37 (1), 58-82.
- DEVESCOVI, Antonella, María Cristina CASELLI, Daniela MARCHIONE, Patrizio PASQUALETTI, Judy REILLY y Elizabeth BATES, 2005: "A crosslinguistic study of the relationship between grammar and lexical development", *Journal of Child Language* 32 (4), 759-786.
- DOCKRELL, Julie, y David MESSER, 2004: "Lexical acquisition in the early school years" en Ruth BERMAN (ed.): *Language Development across Childhood and Adolescence. Trends in Language Acquisition Research Series (3)*, Amsterdam, Netherlands: John Benjamins Publishing Company, 35-52.

FAUL, Franz, Edgar ERDFELDER, Albert-George LANG y Axel BUCHNER, 2007: "G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences", *Behavior Research Methods* 39, 175-191.

FUSTÉ-HERRMANN, Belinda, Elaine R. SILLIMAN, Ruth H. BAHR, Kyna S. FASNACHT y Jane FEDERICO, 2006: "Mental state verb production in the oral narratives of english - and spanish speaking preadolescents: an exploratory study of lexical diversity and depth", *Learning Disabilities Research & Practice* 21 (1), 44-60.

GALLOWAY, Emily Philips, y Paola UCCELLI, 2015: "Modeling the relationship between lexico-grammatical and discourse organization skills in middle grade writers: insights into later productive language skills that support academic writing", *Reading & Writing* 28 (6), 797-828.

GÓMEZ-VERA, Gabriela, Carmen SOTOMAYOR, Percy BEDWELL, Ana María DOMÍNGUEZ y Elvira JELDREZ, 2016: "Analysis of lexical quality and its relation to writing quality for 4th grade, primary school students in Chile", *Reading and Writing* 29 (7), 1317-1336.

HALLIDAY, Michael, 1985: *An Introduction to Functional Grammar*, London: Edward Arnold.

HESS, Karina, 2013: "Desarrollo léxico en la adolescencia: un análisis de sustantivos en narraciones orales y escritas", *Actualidades en Psicología* 27 (115), 113-127.

HUDSON, Richard, 2009: "Measuring maturity" en Roger BEARD, Debra MYHILL, Jeni RILEY y Martin NYSTRAND (eds.): *The Sage handbook of writing development*, Londres: SAGE, 349-362.

JARVIS, Scott, 2013a: "Capturing the Diversity in Lexical Diversity", *Language Learning* 63 (1), 87-106.

JARVIS, Scott, 2013b: "Defining and measuring lexical diversity" en Scott JARVIS y Michael DALLER (eds.): *Vocabulary Knowledge. Human ratings and automated measures*, Amsterdam: John Benjamins, 13-43.

JOHANSSON, Victoria, 2008: "Lexical diversity and lexical density in speech and writing: a developmental perspective", *Working Papers* 53, 61-79.

KARMILOFF-SMITH, Anette, 1979: *A Functional Approach to Child Language*, Cambridge: Cambridge University Press.

KARMILOFF-SMITH, Anette, 1992: *Beyond Modularity: A Developmental Perspective on Cognitive Science*, Cambridge: MIT Press/Bradford Books.

KYLE, Kristopher, y Scott CROSSLEY, 2016: "The relationship between lexical sophistication and independent and source-based writing", *Journal of Second Language Writing* 34, 12-24.

MACWHINNEY, Brian, 2000: *The CHILDES Project: tools for analyzing talk*, Mahwah, NJ: Lawrence Erlbaum Associates.

MALVERN, David, Brian RICHARDS, Ngoni CHIPERE y Pilar DURÁN, 2004: *Lexical diversity and language development. Quantification and Assessment*, New York: Palgrave Macmillan.

MILTON, James, y Jeanine TREFFERS-DALLER, 2013: "Vocabulary size revisited: the link between vocabulary size and academic achievement", *Applied Linguistics Review* 4 (1), 151-172.

McKoon, Gai, y Roger RATCLIFF, 2016: "Adults with Poor Reading Skills: How Lexical Knowledge Interacts with Scores on Standardized Reading Comprehension Tests", *Cognition* 146, 453-469.

NEUENDORF, Kimberly, 2002: *The content analysis guidebook*, Thousand Oaks, CA: Sage.

NIPPOLD, Marilyn. A., 2007: *Later language development: the school-age and adolescent years*, Austin, TX: Pro-Ed.

NIPPOLD, Marilyn, Susan HEGEL, Mackay Moore SOHLBERG e Ilsa SCHWARZ, 1999: "Defining Abstract Entities: Development in Pre-Adolescents, Adolescents, and Young Adults", *Journal of Speech, Language, and Hearing Research* 42 (2), 473-481.

PERFETTI, Charles, 2010: "Decoding, vocabulary, and comprehension: the golden triangle of reading skill" en Margaret G. McKEOWN y Linda LUCAN: *Bringing reading research to life*, Nueva York: The Guilford Press, 291-303.

QUIROGA, Riva, 2015: *El despliegue léxico en estudiantes chilenos en distintos géneros discursivos*. Tesis de maestría, Pontificia Universidad Católica de Chile.

RAKISON, David, y Yevdokiya YERMOLAYEVA, 2011: "How to identify a domain-general learning mechanism when you see one", *Journal of Cognition and Development* 12, 134-153.

RAVID, Dorit, 2005: "Emergence of linguistic complexity in written expository texts: Evidence from later language acquisition" en Dorit RAVID y Hava Bat-Zeev SHYLDKROT (eds.): *Perspectives on language and language development*, Dordrecht: Kluwer, 337-355.

RAVID, Dorit, 2006: "Semantic development in textual contexts during the school years: Noun scale analyses", *Journal of Child Language* 33 (4), 791-821.

READ, John, 2000: *Assessing vocabulary*, Cambridge: Cambridge University Press.

ROCHE, Thomas, y Michael HARRINGTON, 2013: "Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting", *Language Testing in Asia* 3 (12), 1-13.

SCOTT, Cheryl M. y Jennifer WINDSOR, 2000: "General language performance measures in spoken and written narrative and expository discourse of school age children with language learning disabilities", *Journal of Speech, Language, and Hearing Research* 43, 324-339.

SERRAT, Elisabet, Mónica SANZ TORRENT y Aurora BEL, 2004: "Aprendizaje léxico y desarrollo de la gramática: vocabulario verbal, aceleración morfológica y complejidad sintáctica", *Anuario de Psicología* 35 (2), 221-234.

SNOW, Catherine, y Paola UCCELLI, 2009: "The challenge of academic language" en David OLSON y Nancy TORRANCE (eds.): *The Cambridge Handbook of Literacy*, Cambridge: Cambridge University Press, 112-132.

STEMLER, Steven, 2004: "A comparison of consensus, consistency, and measurement approaches to estimating inter-rater reliability", *Practical Assessment, Research & Evaluation* 9 (4) [<http://www.pareonline.net/getvn.asp?v=9&n=4>].

SUN, Lei, y Marilyn A. NIPPOLD, 2012: "Narrative writing in children and adolescents: Examining the literate lexicon", *Language, speech, and hearing services in schools* 43 (1), 2-13.

TOLCHINSKY, Liliana, 2004: "The nature and scope of later language development" en Ruth BERMAN (ed.): *Language development across childhood and adolescence*, Amsterdam: John Benjamins, 233-247.

TOLCHINSKY, Liliana, María Antonia MARTÍ y Anna LLAURADÓ, 2010: "The growth of the written lexicon in Catalan. From childhood to adolescence", *Written Language & Literacy* 13 (2), 206-235.

VALENCIA, Alba, 1994: *El léxico de los estudiantes de 4º año de Educación Media. Centro de interés "Procesos mentales"*, Santiago: CPEIP.

VALENCIA, Alba, 1997: "Disponibilidad léxica. Muestreo y estadísticos", *Onomázein* 2, 197-226.

VALENCIA, Alba, y Max ECHEVERRÍA, 1999: *Disponibilidad léxica en estudiantes chilenos*, Santiago: Ediciones de la Universidad de Chile/Universidad de Concepción.

VALENCIA, Alba, y Violy HERMANSEN, 1985: *El léxico de los estudiantes de cuarto medio. El verbo: materiales*, Santiago: Universidad de Chile.