

Viedma 8 y 9 de noviembre de 2018



Sintaxis y comprensión. Protocolo para una investigación

María Silvia Alasio

CURZA - Universidad Nacional del Comahue. Argentina

Proyecto de Investigación 04/V116. "Enseñanza de la gramática y comprensión de textos". (2018-2021)

Directora: Gabriela Patricia Comezaña

RESUMEN

A partir de la hipótesis de que el deterioro de las habilidades de lectura y escritura en los egresados de la escuela secundaria está causada, entre otras cosas, por el abandono de la enseñanza de la gramática y la reflexión metalingüística en la escuela, el presente trabajo pone en relación distintas perspectivas teóricas que apoyan esa hipótesis y muestra un primer análisis de datos concretos obtenidos a partir de una evaluación diagnóstica. Se presentan los conceptos básicos de la adquisición tardía del lenguaje (Nippold 2004), el rol del componente sintáctico dentro de un modelo multicomponential de la comprensión de textos (Abusamra et al. 2010) y la contribución de la sintaxis en la alfabetización (Scott 2004). Luego se analizan los resultados de la primera evaluación diagnóstica implementada por el equipo del proyecto en relación con los conceptos planteados en el marco teórico.

Palabras clave: Gramática; Adquisición tardía; Comprensión de textos; Educación secundaria.

Viedma 8 y 9 de noviembre de 2018

Introducción

El proyecto de investigación en el que se enmarca este trabajo se propone abordar cuali y cuantitativamente la relación entre la enseñanza de la gramática en la escuela media y la universidad y las habilidades de comprensión textual de los alumnos.

Se parte del supuesto de que el deterioro de las habilidades de lectura y escritura de los estudiantes se debe, en alguna medida, al abandono de la enseñanza de la gramática y la reflexión metalingüística en la escuela.

El primer paso consiste en encontrar una correlación entre el dominio en el componente sintáctico y las habilidades de comprensión de textos y luego, de ser posible, establecer si esa correlación implica o no causalidad. Si se encuentran estas relaciones, uno de los objetivos del proyecto es diseñar un protocolo de entrenamiento para mejorar aquellas habilidades sintácticas que se hayan detectado como pertinentes para una lectura comprensiva.

Marco teórico

Desarrollo tardío del lenguaje

Los estudios psicolingüísticos sobre el desarrollo del lenguaje han demostrado que este es un proceso que va mucho más allá de los primeros años de vida para extenderse, de modo gradual y prolongado, hasta la adultez.

De acuerdo con Marilyn Nippold, los trabajos sobre el desarrollo tardío del lenguaje se han enfocado en tres cuestiones:

- a) ¿Qué es lo que se adquiere tardíamente?
- b) ¿Qué evidencias se pueden tener de ese desarrollo tardío?
- c) ¿Qué factores contribuyen a que ese desarrollo se produzca?

Con respecto a la primera cuestión, la naturaleza y sustancia de lo que se adquiere tardíamente puede englobarse en el concepto de *alfabetización*, que Nippold define como “la adquisición de las herramientas del lenguaje oral y escrito necesarias para la comunicación en situaciones formales, tales como la escuela o el trabajo”¹.

La alfabetización incluye la adquisición de un “lexicón alfabetizado”, un diccionario mental de palabras complejas y de baja frecuencia, interrelacionadas en una red semántica elaborada. Este “sistema rico y poderoso” se expande sostenidamente durante los años de la escolarización y continúa en la adultez. De hecho, no parece haber límite para el aprendizaje del léxico a lo largo de toda la vida.

Otro rasgo de la alfabetización es la aparición de significados multidimensionales y simbólicos. Dice Nippold: “aunque el razonamiento analógico y la comprensión de metáforas aparecen tempranamente, la comprensión avanzada del lenguaje figurativo sólo se alcanza en la adultez”.

También forman parte de este lexicón alfabetizado los verbos metacognitivos de baja frecuencia, como *asumir*, *darse cuenta*, *inferir*, *simbolizar*, etc.

Además del lexicón, las etapas tardías de desarrollo del lenguaje incluyen logros sintácticos. Uno de ellos es la habilidad de usar estructuras poco frecuentes, como la voz pasiva, común en la escritura científica. Esta capacidad se acrecienta gradualmente en la producción de lenguaje escrito a lo largo de la escolarización hasta la adultez.

Otras estructuras sintácticas que evidencian un aumento gradual en el uso durante los últimos años de la escuela y los primeros de la universidad son las que incluyen subordinación en todas sus formas: relativas, especialmente las de objeto (por ejemplo: *El pájaro que el gato atrapó, logró liberarse.*), relativas libres (por

¹ Nippold, M.A. (2004) “Research on later language development: International perspectives”. Traducción propia.

Viedma 8 y 9 de noviembre de 2018

ejemplo: *El que atrapó al pájaro fue el gato.*), cláusulas no finitas (por ejemplo: *Habiendo subido al árbol, el gato intentó atrapar al pájaro.*), construcciones con pasado perfecto o “pasado dentro del pasado” (por ejemplo: *El gato que había subido al árbol trató de atrapar al pájaro*), los auxiliares modales (*poder, deber, soler*) y las conjunciones adverbiales de baja frecuencia (como: *mientras tanto, más aún, consecuentemente*). Todas estas construcciones dan como resultado la producción de oraciones más largas, densas y complejas tanto en el discurso oral como en el escrito.

Dado que el modelo gramatical que sustenta este proyecto asume que la estructura oracional es una proyección del léxico, tanto las cuestiones referidas a la adquisición tardía del lexicón alfabetizado como a los logros sintácticos en etapas avanzadas de escolarización resultan relevantes.

Modelo multicomponencial de la comprensión de textos

El modelo multicomponencial de la comprensión de textos tal como es presentado en los materiales que conforman el *Test Leer para Comprender* (Abusamra et al. 2010), distingue tres grandes áreas o dominios: *Contenido, Elaboración y Metacognición*. El área de comprensión relacionada con la estructura sintáctica se incluye en el ámbito de la *Elaboración*, junto con la cohesión, las inferencias, la jerarquía del texto y los modelos mentales.

En este modelo, las habilidades de comprensión de textos relacionadas con la estructura sintáctica incluyen la capacidad de procesar los nexos gramaticales de las palabras dentro de la oración y de las oraciones entre sí. Una oración es definida como la unidad discursiva mínima con la que se construyen los textos. Las habilidades de alcance oracional se refieren sobre todo a la capacidad de interpretar las diferencias de significado debidas a diferencias estructurales. También forman parte del ámbito oracional las habilidades relacionadas con la cohesión (en caso de que aparezcan pronombres reflexivos, clíticos duplicados, conjunciones coordinantes o subordinantes, elipsis, etc.) y algunas inferencias o presupuestos (por ejemplo, los argumentos implícitos de algunos verbos).

Estas últimas habilidades también se ponen en juego para la comprensión de unidades mayores, como el texto o los párrafos (a los que podríamos llamar “micro textos”). En este caso, se encontrarían nuevamente la cohesión y las inferencias, junto con la jerarquía del texto y los modelos mentales.

Entonces, si definimos la oración como la unidad mínima con la que se construyen los textos, parece lógico plantear que este sea el primer objeto para estudiar las estrategias de comprensión. Se trata de un objeto privilegiado: su extensión es limitada, es el bloque constructivo de todas las unidades mayores y su estructura está fuertemente reglada por la sintaxis, el componente lingüístico más autónomo.

El componente sintáctico en la comprensión de textos

Cheryl Scott, en su artículo “Syntactic contributions on literacy learning”² se pregunta por qué la sintaxis, o más precisamente la gramática oracional, no figura entre las preocupaciones centrales cuando se analizan las dificultades en la comprensión y producción de textos. Sin embargo, las operaciones sintácticas son centrales en los estudios sobre adquisición del lenguaje, sobre todo en las descripciones de los trastornos específicos del lenguaje. Se encuentra entonces la siguiente paradoja: uno de los temas relevantes de los estudios psicolingüísticos es la adquisición del lenguaje y dentro de ésta, las patologías en el desarrollo. En las etapas tempranas se ha comprobado que la sintaxis juega un papel central en muchos de esos déficits. Sin embargo, al estudiar déficits tardíos, como los problemas en la comprensión y producción de textos, se presupone que la sintaxis pierde por completo ese rol central, o más bien, que la sintaxis no tienen ningún rol en esos trastornos.

En los niños con diagnóstico de trastorno específico del lenguaje, la sintaxis siempre ha sido vista como un déficit central (Leonard, 1998; Paul, 2001; Rice, 2000), y estos mismos niños tienen un alto riesgo de

² Scott, Cheryl (2004) “Syntactic contributions on literacy learning” en *Handbook of Language and Literacy*. Traducción de Paola Galbarini para la cátedra de Psicolingüística y Adquisición del lenguaje, CURZA-UNCo, 2018

Viedma 8 y 9 de noviembre de 2018

padecer desórdenes de lectura y escritura (Catts, Fey, Zhang, & Tomblin, 1999; Scarborough, 1991). ¿Por qué, entonces, la sintaxis, parece haber quedado en segundo plano en las agendas de desarrollo e instrucción de alfabetización en lecto-escritura?³

¿No es un trastorno del lenguaje la dificultad para comprender y producir textos? ¿Solamente tiene que ver con cuestiones más generales como conocimiento de mundo, memoria de corto y largo plazo, capacidad simbólica o conocimiento léxico (adquisición de vocabulario)?

Los textos están “hechos de” oraciones. Si bien el objetivo al leer un texto es saber de qué se trata, descubrir y recordar un tema o una idea, acordar o no con una postura, y no “comprender oraciones”, son las oraciones las que “arman” esos significados globales. Resulta relevante entonces estudiar la “forma” de esos constructos y qué características ayudan o entorpecen la operación cognitiva de construcción de significados.

La mayoría de los modelos de comprensión y producción de textos asumen que el procesamiento de la gramática oracional (entendiendo gramática en un sentido amplio que incluye la sintaxis y los niveles de interfaz léxico-sintaxis y sintaxis-semántica) se hace en forma automática o inconsciente.

La caracterización del procesamiento sintáctico como automático es relevante desde el punto de vista teórico y permite explicar adecuadamente ciertos rasgos de la adquisición del lenguaje en las primeras etapas del desarrollo. Sin embargo, las habilidades de lectura y escritura, relacionadas con etapas tardías de adquisición del lenguaje, parecen requerir la desautomatización del conocimiento gramatical para poder controlarlo y utilizarlo en la producción de significados.

[...] si consideramos la lectura y la escritura, la distinción entre un nivel de procesamiento automático y uno más consciente se vuelve importante. Se cree que ambas actividades, sobre todo en las etapas de desarrollo, a veces requieren un mayor nivel de conciencia que, presumiblemente, incluiría algún tipo de conciencia sintáctica.⁴

La mayor o menor complejidad sintáctica de las oraciones es una cuestión que ha sido considerada ampliamente en la bibliografía relacionada con la comprensión de textos. Usualmente se consideran oraciones más complejas las que incluyen algunos de los siguientes rasgos:

- Estructuras poco usuales, esto incluye las formas más literarias y los órdenes de palabras marcados, como las oraciones hendidas y pseudohendidas
- Negación, sobre todo las negativas no explícitas y las literarias
- Sintagmas nominales extensos, frecuentemente en posición de sujeto
- Nominalizaciones
- Elipsis, sobre todo cuando el referente elidido está alejado
- Palabras funcionales ambiguas
- Ambigüedades estructurales
- Estructuras que recargan la memoria de trabajo

Estos rasgos han sido tenidos en cuenta en la elaboración de la primera prueba diagnóstica del proyecto

Metodología

En una primera etapa, el proyecto desarrollará los instrumentos de diagnóstico adecuados a sus intereses. Se tomará como referencia el *Test Leer para Comprender* ya mencionado y las consideraciones acerca de las marcas o indicios de desarrollo lingüístico tardío así como los índices de complejidad oracional resumidos anteriormente.

³ Idem anterior.

⁴ Op. Cit. p.16

Viedma 8 y 9 de noviembre de 2018

Esos instrumentos de diagnóstico se aplicarán a grupos seleccionados de alumnos de nivel medio de Viedma⁵ y su zona de influencia, con el propósito de evaluar cuali y cuantitativamente la comprensión textual y los conocimientos gramaticales.

Se trata, por lo tanto, de un diseño de corte transversal destinado a distinguir lectores (o comprendedores) buenos y malos dentro de individuos de la misma edad y a correlacionar las habilidades de comprensión con el dominio de la sintaxis.

La evaluación diagnóstica

La primera versión de la evaluación diagnóstica (ver Anexo) parte de la lectura de un texto para responder luego a un conjunto de 16 preguntas de selección múltiple, emparejamiento, verdadero / falso y respuesta breve. Una de las preguntas, la número 7, se refiere a un dato negativo y otra, la número 15, indaga la comprensión global del texto⁶.

Para el diseño de este instrumento se tuvo en cuenta que las tareas prototípicas usadas en este tipo de pruebas son el completamiento y los juicios sobre la corrección.

Implementación: logros y dificultades

Para conocer su calidad, se puso a prueba el instrumento de diagnóstico en tres muestras de alumnos de distintos establecimientos educativos, dos de la ciudad de Viedma y uno de San Javier (una pequeña localidad semi-rural ubicada a 45 km. de Viedma).

El total de la muestra fue de 59 individuos, 24 varones y 35 mujeres, todos alumnos que cursan el primer año de la escuela media y tienen entre 13 y 14 años de edad.

El test se implementó durante el horario de clase y se presentó a los alumnos como una actividad de evaluación diagnóstica. Todos los alumnos completaron la tarea, y las docentes a cargo del curso tomaron notas acerca del comportamiento durante el desarrollo, el mayor o menor apego a la tarea, las preguntas que surgieron y otras situaciones que les parecieron importantes. Todavía no se ha realizado el análisis de estos registros pero no surgió ninguna situación que impidiera llevar a cabo la evaluación.

Algunos alumnos plantearon que la actividad era muy larga en comparación con las que suelen hacer en clase. Otras dificultades que manifestaron fue el desconocimiento de algunas palabras presentes en el texto o en las consignas (por ejemplo, muchos desconocían la palabra *prólogo*).

A partir de los resultados se analizará, por un lado, la calidad del instrumento de diagnóstico diseñado, para su ajuste posterior; por otro, se evaluará el desempeño de los alumnos, con el fin de extraer algunas conclusiones iniciales.

Análisis de los datos

El primer procesamiento de la muestra consistió en asignar un código o referencia a cada uno de los individuos para poder tratar los datos preservando la identidad.

Se consideraron como datos, además de la referencia mencionada, el sexo, la edad y las respuestas a las preguntas. Para todas las preguntas sobre procesamiento sintáctico o gramatical se consideró un punto por respuesta correcta y cero por respuesta incorrecta. Para la última pregunta, la que indaga la comprensión global, ya que pide emparejar un número de párrafo con un posible título, la valoración se estableció en relación con el número de aciertos en el emparejamiento:

⁵ La ciudad de Viedma es la capital de la provincia de Río Negro, en la región sur de Argentina.

⁶ Dado que ya se ha presentado en otro trabajo, no se incluye aquí la descripción en detalle de la evaluación diagnóstica.

Viedma 8 y 9 de noviembre de 2018

0	Mala
1	Regular
2	Buena
3	Muy buena
5	Excelente

No se tuvo en cuenta la posibilidad de 4 aciertos ya que en ese caso, por descarte, queda identificada la quinta opción. Las valoraciones hacen referencia a la comprensión.

Tratamiento estadístico de los primeros resultados⁷

Se hizo el análisis del comportamiento de los ítems a partir de dos indicadores fundamentales: el nivel de dificultad que presenta y el poder para discriminar entre buenos y malos comprendedores.

El nivel de facilidad / dificultad se establece a partir de un cociente entre la cantidad de respuestas correctas y el total de individuos de la muestra. El valor que se obtiene estará entre 0 y 1, cuanto más cerca de 0 más difícil será la pregunta y cuanto más cerca de 1 más fácil.

Dado que se trata de un diagnóstico para establecer los niveles de dominio de un tema determinado, son preferibles los valores intermedios. Las preguntas con un índice superior a 0.86 son consideradas muy fáciles y aquellas con un índice menor a 0.32 son consideradas muy difíciles. Este valor se usa también para ordenar las preguntas de una evaluación en un grado creciente de dificultad, que es el modo en que se obtienen mejores desempeños. Sin embargo, no siempre se puede respetar esto cuando se trabaja con un texto ya que las preguntas normalmente siguen el orden de la información en el texto.

Los índices de facilidad / dificultad para las preguntas del diagnóstico de prueba son los siguientes:

Preg.	1	2	3	4	5a	5b	6	7	8	9	10	11	12	13	14
Índice	0.695	0.91	0.89	0.71	0.55	0.39	0.86	0.18	0.54	0.61	0.62	0.47	0.35	0.44	0.52

Estos resultados muestran que el grado de facilidad / dificultad del instrumento de diagnóstico es adecuado, ya que la mayoría de las preguntas (47%) son de dificultad media. Un 27% es medianamente difícil, no hay ninguna pregunta medianamente fácil y un 20% es altamente fácil. Solamente hay una pregunta con un valor que indica una alta dificultad: la pregunta 7.

En cuanto al índice de discriminación se trata de un valor que mide la capacidad de las preguntas para distinguir entre buenos y malos comprendedores. Hay varias maneras de calcular este índice. En este primer análisis no se siguió el usado por Abusamra y otros (2010) sino que se prefirió el método del cálculo del coeficiente biserial, dado que permite tener en cuenta todos los individuos encuestados y todas las respuestas, algo posible dado el tamaño todavía reducido de la muestra.

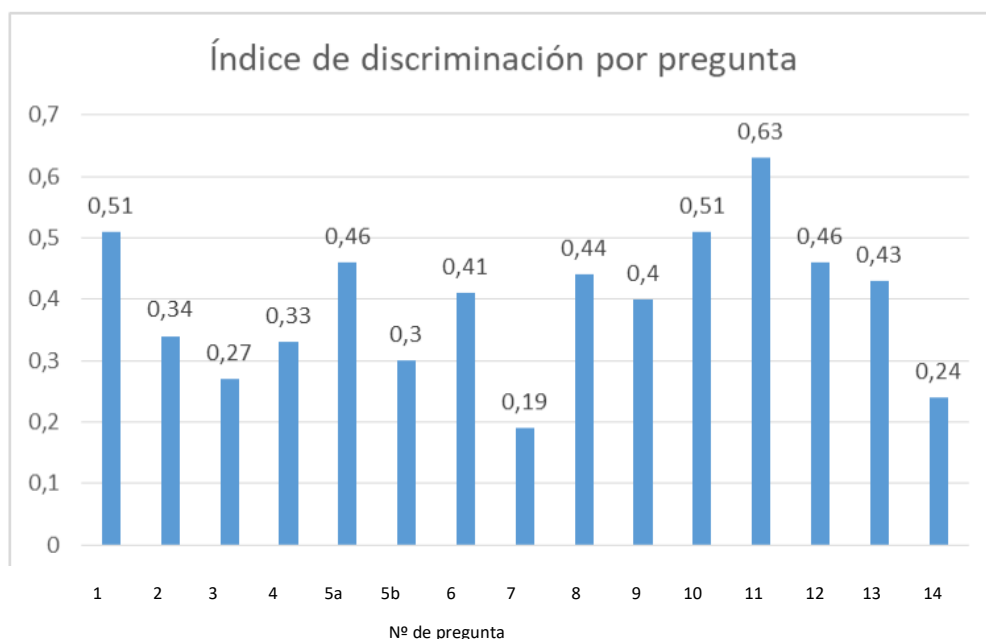
La interpretación del coeficiente se hace sobre la base de una tabla:

r_{pbis}	Poder de discriminación del ítem
$> 0,39$	Excelente
$0,30 - 0,39$	Bueno
$0,20 - 0,29$	Regular
$0,00 - 0,19$	Pobre
$< - 0,01$	Pésimo

⁷ En este punto se tiene en cuenta parte del análisis realizado por la integrante del proyecto, Prof. Gabriela Comezaña como trabajo práctico para un curso de Estadística Experimental.

Viedma 8 y 9 de noviembre de 2018

A partir del cálculo de este coeficiente se obtuvo que el poder de discriminación de las preguntas está entre bueno y excelente: un 60% de las preguntas tiene un índice de discriminación excelente, un 20% es bueno, un 13% es regular y solo una pregunta tiene un índice bajo de discriminación. Se trata de la misma que tiene el mayor grado de dificultad: la pregunta 7.



Del análisis comparativo de estos dos índices surgieron conclusiones muy iluminadoras acerca de las preguntas que componen el diagnóstico. Por ejemplo, de las tres preguntas altamente fáciles no todas se comportan igual ante el poder discriminatorio, de modo que sólo una resulta inaceptable, mientras que otra debería ser mantenida y una tercera necesita mejorarse.

El resultado más llamativo es el de la pregunta 7, tiene al grado más alto de dificultad y el más bajo de poder discriminatorio. La pregunta es la siguiente:

7. *¿Por qué el autor dice que las personas auténticas son malas? (línea 20)*

- a. *porque lastiman sin querer.* b. *porque son una amenaza.*
 c. *porque se hacen solas* d. *El texto no dice eso*

Esta pregunta plantea una falsa relación causal, es decir, algo que el texto no dice. Hay varias interpretaciones posibles para rastrear la falla de esta pregunta. Una de ellas es que la dificultad es doble: por un lado se trata de una pregunta por un "dato negativo" y por otro, ese dato negativo está planteado como una relación causal, una de las relaciones intra clausales que de acuerdo con Nippold aparecen en el desarrollo tardío del lenguaje. Otra es que lo que aquí falló es el componente metacognitivo (en el modelo multicomponencial de la comprensión), en aquel módulo que permite detectar errores o incongruencias en la propia comprensión. Así, se elige la opción que contiene las palabras más parecidas a las que hay en esa línea del texto sin tener en cuenta si se corresponde con la interpretación de la frase (hay que aclarar que la opción más elegida fue la c)).

Primeras conclusiones

El análisis de la prueba diagnóstica permite extraer algunas conclusiones importantes:

- La mayoría de los reactivos son adecuados para medir la comprensión de microestructuras textuales.
- Algunas preguntas, como la 2, la 3 y la 5b, necesitan ser mejoradas para que cumplan con el objetivo de discriminar entre buenos y malos comprendedores.

Viedma 8 y 9 de noviembre de 2018

- Algunas preguntas, como la 7, deben ser directamente reemplazadas, ya que unen a su alto índice de dificultad una muy pobre capacidad discriminatoria. Además, parece estar evaluando un componente metacognitivo que no forma parte de los objetivos del proyecto.

Al hacer una valoración de los resultados de la pregunta 15 (la que releva comprensión global), se observa que el desempeño general de los estudiantes es mejor de lo esperado si se lo correlaciona con los resultados para las competencias gramaticales. Esto podría indicar que no hay una correlación clara entre ambas variables. Sin embargo, en este punto se pueden tener en cuenta otras posibilidades:

1. que alguno se copió o respondió bien al azar. Esta opción se ve potenciada dado el tamaño reducido de la muestra.
2. que el diseño de la pregunta 15 no es adecuado para dar cuenta en forma fehaciente de la comprensión global de un texto.
3. que la idea de la correlación entre conocimiento gramatical y comprensión general no es acertada.
4. que hay una diferencia en las habilidades que se ponen en juego para la interpretación de las microestructuras, como las oraciones, y las que se utilizan para comprender textos globalmente.

En cualquier caso, es necesario ampliar la muestra y continuar la investigación.

Viedma 8 y 9 de noviembre de 2018

Referencias

Ampudia, Marina. "La educación de jóvenes y adultos en contextos de desigualdad social. Relatos de Abusamra, V., Ferreres, A., Raiter, A., De Beni, R., & Cornoldi, C. (2010). Test. Leer para Comprender TLC. Evaluación de la comprensión de textos.

Comezaña, Gabriela (2018). Informe final del Curso de Posgrado "Estadística descriptiva", dictado por el Mg. Pablo Bohoslavsky, CURZA, Universidad Nacional del Comahue, Argentina.

Nippold, Marilyn (2004). "Research on later language development: International perspectives". En R. A. Berman (Ed.), *Language development across childhood and adolescence. Psycholinguistic and Crosslinguistic Perspectives* (pp. 1-9). Volume 3, Trends In Language Acquisition Research (TILAR). Amsterdam, The Netherlands: John Benjamins

Scott, Cheryl (2004). "Syntactic Contributions to Literacy Learning". En Stone, C. A., Silliman, E. R., Ehren, B. J., & Wallach, G. P. (Eds.). *Handbook of language and literacy: Development and disorders*. Guilford Publications. pp. 340-362.