

Las unidades de procesamiento en la lectura infantil de textos

JORGE ENRIQUE VACA URIBE*

Resumen:

Se expone aquí un estudio acerca de *las unidades de procesamiento de texto* que adoptan algunos niños hispanófonos mexicanos de dos grados escolares diferentes de 8.4 y 9.8 años en promedio, al leer un cuento de 364 palabras. Se grabó la lectura en voz alta de los niños, se transcribió de acuerdo con un sistema específico y se analizan cuantitativamente los *agrupamientos*, evento definido como *el fragmento de lenguaje oral que un observador externo percibe sin segmentación interna alguna y que se corresponde con elementos del texto escrito que se lee*. El análisis de dichos *agrupamientos* nos permite establecer el hecho de que la unidad de procesamiento predominante cambia con respecto a la edad de los niños, pasando de unidades pequeñas, menores o iguales a una palabra, a unidades mayores, que agrupan a dos o más palabras. Conforme aumenta la experiencia de los lectores, ellos parecen estar en mejores condiciones de procesar de manera simultánea más información del texto escrito, lo que a su vez indicaría que los procedimientos de acceso a esas unidades son cualitativamente diferentes. Se discute, además, la importancia de complementar los estudios experimentales de acceso léxico que se realizan al nivel de la palabra, con los estudios de procesamiento a nivel textual.

Abstract:

This article presents the results of a study on the text processing units adopted by twenty Spanish-speaking Mexican children in two different elementary grades (ten second-graders and ten fourth-graders, with an average age of 8.4 and 9.8 years, respectively) when reading a 364-word story. We recorded the children's reading aloud and made a transcription based on a specific system. Then we

* Investigador y director del Instituto de Investigación en Educación de la Universidad Veracruzana, Diego Leño núm. 8, Col. Centro, Xalapa, Veracruz, cc: jvaca@uv.mx

quantitatively analyzed events known as “clusters”, defined as *the oral language segment that an external observer perceives without any internal fragmentation and that corresponds to (a group of) written elements of the text being read*. The analysis of these *clusters* permits us to establish that the predominant processing unit changes according to the child’s age, by passing from small units of one word or less to larger units of two or more words. As a reader’s experience increases, he seems to be more able to process larger fragments of written text simultaneously. This phenomenon indicates that the procedures used to access units are qualitatively different. In addition, we discuss the importance of complementing experimental studies on lexical access at the word level with studies on reading processing at the textual level.

Palabras clave: procesamiento, texto, unidades, niños, evolución.

Key words: text processing, units, children, evolution.

Introducción

Los estudios psicolingüísticos actuales de lectura siguen dos vertientes principales: la primera explora el “procesamiento superficial” del texto y específicamente los mecanismos o procedimientos de reconocimiento de palabras en diferentes poblaciones; muchos son estudios experimentales que trabajan a nivel de “la palabra”, es decir, diseñando interesantes situaciones de reconocimiento de palabras en condiciones experimentales diversas, que controlan variables como el tiempo de exposición, la frecuencia del léxico presentado, sus características ortográficas, la velocidad de las decisiones acerca de si se está frente a una palabra o a una no-palabra, etcétera.

Muchas de estas investigaciones buscan apoyar o refutar (aspectos de) algunos modelos propuestos, uno de los cuales es el elaborado por Uta Frith de 1985 (citado en Sprenger-Charolles y Casalis, 1996), quien tiene gran influencia en este dominio, por ejemplo en la investigación francesa contemporánea de la lectura sobre autores tales como Sprenger-Charolles y Casalis (1996) y Fayol y Jaffré (1999).

La segunda vertiente importante de investigación es aquella que busca entender el procesamiento semántico del texto. Esta vertiente tiene como referente teórico principal el modelo estratégico de comprensión del discurso publicado por van Dijk y Kintch en 1983 (aunque puede verse también van Dijk, 1995). Un ejemplo de la influencia que empieza a tener en los investigadores de habla hispana puede apreciarse en los trabajos reportados por García, Elosúa, Gutiérrez, Luque y Gárate (1999); también puede comenzar a verse su influencia en países de habla francesa (Fayol, 1997).

Desafortunadamente, conocemos pocos estudios que aborden actualmente el procesamiento superficial pero a nivel textual. Es en este contexto que nos situamos.

Consideramos que la psicolingüística de la escritura está aún en sus fases iniciales de desarrollo, por lo que la planificación de “rigurosos” estudios experimentales pueden adolecer de una teoría suficientemente desarrollada que dote de un significado profundo a dichos experimentos. Es así que en la actualidad podemos encontrar cientos, si no miles, de estudios que reportan correlaciones, varianzas y posibles significados de complejísimo análisis estadísticos (muchas veces semi-automatizados) cuya significación en realidad puede ser pobre, por no estar inmersos en una teoría que los signifique. Como Minsky sostuvo con relación al nacimiento de la inteligencia artificial:

El pensamiento siempre comienza con planes e imágenes sugestivas pero imperfectas; éstas son reemplazadas progresivamente por mejores ideas —aunque usualmente aún imperfectas—. [...] Es comprensible que los psicólogos se sientan incómodos con propuestas complejas que no están basadas en mecanismos bien establecidos, pero creo que la parsimonia es aún inapropiada en este estadio, valiosa como lo es en fases posteriores de desarrollo de cada ciencia (Minsky, 1974: 112-113).

Así, estudiamos la evolución del procesamiento de la superficie del texto con el objetivo de complementar los estudios experimentales que actualmente se realizan sobre la palabra. Pretendemos hacer una observación de la evolución de la lectura a cierta distancia, con

cierta perspectiva alejada, que nos permita obtener una imagen de conjunto (necesariamente imperfecta) aunque no podamos observar detalles.

De hecho, hubo una época de la psicolingüística (si podemos hablar de una época en una disciplina tan joven) en la que estos estudios fueron los dominantes, cuando a finales de los años sesenta y principios de los setenta, P. Kolers (1972), K. Goodman (1967 y 1973) y F. Smith (1973) elaboraban ya modelos que daban cuenta, con sus obvias limitaciones, de los procesos de lectura de textos, incluida su comprensión, que comparten principios y perspectivas fundamentales con algunos modelos muy actuales, como se puede apreciar a continuación:

El proceso receptivo comienza con la exposición al *input* fonológico o gráfico y termina con el significado como *output*, pero el usuario eficiente del lenguaje toma la ruta más directa y toca el menor número de puntos intermedios para alcanzar su objetivo. Él logra esto *muestreando*, confiando en la redundancia del lenguaje y en su conocimiento de las construcciones lingüísticas. Él *predice* estructuras, las *prueba* contra el contexto semántico que construye a partir de la situación y del hilo del discurso y luego las *confirma* o las refuta conforme procesa más lenguaje. Los procesos receptivos del lenguaje son ciclos de *muestreo*, *predicción*, *prueba* y *confirmación*. El usuario del lenguaje confía en estrategias que producen la predicción más confiable con el mínimo uso de la información disponible... Ni escuchar ni leer son procesos precisos; y, de hecho, lo que el usuario del lenguaje percibe es en parte lo que ve o escucha y, en parte, lo que espera ver o escuchar (Goodman, 1973:23)

El modelo estratégico de procesamiento del discurso de van Dijk y Kintsch (1983) comparte con Goodman la idea de que el proceso de lectura no está basado en reglas y por lo tanto no es preciso, algorítmico. Más bien se trata, según su perspectiva, de un proceso muy flexible y dinámico, estratégico, determinado en gran medida por las características del sujeto que lee (o que escucha), de su conocimiento del mundo, de sus expectativas, del modelo de situación que él sea capaz de elaborar y a través del cual dará significación y cohesión a los elementos leídos. Sin embargo, este modelo estratégico no aborda el

procesamiento superficial del texto, por un necesario recorte metodológico:

La teoría asume que un *input* verbal es decodificado en una lista de proposiciones atómicas que son organizadas en unidades más grandes con base en estructuras de conocimiento, para formar una base de texto (*text base*) coherente. A partir de esta base de texto, es construida una macroestructura, que representa la información más esencial presente en la base de texto. No solamente el conocimiento del escucha (*comprehender*), sino también sus creencias y objetivos juegan un papel crucial en este proceso. Paralelamente a la construcción de esta base jerárquica de texto, se elabora un modelo de situación (*situation model*), que integra el conocimiento actualmente existente por parte del escucha con la información derivada del texto que está siendo procesado. Entonces, el producto final de la comprensión es un registro de procesamiento multinivel, que incluye trazos en la memoria del *input* lingüístico actual y del significado del texto, ambos a un nivel local y global, y también incluye trazos del efecto que el texto tuvo en el conocimiento del mundo del escucha. El principal interés de este libro es el estudio de las estrategias de procesamiento en línea que producen esta estructura (van Dijk y Kintsch, 1983: x).

En estudios actuales sobre la oralidad, también se toma en cuenta el conocimiento que el escucha tiene de su propia lengua y de los hablantes, así como de sus expectativas:

Todos los estudios muestran que, cuando se trata de nuestra propia lengua, más que escuchar, reconstruimos los enunciados; los reconstruimos en función de las previsiones que acostumbramos hacer sobre nuestra lengua, y según el grado de probabilidad de tal o cual elemento. Por ejemplo, muchos transcriutores “oyen”: “mi cuñado” [mon beau-frère] allí donde el hablante dice “mi pobre hermano” [mon pauvre frère], porque “cuñado” es para ellos más probable. El transcriptor “oye” también en función de los juicios que hace sobre el hablante; si el hablante es una persona poco educada, tenderá a oír muy pocos “ne” de negación [se trata del primer elemento de la negación doble del francés, facultativo en ciertos contextos, especialmente en lengua oral], porque presupone que con esa educación es probable que diga “on sait pas où on va” [uno no sabe adónde va] antes que “on ne sait pas où on va”. Del

mismo modo, el transcriptor corre el riesgo de oír giros sintácticos triviales ahí donde puede oírse un giro muy próximo pero más refinado, si se hace un gran esfuerzo de atención... (Blanche-Benveniste, 1998:24).

Por lo anterior, el estudio sistemático del procesamiento superficial del texto escrito es pertinente y, de hecho, urgente y complementario de los actuales trabajos en psicolingüística de la lectura. Hay una laguna enorme que debe tomarse en cuenta. No se trata de una pretensión personal el decidir abordar el estudio del procesamiento del texto, sino de una necesidad teórica y metodológica, ya que al eliminar el texto de la investigación, suprimimos con él la lectura, proceso que pretendemos estudiar. Porque no podemos reducir la lectura a la identificación de palabras aisladas, por las mismas razones por las que una lengua no es igual a la suma de sus palabras. La investigación acerca de las dificultades o los mecanismos de identificación de palabras aisladas es, sin duda, importante pero representa tan sólo una parte de la problemática del lector. No podemos identificar “el acto léxico” con la lectura:

Todos estamos de acuerdo en que se lee para comprender. Sin embargo, evaluar la lectura únicamente a partir de la comprensión, puede ocultar el hecho de que ciertas dificultades léxicas (término que será utilizado como sinónimo de dificultades de lectura en el resto de la obra) no tienen que ver con esta competencia. Es el caso del niño que comprende perfectamente un texto que se le lee en voz alta pero que no comprende cuando debe leerlo por sí mismo. [...] Para identificar la especificidad de la lectura, es por lo tanto necesario diferenciar lectura y comprensión. La comprensión es la finalidad de la lectura, como la de todas las actividades lingüísticas. Esta finalidad no puede sin embargo ser alcanzada con el lenguaje escrito más que si ciertos mecanismos específicos al acto léxico son bien dominados (Sprenger-Charolles y Casalis, 1996:2-3).

No compartimos esta separación radical entre comprensión y lectura, y mucho menos la reducción que sufre la lectura al acto léxico.

Por añadidura, en este contexto de discusión se inserta otro, también muy complejo, referente a la(s) unidad(es) de procesamiento del texto escrito. Al respecto hay diferentes posiciones, según las aproximaciones

adoptadas. Se defiende a la *letra* como unidad básica de procesamiento, o bien al *grafema*; a la *sílaba*, como unidad fonológica privilegiada y de fácil acceso para el niño o bien unidades intermedias que siguen siendo menores a una palabra (*coda* o *rima*).

También se defiende al *morfema*, por ser la unidad mínima significativa o a la *palabra*, unidad natural de la escritura (en la actualidad). Se defienden unidades superiores a la palabra como la *cláusula*, la *frase* (o sintagma) o a la *oración* misma. Lo que la mayoría de los autores parece compartir es la idea de que el texto escrito, con su linealidad, debe ser organizado, jerarquizado y estructurado por el lector, y que dicha operación de organización supone la identificación y la jerarquización de las unidades del texto.

Percibir el lenguaje significa llevar a cabo varias operaciones psicológicas tales como aislar y segmentar palabras, frases y unidades más largas, además de atribuirles un significado [...] para organizar la señal que nos llega perceptivamente, tenemos que ser capaces de asignarle una estructura lingüística apropiada (Flores d'Arcais, 1988: 123-124).

En este artículo describiremos y analizaremos un grupo de fenómenos aparentemente poco informativos de la lectura en voz alta. Si revisamos con suficiente detalle aquellos fragmentos de la lectura oral de los niños en los que no cometen errores, encontramos indicadores importantes de su manera de abordar el texto y procesarlo. Para establecer un paralelismo con estudios actuales acerca de la oralidad, diremos, con Blanche-Benveniste (1998:28): “Debemos tomar en serio lo que aparece a primera vista como apuros de la producción: son, de hecho, reveladores de funcionamientos esenciales de la lengua”.

Parafraseando a Blanche-Benveniste diríamos que, al estudiar la lectura, debemos tomar en serio lo que a primera vista aparece como apuros en la lectura, ya que ellos pueden ser reveladores importantes de los procesos psicolingüísticos del lector.

Aquellos “apuros” que en los setenta se privilegiaron, fueron los errores de lectura, y fue conocido como el *miscue analysis* de Goodman

(aunque podemos encontrar antecedentes de hace 200 años). Actualmente, también en el contexto del estudio de la oralidad, encontramos estudios como el de Rossi y Peter-Defare (1998), *Les lapsus*, explotado con fines de investigación psicolingüística. En este trabajo nos concentraremos en la descripción y el análisis solamente de algunos “apuros” de lectura que no son errores y que, sin embargo, merecen nuestra atención.

Según el grado de experiencia que el niño tenga como lector, su lectura oral puede variar de una “con tropiezos”, a una muy rítmica y “entonada”, lo que generalmente se llama “fluida”. Buscamos hacer una descripción de la lectura de algunos niños hispanófonos en términos de ciertos fenómenos que hemos privilegiado en la descripción: silabeos, lectura palabra por palabra, y lectura de grupos de palabras. En otras publicaciones (Vaca, 1997a y 2000) hemos llamado *agrupamientos* a los elementos oralizados durante la lectura que se ubican entre pausas, o bien formando contextos entonativos que nos hacen percibirlos como aislados de lo que les precede y de lo que les sucede. Estos elementos pueden ser subléticos (sílabas, pseudosílabas, grupos de fonemas, etcétera), léxicos y supraléticos.

Los datos por analizar pertenecen a lo que llamamos el “corpus mantarraya”, que es un pequeño aunque complejo cuerpo comparativo de datos, recabado y minuciosamente transcrito con el objeto de estudiar algunas relaciones entre las conceptualizaciones ortográficas de los niños y su procesamiento de texto, en niños hispanófonos y francófonos cuyas edades fluctúan entre los 8 y los 10 años y, actualmente, en estudiantes universitarios hispanófonos. En este artículo exponemos solamente los datos relativos a la población infantil hispanófona.

Nuestro *objetivo* es describir y analizar cuantitativamente la proporción de texto que es leído por cada niño mediante agrupamientos subléticos, léxicos y supraléticos en función de su edad (o grupo escolar al que asisten).

Suponemos que si hay una variación significativa, se debe a que la unidad dominante de procesamiento va cambiando respecto de la expe-

riencia que el niño va acumulando como lector y postulamos como *hipótesis* que los niños más pequeños tienden a utilizar unidades menores de procesamiento mientras que los de más edad logran utilizar unidades mayores a la palabra, de manera predominante.

Comprobar la existencia de patrones claros en esos comportamientos lectores y constatar cambios significativos en dichos patrones respecto de la edad, sería muy relevante para replantear el problema de la unidad de procesamiento y también para señalar que la metodología psicogenética podría aportar mucho a nuestra comprensión de la evolución de la lectura.

Problema

En la literatura actual se intenta comprender cuál es la unidad de procesamiento del texto escrito por parte del lector. Evidentemente, este es un problema muy complejo. Nosotros tratamos de abordarlo psicogenéticamente, es decir, de analizar si existen cambios en los patrones de comportamiento lector asociados con el incremento de su experiencia o, lo que es generalmente lo mismo, a la edad o a la escolaridad.

Definimos nuestra variable dependiente en términos de los agrupamientos enunciados por los lectores durante la lectura en voz alta de un texto, como los grupos de elementos lingüísticos (sean fonemas, grupos de fonemas, sílabas, palabras o grupos de palabras) que son percibidos de manera aislada. Nuestra variable independiente corresponde a la edad de los niños (equivalente al grado escolar de los mismos en la muestra que nos ocupa).

Método

El “corpus mantarraya” consta de una serie de entrevistas clínicas realizadas a 20 niños hispanófonos y a 17 niños francófonos, que recaba información diversa sobre la conceptualización del sistema de escritura y sobre el desempeño en lectura. Describiremos, a continua-

ción, solamente aquellos rasgos metodológicos que son pertinentes para este artículo, centrado en la lectura del texto por parte de los niños hispanófonos mexicanos.

Sujetos

Los sujetos del corpus de quienes nos ocupamos tienen las siguientes características:

Edad y grado escolar:

	7-8 años 2° de primaria	10-11 años 4° de primaria
Número de niños(as)	10	10
Edad promedio	8.4	9.8

La diferencia promedio de edad es de 1.4 años entre un grupo y otro; el género no es considerado relevante.

Los criterios de preselección de los niños aseguraban que tuvieran un dominio de la lengua materna como el esperado a su edad, que fueran con un avance regular en la escuela al asistir al grado escolar correspondiente a su edad, que provinieran de grupos escolares diferentes y que pudieran escribir y leer algunas palabras y oraciones acompañadas por imágenes.

El texto leído por los niños es un cuento de 364 palabras que tiene características específicas cuya descripción detallada no es relevante para este artículo, pues su construcción fue pensada para explorar, además, el procesamiento de algunas características ortográficas del sistema de escritura.

Procedimientos de entrevista

Las entrevistas se realizaron a medio camino entre la clínico-crítica propia de las investigaciones psicogenéticas y el método de tests, que

busca fijar las preguntas y la manera de proponerlas a los niños. Diseñamos un orden y una serie de preguntas o intervenciones a proponer a los niños que nos permitió recabar datos sistemáticos en ambas poblaciones, aunque también pudimos indagar con cierta profundidad algunos comportamientos o justificaciones encontradas en los niños. En lo que respecta a la lectura del cuento, el procedimiento era el siguiente:

Consigna: “aquí hay un cuento. Quiero que lo leas en voz alta. Lo más importante es que comprendas lo que vas a leer. Cada vez que llegues a un signo como este (☛) me vas a decir de qué se va tratando, lo que has entendido, ¿de acuerdo? Comencemos”.

Procedimientos y precauciones de transcripción

La tarea de transcribir no es una actividad ingenua. No existe una transcripción que “transvase” lo oral a lo escrito. Sin embargo, para reflexionar sobre lo oral, es necesario escribirlo. Desde la transcripción misma hay un compromiso entre “el dato” y “la interpretación”; asimismo, hay un compromiso entre “su exactitud” y “su legibilidad”. Así, para el caso de la transcripción de la entonación, Martens nos dice:

No existe un sistema único de transcripción, aceptado por todos, para el de la entonación; existe un gran número de ellos, y se puede decir que cada monografía sobre la entonación introduce una nueva notación. Para evitar la trampa de una transcripción que sería el reflejo de un análisis intuitivo de los datos más que de los datos mismos, la notación debe limitarse a los hechos de la observación auditiva, o al menos debe separar claramente el nivel de la observación y el de la interpretación. Una transcripción auditiva que se diga fiel e independiente de quien la realiza, debe igualmente poder ser obtenida por un procedimiento automático, es decir, por una máquina (Martens, 1990:159)

Ninguna transcripción, por fiel y detallada que fuera, permitiría al lector de este trabajo “escuchar con los ojos” las lecturas de los niños, como si fueran las grabaciones. De hecho, esto no sería la solución al

problema del análisis de las lecturas. Si lo fuera, habría otra, moderna, nada despreciable: digitalizar las lecturas, grabarlas en un disco compacto y ponerlas así a disposición. Eso sería equivalente a distribuir copias de las cintas junto con el documento escrito, pero volveríamos al punto de partida: ¿qué analizar?, ¿desde qué óptica?, ¿cómo cuantificar los fenómenos encontrados?, ¿cómo segmentarlos?, ¿cómo clasificarlos?

La transcripción es ya una parte del análisis. En ese sentido, para reflexionar sobre lo oral, es necesario escribirlo.

Concebimos la lectura en voz alta como la oralización de un texto escrito. Esto implica que debemos establecer el texto oral correspondiente, por escrito (sin pretender jugar con las palabras), en el sentido técnico del establecimiento del texto, tal y como lo concibe Blanche-Benveniste (1998).

La lectura en voz alta comparte seguramente características con los textos orales espontáneos, pero también se diferencia en que su objetivo es la reconstrucción del texto escrito que le sirve de base. El lector no debe elaborar (o construir) las ideas propias por expresar ni las formas gramaticales, sino reconstruir (o asignar) un texto oral (posible, de entre varios) al escrito correspondiente, con base en la información proveniente de dicho texto escrito, independientemente de que para relaborar las ideas o estructuras gramaticales producidas por el autor, deba de reconstruirlas.

Lo que debemos de retener en la transcripción y sus niveles de análisis o “exactitud” depende de los objetivos fundamentales de la investigación en su conjunto. Para este trabajo, retendremos los *agrupamientos* de elementos oralizados (y presumiblemente) procesados simultáneamente, por considerarlos un índice valioso de acceso a los mecanismos de procesamiento puestos en juego por el niño. Esto implica poner especial atención a los agrupamientos de elementos oralizados durante la lectura y la elaboración de un sistema simple pero confiable y consistente para su transcripción. En muchas ocasiones es la *pausa* el índice utilizado para detectar el agrupamiento, pero no siempre lo es:

Las bases materiales sobre las que se funda la prosodia, los tonos melódicos, la duración y la intensidad son objetivamente mensurables. Pero se sabe que la percepción auditiva no coincide directamente con esas características objetivas. Ahí donde nosotros creemos escuchar claramente una pausa, sucede frecuentemente que los aparatos de medida muestran que no hay nada. El efecto de pausa de la percepción corresponde a una frontera abstracta que separa dos pedazos de enunciado, que imaginamos hecho de un silencio, mientras que está materialmente compuesto de un conjunto de parámetros que incluye melodías, alargamientos e intensidades (Blanche-Benveniste, 1997:72).

De hecho, intentamos verificar nuestra propia sistematicidad en la transcripción de los agrupamientos mediante la digitalización de fragmentos de lectura y su graficación (con el apoyo del programa CoolEdit). Pudimos medir aproximadamente las pausas, cuando las hay, pero nos encontramos en algunas ocasiones con que, a pesar de no existir una pausa física observable mediante este instrumento, seguimos percibiendo auditivamente una agrupación particular de elementos, producto de un conjunto de parámetros entonativos difícil (o imposible) de analizar (al menos con los medios técnicos con los que contamos). Así, deseamos la idea de “refinar” nuestra transcripción con el apoyo de la digitalización de las lecturas. Nuestro sistema de transcripción de límites entre agrupamientos se restringe a lo siguiente:

- límite percibido entre agrupamientos y/o pausa
- pausa larga
- Ø agrupación de elementos de dos fragmentos diferentes

Resultados

Los *agrupamientos* informan principalmente acerca de la unidad o unidades predominantes de procesamiento que el lector utiliza, y pueden ser definidos como *el fragmento de lenguaje oral que un observador externo percibe sin segmentación interna alguna y que se corresponde con elementos del texto escrito*. Consideremos el siguiente ejemplo, tomando en cuenta que tan sólo deseamos ilustrar nuestro procedimiento de transcripción

y de conteo, y no describir detalladamente nuestro sistema completo de transcripción:

Fragmento de transcripción de lectura

Kri (8.08, 2º)

1 e = 3

{dice} lan r:a-
la raya sal-
la raya salvavidas-

2 r = 2 1 1 r = 2 s = 2 e = 4 2 e = 1

como-to-[carraspeo]
como todos-los-días- un- ray-
un: rayo-de so:l-en-tró en-
en mi cuer-
cuer:-tó--y me-
despe:r-
desper--tó [¿y qué?] -- [agárralo si quieres, para que lo puedas leer bien]
/despértó/ [fuerte]
des-
/despértó/-

3 1 1 1 1 1 r = 2 s = 1 1 1 s = 1

como-todos-los-días-leí- un-
un rato-mien-tras-mi-papá-des-per-taba-

4 1 1 1 1 1 e = 1 s = 3 r = 1 1 1

abrí-un-libro-y-me- /é-ko/-
encontré-con la fo-tografía- de-
de-una-raya-

Hemos analizado así cada fragmento del texto leído por cada niño y construimos, con fines de análisis, una cadena que recupera lo esencial de los eventos encontrados durante el procesamiento. En el fragmento 2 del ejemplo, tenemos la siguiente cadena:

2 r = 2 1 1 r = 2 s = 2 e = 4 2 e = 1

en donde,

r = 2 codifica la repetición de dos palabras (como todos). Cada repetición es codificada en el renglón siguiente, para representar el trabajo que posiblemente se realiza sobre el eje paradigmático, siguiendo los sistemas de transcripción del GARS en Francia (Groupe Aixois de Recherches en Syntaxe), véanse los textos citados de Blanche-Benveniste.

1 1 codifican las palabras leídas en forma aislada (-los-días-)

s = 2 codifica la lectura de dos palabras (-de so:1-) mediante un agrupamiento subléxico en el que se agrupa una palabra (de) con un grupo de fonemas (so:); a este tipo de agrupamiento subléxico lo hemos denominado “mixto”, por agrupar una o más palabras con un elemento menor que una palabra; en este caso es limitado por el alargamiento vocálico, transcrito por : (que fue considerado sistemáticamente como equivalente a una pausa, ya que los niños tienden a producir dichos alargamientos para darse un cierto tiempo adicional de exploración del texto). Además, queda otro agrupamiento subléxico, correspondiente a un solo fonema consonántico aislado (:1-).

e = 4 codifica la aparición de un error de lectura (subrayado en la transcripción) sobre la palabra cuarto, que “arrastra” a otras tres palabras en un proceso en el que, a su vez, hay repeticiones y correcciones parciales (entró en mi cuarto).

2 representa un agrupamiento biléxico, es decir, de dos y sólo dos palabras (--y me-).

La suma de los numerales de la cadena equivale a 15, número de palabras del fragmento dos del texto (como todos los días, un rayo de sol entró en mi cuarto y me despertó).

El fragmento 3 es leído por Kri con una predominancia clara de agrupamientos uniléxicos (1 1 1...), una repetición y dos palabras leídas mediante agrupamientos subléxicos, unisilábicos la primera (-mientras-) y uni- y polisilábicos la segunda (des-per-taba-).

Considérese, en contraste, las características del procesamiento de Kar (se indica entre paréntesis la edad en sistema decimal y el grado escolar):

Fragmento de transcripción de lectura

Kar (9.25, 4°)

1 2 s = 1

la raya-salva:-vidas

2 <4> 4 4 3

<como todos los días>-un rayo de sol-entró en mi cuarto-y me despertó-

3 <4> 3 4

<como todos los días>-leí un rato-mientras mi papá despertaba-

4 3 3 6

abrí un libro-y me encontré-con la fotografía de una raya-

La diferencia entre los procesamientos de Kri y Kar son evidentes: con muchos “apuros” el primero y asumiendo unidades pequeñas de procesamiento y una lectura muy fluida la segunda, sin repeticiones, con pocos errores y en unidades mayores (agrupamientos poliléxicos, generalmente de tres o más palabras).

Aplicando sistemáticamente este procedimiento de codificación, hemos cuantificado qué proporción del texto es oralizado mediante agrupamientos subléxicos, uniléxicos, biléxicos y poliléxicos, de tal forma que obtuvimos los porcentajes de acuerdo con el total de palabras del texto (364 en el caso del español). El cuadro 1 expresa los datos de manera concentrada, del que queremos sólo subrayar los siguientes aspectos:

- 1) Hay niños de 2° grado que muestran una proporción importante de procesamiento subléxico (Edu, Kri y Arl), de alrededor de 25% sobre la totalidad del texto. Tomando en cuenta que se trata de niños de poco más de ocho años, la proporción es elevada.
- 2) Estos mismo niños, junto con Ale, tienen también una tendencia muy marcada al procesamiento uniléxico, quienes presentan una proporción cercana a un tercio (entre 29% y 36%) de la totalidad del texto.

Estos niños con predominancia de agrupamientos subléxicos y uniléxicos son generalmente los pequeños, aunque hay algunos que no se comportan así. Sabemos que la edad o el grado escolar es tan sólo una variable aproximada de los procesos y ritmos evolutivos.

- 3) Como se puede observar en el cuadro 1, las proporciones predominantes van desplazándose hacia los agrupamientos bi- y poliléxicos conforme observamos el comportamiento de los niños mayores, lo que refleja una lectura cada vez más fluida de su parte.

Cuadro 1
Porcentaje de agrupamientos por niño y grupo escolar

Nombre	Edad	Grado	Sub- léxicos	Uni- léxicos	Bi- léxicos	Poli- léxicos	Total
Edu	8.5	2	24.2	36	13	1.6	74.8
Kri	8.08	2	25.3	35	6	0.8	67.1
Ale	8	2	9.3	35	12	8.4	64.7
Arl	8.58	2	22.8	29	16	4.9	72.7
Adr	8.5	2	6.9	21	27	28	82.9
Mih	8.4	2	3.3	19	25	36.7	84
Mar	8.25	2	0.3	9.9	20	51.7	81.9
Irv	8.33	2	1.1	6.9	8	58.4	74.4
Lau	8.58	2	3	4.7	9	60.1	76.8
JuJ	8.17	2	1.9	3	7	54.4	66.3
Promedio	8.33		9.8	20	14.3	30.5	74.6
Ana	9.64	4	1.4	5.2	10	28	44.6
Pas	10.1	4	2.7	8.2	18	29.1	58
MaF	10.2	4	3.6	6.9	10	49.2	69.7
Arc	10.1	4	0	1.4	4	53.7	59.1
Car	9.5	4	5.8	5.2	8	55	74
MaL	9.83	4	2.7	5.8	14	58	80.5
AnM	9.58	4	0	3.6	12	55.2	70.8
Ing	9.5	4	0.3	4.7	10	58.9	73.9
Kar	9.25	4	3.8	4.1	11	64.6	83.5
Wen	10.2	4	3.8	3.6	11	66	84.4
Promedio	9.78		2.4	4.9	10.9	51.8	70

Los datos indican claramente que la unidad de procesamiento va cambiando. Si tomamos en cuenta, además, nuestros datos del estudio longitudinal (Vaca, 1997a), diríamos que la unidad de procesamiento predominante es inicialmente la sílaba y luego la palabra. Considerando ahora la distribución porcentual grupal, obtenemos lo siguiente:

Comparación grupal porcentual de los agrupamientos respecto de la edad

Grado esc.	Agrupamientos				
	<i>Subléxicos</i>	<i>Uniléxicos</i>	<i>Biléxicos</i>	<i>Poliléxicos</i>	<i>Subtotal</i>
2°	9.8	20.0	14.3	30.5	74.6
4°	2.4	4.9	10.9	51.8	70.0

Podemos observar que del 2° al 4° grado, es decir, con 1.4 años de diferencia de edad en promedio:

- 1) La proporción de agrupamientos subléxicos disminuye, así como la de los uniléxicos.
- 2) La proporción de agrupamientos biléxicos es similar (entre 11 y 14% del texto), aunque es necesario aclarar que su significado es diferente según las poblaciones o el proceso evolutivo que podemos formular hipotéticamente: mientras que para los niños pequeños representa un logro y un “límite superior” respecto del predominio de agrupamientos menores en su lectura (en ese momento de su proceso de evolución), para los niños mayores representa un tipo de respuesta que tiende a ser abandonada, por acceder ya a la posibilidad de agrupamientos mayores durante la lectura. El agrupamiento biléxico correspondería a una “respuesta intermedia” de procesamiento.
- 3) Los agrupamientos poliléxicos llegan escasamente a un tercio en la muestra de los niños de segundo grado, mientras que representan la mitad de los agrupamientos realizados por los niños mayores. La diferencia de proporción (20%) es considerable.

Hemos realizado un análisis complementario de las características superficiales de la lectura con el auxilio de un programa computacional de búsqueda recursiva de patrones que se repiten en una cadena de símbolos, denominado Grammar (Ebeling y Jiménez-Montaño, 1980)¹ que comprime cadenas de caracteres que describen ciertos fenómenos (originalmente creado para estudiar problemas de secuencias de proteínas en el código genético).

Su funcionamiento, grosso modo, consiste en analizar grandes secuencias de caracteres de una forma recursiva: se elige el primer par de caracteres, y ese par es buscado en toda la cadena. Si éste se repite al menos tres veces, es sustituido por un símbolo no terminal (S_0 , por ejemplo). El programa registra la frecuencia de ese par de caracteres y sustituye, en la cadena original, los pares por el símbolo no terminal asignado.

Se reitera el proceso tantas veces como sea necesario, hasta que no se encuentren ya pares que se repitan tres veces o más. Debe notarse que, como los pares de caracteres son sustituidos por símbolos no terminales, en el ciclo siguiente son considerados en el análisis, lo que permite extraer no solamente parejas, sino tercias, cuartetas, etcétera. de elementos que se repiten en determinada secuencia —patrones—. Al final, Grammar da un reporte, ordenando los patrones del más al menos frecuente, así como otros datos complementarios como la longitud total de la cadena y su complejidad (un índice asociado con la reductibilidad de la cadena: mientras menos se reduzca —o comprima— la cadena original, más compleja es, ya que no se localizaron patrones regulares).

Puesto que nosotros, como se ha expuesto, habíamos codificado con secuencias de caracteres los eventos de lectura encontrados en cada fragmento, de cada niño y de cada grupo, decidimos analizar con Grammar estos datos. Una parte mínima del análisis refuerza los datos que ya hemos presentado.

El cuadro 2 muestra los cinco patrones más frecuentes encontrados al analizar nuestras cadenas. Se procedió a hacer una cadena de caracteres con todos los niños de cada grupo, obteniendo los siguientes resultados:

Cuadro 2
Los cinco patrones más frecuentes extraídos
por Grammar de las cadenas

Población	Esc.	Patrones más frecuentes					% reducción
		1°	2°	3°	4°	5°	
Hispanófono	2	11	1s	12	111	rr	42
	4	21	23	rr	22	43	32

Del cuadro 2 podemos destacar lo siguiente:

- 1) Mientras que en el grupo de niños menores, cuatro de los cinco patrones más frecuentes tienen al menos un agrupamiento uniléxico, en los mayores sólo un patrón lo tiene, siendo los demás sustituidos por agrupamientos bi- y/o poliléxicos (respecto de los patrones encontrados en el grupo de menores). Este análisis reafirma el anterior: conforme aumenta la experiencia en lectura, el niño tiende y es capaz de procesar unidades mayores de texto de manera simultánea, cambiando la unidad predominante de procesamiento.
- 2) Los porcentajes de reducción de la cadena son importantes en todos los grupos: el menor es de 32% (lo que significa que Grammar pudo reducir la cadena en una tercera parte); lo anterior es muy significativo, ya que indica que se encontraron patrones regulares que se repiten al menos tres veces y que fueron sustituidos por símbolos no terminales. Además, se puede apreciar que el porcentaje de reducción es mayor para la población de niños pequeños que para los mayores, lo que puede significar que los más pequeños aplican sus procedimientos muy sistemáticamente, que giran evidentemente en torno a la palabra como unidad de procesamiento.

Debemos suponer que hay cambios cualitativos en los procedimientos utilizados por los niños al leer que les permiten procesar el texto escrito en unidades cada vez más grandes. Explicar qué es lo que hace

posible esos cambios, o cómo se logran, es una tarea difícil. Por el momento, lo que nos interesa subrayar es, simplemente, que la evolución que parecen seguir los niños implica cambios en la unidad de procesamiento y, seguramente, en los procedimientos mismos de acceso al léxico, que posibilitan un camino más rápido a unidades más grandes, quizás a costa de la precisión en el mismo.

Conclusiones

En este artículo analizamos solamente lo que hemos denominado *agrupamientos*, y sólo cuantitativamente en un reducido rango de edad: de los 8 a los 9 años. Si extendiéramos este análisis de manera sistemática hacia niños menores (5-6 años) y hacia mayores (12 años), obtendríamos un panorama más preciso de los cambios que sufren los agrupamientos en la evolución de la lectura.

Anteriormente (Vaca, 1997a) realizamos algunos análisis cualitativos de los agrupamientos, construyendo una clasificación mucho más refinada y también un análisis más detallado de cuáles palabras forman grupos, en función de su pertenencia a un mismo constituyente gramatical o a diferentes constituyentes (entre otras características analizadas). Sin embargo, en este artículo hemos querido restringir el análisis al aspecto cuantitativo, para establecer claramente el hecho de que, efectivamente, aquí hay un fenómeno regular que merece ser estudiado en detalle.

Los resultados expuestos, en coordinación con los anteriormente publicados, nos permiten establecer el hecho de que *hay una regularidad en la evolución de los agrupamientos mediante los cuales los niños procesan el texto escrito*. Hay una diferencia significativa que está de acuerdo con la hipótesis establecida: los niños pequeños tienden a procesar el texto mediante unidades pequeñas, subléxicas o uniléxicas, mientras que los niños mayores lo pueden hacer ya mediante unidades mayores, predominantemente bi- y poliléxicas.

Es importante mencionar que los resultados que estamos presentando aquí referentes a los niños hispanófonos, son similares (aunque no

idénticos) a los obtenidos por la muestra francófona del “corpus mantarraya”. Este hecho, aparentemente simple, debe estar estrechamente relacionado con los procedimientos de procesamiento que el niño va usando y éstos, a su vez, pueden estar relacionados con la evolución de las conceptualizaciones que el niño va logrando sobre el sistema de escritura (Vaca, 2001).

Por su aparente trivialidad, los hechos que aquí intentamos establecer o bien no han sido considerados en las descripciones y estudios sobre la lectura de textos, o bien han sido analizados por diferentes autores y en distintas épocas bajo perspectivas diversas. Smith (1973) y Goodman (1973) interpretaron los agrupamientos subléxicos, uniléxicos y las repeticiones (bajo otras denominaciones) como productos de la influencia escolar. Ferreiro y cols. (1982), adoptando los argumentos de aquellos autores, ha englobado esos comportamientos bajo el rubro de “descifrado”, que es también para ella “un producto típicamente escolar” que no pertenece a la psicogénesis “natural”. Estas interpretaciones son discutibles (Vaca, 1997b y 2000).

Estos resultados deben también ser vinculados con algunas de las tendencias actuales de investigación, ya que los agrupamientos subléxicos que observamos pueden estar estrechamente relacionados con los procedimientos de acceso al léxico por la “vía indirecta de ensamblaje fonológico”; su drástica disminución podría estar ligada a la emergencia de la “vía directa lexical” o “vía ortográfica” (según los autores), que en la actualidad se discute, según el mismo modelo y sus zonas de influencia (Rieben, Fayol, Perfetti, 1997).

Obviamente, no se pueden hacer corresponder mecánicamente los resultados de las investigaciones realizadas sobre la base de unidades pequeñas (palabras) con los provenientes de investigaciones de lectura de textos. Sin embargo, podemos comenzar a explorar las similitudes y las diferencias entre los modelos (o teorías) generados por aproximaciones más parsimoniosas —experimentales y con una búsqueda de control estricto de variables— con otras más exploratorias, pero que se arriesgan a considerar otro tipo de variables como la influencia de la sintaxis y la semántica en el procesamiento mismo, la intervención de aspectos pragmáticos, etcétera. De ninguna manera podemos

considerar excluyentes a estas aproximaciones metodológicas, mucho menos en una disciplina tan joven —escasamente tiene 40 o 50 años— como la psicolingüística. Decir que una y solamente una de las aproximaciones es válida, sería puro dogmatismo y cerrazón, actitudes pre- y anticientíficas por excelencia.

Queremos también marcar el punto de contacto de nuestros datos con los estudios que actualmente realiza el GARS en Francia, a través del establecimiento y análisis de un gran corpus de datos orales recabado y transcrito recientemente sobre el francés actual. Bajo esta perspectiva, hay un retorno al análisis de la especificidad de la oralidad, lo que permite a su vez definir mejor la especificidad de la escritura y sus usos. En estos trabajos, como hemos dejado entrever, hay una intención a la vez descriptiva e interpretativa, que da una nueva dimensión al estudio de la oralidad o de aspectos específicos de ella (la sintaxis, por ejemplo).

De la misma manera, consideramos que el estudio de lectura requiere de una aproximación que renueve y, a la vez, retome el análisis de los procesos psicolingüísticos de lectura de textos y su evolución. Dicha aproximación se interrumpió durante los años ochenta y noventa, porque los psicolingüistas adoptaron concentraciones diferentes (el estudio de la comprensión del discurso es un claro ejemplo), sin que hayan quedado cerrados los capítulos abiertos en las primeras dos décadas de la psicolingüística.

Nos parece importante destacar, como una contribución complementaria de este artículo, el haber rescatado —sobre la descripción de la lectura— lo que sucede cuando parece no haber nada “anómalo”. Lo que hemos llamado *agrupamientos* no ha sido, hasta donde sabemos, tematizado por los investigadores que conocemos. Generalmente se describe la lectura en términos de velocidad, de número de errores cometidos, etcétera.

Los eventos de lectura que hemos destacado, los *agrupamientos*, junto con los errores y las repeticiones son elementos importantes en la descripción de la lectura de una persona y pueden ocultar valiosas informaciones para que construyamos una imagen cada vez más clara del proceso

psicolingüístico de la lectura, así como de la evolución de este proceso. Estamos dejando ver una pequeña parte de nuestro intento por sistematizar una descripción que sirva de base a la inferencia de los procesos psicolingüísticos subyacentes a la lectura. En términos más generales, queremos destacar la necesidad de construir sistemas descriptivos consistentes y que tiendan a ser comunes. Éstos serían la base de los intercambios comunicativos entre los investigadores en esta disciplina.

Nota

- ¹ Agradecemos al doctor Manuel Martínez Morales y a Julio Sandria R., de la maestría en Inteligencia Artificial de la Universidad Veracruzana que nos apoyaron con la realización de este análisis, en el contexto del Proyecto Conacyt (ref. 25977-A) “Modelaje y caracterización de patrones en las ciencias biológicas y del comportamiento por métodos computacionales”.

Referencias bibliográficas

Blanche-Benveniste, C. (1990). *Le français parlé: Études grammaticales*, París: CNRS.

Blanche-Benveniste, C. (1997). *Approches de la langue parlée en français*, París: Ophrys.

Blanche-Benveniste, C. (1998). “Algunas características de la oralidad”, en C. Blanche-Benveniste (ed.) *Estudios lingüísticos sobre la relación entre oralidad y escritura* (L. Varela, trans.), Barcelona: Gedisa, pp. 19-28.

Ebeling, W. y Jiménez-Montaña, M. A. (1980). “On grammars, Complexity, and Information Measures of Biological Macromolecules”, en *Mathematical Biosciences* 52, pp. 53-71.

Fayol, M. (1997). *Des idées au texte*, París: PUF.

Fayol, M. y Jaffré, J-P. (1999). “L’acquisition/apprentissage de l’orthographe”, en *Revue française de pédagogie*, núm. 126, enero-marzo, pp. 143-170.

Ferreiro, E.; Gómez-Palacio, M. y cols (1982). *Análisis de las perturbaciones en el proceso de aprendizaje de la lecto-escritura*, fascículo 3, México: SEP-OEA.

Flores d'Arcadis, G. (1988). "La percepción del lenguaje" en Newmeyer, F. J. (comp.) (1988) *Panorama de la Lingüística Moderna*, vol. III, Madrid: Visor.

García, J. A.; Elosúa, M. R.; Gutiérrez, F.; Luque, J. J.; Gárate, M. (1999). *Comprensión lectora y memoria operativa: aspectos operativos e instruccionales*, Barcelona: Paidós.

Goodman, K. (1967). "Reading: A Psycholinguistic Guessing Game" en Singer, H., Ruddell, B. (eds.) (1976), *Theoretical models and processes of reading*, Newark: IRA.

Goodman, K. (1973). "Analysis of Oral Reading Miscues: Applied Psycholinguistics" en Smith, F. (1973) *Psycholinguistics and Reading*, Toronto: Holt, Rinehart and Winston.

Minsky, M. (1974). "A Framework for Representing Knowledge", en Haugeland, J. (ed.) (1997) *Mind Design II*, Cambridge: MIT Press.

Martens, P. (1990). "Intonation", en C. Blanche-Benveniste. *Le français parlé: Études grammaticales*, París: CNRS.

Kolers, P. (1972), "Experiments in reading", en *Scientific American*, julio, vol. 227, 1, pp. 84-92.

Rieben, L.; Fayol, M.; Perfetti, Ch. (1997). *Des orthographes et leur acquisition*, Lausanne: Delachaux et Niestlé.

Rossi, Mario y Peter-Defare, E. (1998). *Les lapsus*, París: PUF.

Smith, F. (1973) "Decoding: The great Fallacy", en Smith, F. (1973) *Psycholinguistics and Reading*, Toronto: Holt, Rinehart and Winston.

Sprenger-Charolles, L. y Casalis, S. (1996). *Lire*, París: PUF.

Van Dijk, T. (1995). "On Macrostructures, Mental Models, and Other Inventions: A Brief Personal History of the Kintsch-van Dijk Theory", en Weaver, Ch. A.; Mannes, S. y Fletcher, Ch. R, *Discourse comprehension: Essays in honor of Walter Kintsch*, Hillsdale: LEA.

Van Dijk, T. y Kintsch, W. (1983), *Strategies of discourse comprehension*, Nueva York: Academic Press.

Vaca, J. (1997a). *El niño y la escritura*, Xalapa: UV.

Vaca, J. (1997b). “El descifrado en la lectura inicial convencional: ¿momento psicogenético o producto escolar?” en Waldegg, G. y Block, D. (1997) (eds.), *Estudios en Didáctica*, México: Grupo Editorial Iberoamérica.

Vaca, J. (2000). “Jérôme no lee como debe... según las teorías contemporáneas de la lectura” en *Lectura y Vida*, año 21, núm. 4, diciembre, pp. 16-30.

Vaca, J. (2001). *Morphogrammes, Logogrammes et lecture enfantine: une recherche comparative espagnol-français*, tesis doctoral, universidad de Lumière, Lyon 2.

Recepción del artículo: 14 de enero de 2002

Aprobado: 29 de octubre de 2002