

# Colección Pedagógica Universitaria

No. 36

julio-diciembre 2001

Corpus Mantarraya:  
Análisis de desaciertos de lectura

Alfonso Javier Bustamante Santos

Departamento de Investigación Educativa del CINVESTAV-IPN

## Abstract

En este trabajo se analizan desaciertos de lectura que pertenecen al *corpus mantarraya*, cuerpo de datos recabado con el objeto de estudiar algunas relaciones entre conocimiento ortográfico y procesamiento de texto en niños y adultos.

El trabajo se estructura alrededor de tres aspectos principalmente, se aplica la clasificación propuesta por Vaca (1997) y se compara con los resultados obtenidos en sus trabajos, se hace un análisis más específico de las omisiones de palabras, ya que presentan dificultades particulares de clasificación y por último se demuestra pormenorizadamente que el tamaño de la palabra es una información básica que los lectores utilizan para realizar las inferencias acerca de qué palabras están escritas en el texto.

Algunas conclusiones a las que se llegó a través del análisis de los datos son que los desaciertos en la lectura son producto de un procesamiento activo del lector, ya que todos los niños de la muestra presentan desaciertos, lo cual confirma que son inherentes al proceso de lectura.

## Corpus Mantarraya: Análisis de desaciertos...

Se encontraron resultados similares a los reportados por Vaca (1997). El análisis de las omisiones permite reportar cómo su aparición puede ser favorecida por ciertas estructuras sintácticas y generalmente son corregidas cuando afectan la sintaxis de lo leído. Por otra parte, el análisis de los datos del *corpus mantarraya*, nos permitió verificar que el tamaño de la palabra también es una información básica, que el lector usa para realizar sus inferencias.

\* \* \*

### Introducción

Existen diversos elementos por medio de los cuales es posible evaluar la lectura de los niños, por ejemplo: la entonación, la fluidez, la velocidad y otros que nos parecen indeseables y que aparentemente expresan un problema de aprendizaje, como es el caso de los errores o desaciertos,<sup>1</sup> a los cuales se les podría considerar como resultado de dificultades viso-perceptuales o del habla; no obstante, pocas veces los consideramos como parámetros que nos permiten observar la evolución en el dominio de la lectura.

En este trabajo, realizamos un análisis de desaciertos de lectura<sup>2</sup> en voz alta hechos por niños de segundo y cuarto grado de primaria, a partir del cual se muestra cómo estas variaciones entre lo escrito y lo dicho expresan las estrategias utilizadas por los niños y el conocimiento que tienen del lenguaje para acceder al significado del texto, y que las características de estos desaciertos evolucionan paralelamente al dominio de la lectura.

Tradicionalmente, ha habido dos posturas acerca de los errores que cometen los niños al leer en voz alta: o bien se considera que la buena lectura es la lectura sin errores, entonada y fluida, o bien se considera que el error es inherente al proceso de la lectura, y se le considera entonces como una vía de acceso al proceso psicolingüístico subyacente a la misma. Es esta segunda postura la que asumiremos y justificaremos.

Los desaciertos de lectura que analizaremos pertenecen al *corpus mantarraya*, un complejo cuerpo de datos recabado con el objeto de estudiar algunas relaciones entre conocimiento ortográfico y procesamiento de texto en niños y adultos.

El desarrollo del presente trabajo contempla tres aspectos centrales:

Comparación de los resultados obtenidos con los encontrados por Vaca (1997), se hace un análisis específico de los desaciertos que tienen que ver con omisiones de palabras debido a que presentan dificultades particulares de clasificación y por último se demuestra pormenorizadamente que el tamaño de la palabra es información básica que los lectores utilizan para realizar las inferencias acerca de qué palabras están escritas en el texto.

### Concepciones de la lectura

La lectura, considerada como un proceso donde el lector tiene un papel activo, nos permite centrarnos en observar cómo el lector genera estrategias para articular la información obtenida que le permita construir los significados. Goodman (1967) la describió como:

La lectura es la fase receptiva de la comunicación escrita. En el lenguaje escrito el escritor ha codificado un mensaje en símbolos gráficos distribuidos espacialmente en la página. El lector no se limita a pasar simplemente los ojos sobre el lenguaje escrito, recibiendo y registrando un flujo de imágenes perceptivo-visuales. Debe aportar activamente su conocimiento del lenguaje, su experiencia previa, sus logros conceptuales, al procesamiento de la información lingüística codificada en la forma de símbolos gráficos, para poder codificar el lenguaje escrito. Por lo tanto, la lectura debe ser considerada como una interacción entre el lector y el lenguaje escrito, a través de la cual el lector intenta reconstruir el mensaje del escritor.

La postura de Smith (1986) concuerda con la de Goodman al mencionar que:

La lectura no es una actividad pasiva –los lectores deben realizar una contribución activa y sustancial si pretenden darle sentido a lo impreso.

... la lectura no solamente es una actividad visual, tampoco una cuestión de decodificar el sonido, el lector debe utilizar la información ortográfica, sintáctica y semántica para reducir las probabilidades de cometer errores.

## Corpus Mantarraya: Análisis de desaciertos...

Carrasco (1999) agrega aspectos socio-constructivistas en su noción de lectura al definirla como:

un proceso de construcción de significados determinados culturalmente, durante el cual el lector, con los referentes y esquemas socialmente adquiridos, aplica estrategias diversas para construir una comprensión de un mensaje, comunicado a través de un texto escrito.

Por otra parte, Vaca (2000) afirma que el texto leído nunca es igual al texto impreso, ya que el texto, como objeto de conocimiento, se ve modificado por el conocimiento previo y por las estrategias empleadas de quien accede a él.

Retomando las concepciones anteriores, podemos decir que para leer como lo hace un lector experto es necesario, a partir de nuestra experiencia, seleccionar de toda la información disponible aquella que nos permita anticipar con rapidez y cierta precisión lo que está escrito con ello posiblemente podríamos asegurar una lectura fluida y con pocas probabilidades de cometer desaciertos. No obstante, existen diversos elementos en el proceso de lectura que pueden provocar diferencias entre el texto escrito y lo leído. Goodman (1967) menciona que:

El proceso de decodificación directa del input gráfico se ha vuelto tan habitual que el lector debe decodificar primero y luego codificar el significado como output oral. Como cabe esperar, esto resultará en cambios considerables del input gráfico original al output oral.

Por otro lado, puede suceder que las unidades del texto hayan sido decodificadas y procesado su significado correctamente, pero la codificación orientada a la producción no, dando lugar a la oralización de algo diferente.

Si profundizamos un poco más al respecto, encontramos en el mismo texto de Goodman que:

La lectura es una serie rápida de anticipaciones, de procesamientos tentativos de la información. Cuanta menos información disponible use el lector, más rápida y eficiente será su lectura.

Las anticipaciones son realizadas con base en la información que posee el lector, su conocimiento del lenguaje, su experiencia en la lectura, su contexto social y cultural, el tipo de texto, etc.

Carrasco (1999) menciona que:

Un buen lector es aquel que utiliza información proveniente de dos fuentes para construir comprensión al leer: del texto y de su conocimiento previo. (...) Los mejores lectores, son los que logran aprovechar las dos fuentes de información: el texto y sus saberes equilibrando las aportaciones de ambas para construir inferencias adecuadas.

El concepto de anticipación manejado por Goodman y el de inferencias mencionado por Carrasco, también es encontrado en el modelo propuesto por Vaca (1997) al cual le llama proceso inferencial.

En su modelo Vaca destaca un proceso inferencial para poder reconstruir el texto, según el cual el lector accede a él realizando inferencias acerca de lo que está escrito. Muchas de estas inferencias son acertadas, lo cual permite una lectura fluida. Sin embargo, hay inferencias que no son acertadas, dando como resultado "errores" o desaciertos:

El estudio de los desaciertos nos permite observar los elementos que intervienen en la generación de las inferencias, tanto para los desaciertos como para los aciertos.

Se trabaja con el postulado de que al conocer la naturaleza de los errores podemos comprender los mecanismos que llevan a cabo para que éstos aparezcan y de esa manera nos puede dar una orientación de las actividades que se realiza cognitivamente cuando no se comenten errores (Vaca 1997).

En el análisis de los errores hay diferentes metodologías que se han modificado a través del tiempo; Carrasco (1999) dice:

...en el pasado se asignaba una mayor importancia a los resultados y que en la actualidad nos interesamos también por los procesos, particularmente por los procesos de pensamiento del lector y por la forma en la que él logra transar, a partir de lo que sabe, con el texto para formular y contestar preguntas, construir explicaciones, establecer relaciones e inferir como estrategias que el lector pone en juego para construir significados de lo leído.

Una metodología clásica en el análisis de los errores es el "mische analysis" de Goodman; en éste se analizan los errores con base en la categoría sintáctica a la que pertenece la palabra efectivamente escrita y la que es leída en su

## Corpus Mantarraya: Análisis de desaciertos...

lugar; es decir, se analiza la cantidad de sustantivos, verbos, preposiciones, adjetivos, etc., y sus respectivas palabras sustitutas.

En este tipo de estudios se encontró que la mayoría de los errores cometidos son sustituciones de una palabra por otra y que además esas sustituciones corresponden a un mismo tipo de categoría sintáctica; o sea que el lector sustituye, por lo general, un verbo por otro o un sustantivo por otro sustantivo, etc. Un ejemplo de estos estudios es el de Kollers (1985), en el que menciona:

When the student misread a noun, they tended to substituted a noun, a misread verb was supplanted by a verb, and so on.

Sobre las causas de estas sustituciones, Goodman menciona:

Los errores ocurrirán cuando ciertos elementos léxicos, ciertas estructuras sintácticas, conceptos o eventos introducidos en la historia son inesperados, poco familiares o de algún otro modo son difíciles de predecir para el lector (Goodman, 1967)

Otra metodología para el análisis de errores mencionada por Carrasco (1999) es por medio de la inferencia, que también se encuentra en el modelo analítico de McCormic; este modelo analiza los tipos de errores al responder preguntas escritas.

Para fines de este trabajo, nos basamos en el modelo de análisis de errores propuesto por Vaca (1997), que de manera resumida expresa lo siguiente:

1. El niño intentará acceder a la “unidad”, ya sea sílaba, palabra, etc., dependiendo del momento de su desarrollo en la adquisición de la lecto-escritura, por medio del proceso llamado inferencia.
2. Por inferencia se entiende la reconstrucción de un elemento lingüístico (“unidad”) a partir de algunas partes conocidas de esa unidad.
3. Por partes conocidas se entiende los diferentes tipos de información que el niño puede utilizar para realizar la inferencia: el tamaño de la palabra, sus letras iniciales, su posición en el sintagma, lo leído anteriormente (el contexto semántico).

4. El acceso rápido y acertado a la unidad depende de las clases de información que el niño utilice, así como de la equilibración entre ellas, es decir, que lector no se quede centrado en sólo una clase de información.

Vaca (1997) presenta un sistema de clasificación que diferencia las clases de desaciertos que se comenten al momento de leer:

Clase 1. Los desaciertos léxicos son los que consisten en oralizar una palabra en vez de otra, en omitir la oralización de una palabra o en introducir una palabra real del lenguaje.

Clase 2. Los desaciertos no léxicos, consistentes en oralizar palabras no existentes (pseudopalabras) o fragmentos de palabras que no pueden llegar a formar palabras reales del lenguaje.

Los desaciertos léxicos los subclasifica en nueve tipos. Puesto que toman al menos tres tipos de información para realizar la inferencia: La información *fonográfica* (los elementos gráficos de la palabra y su correspondencia fonológica), la información *sintáctica*, correspondiente a la consideración de aspectos sintácticos, y la información *semántica*, es decir, la adecuación semántica del elemento inferido respecto del significado del texto que se lee. Así, en función de la consideración de esas tres clases de información, se constituyen ocho tipos de los nueve.

Consideración de la información (ver tabla 1):

**TABLA 1. CLASIFICACIÓN DE ACUERDO A LA INFORMACIÓN EMPLEADA COMO BASE PARA LA INFERENCIA.**

| Tipo de desacuerdo | Fonográfica | Sintáctica | Semántica |
|--------------------|-------------|------------|-----------|
| 1                  | Sí          | Sí         | Sí        |
| 2                  | Sí          | Sí         | No        |
| 3                  | Sí          | No         | Sí        |
| 4                  | No          | Sí         | Sí        |
| 5                  | Sí          | No         | No        |
| 6                  | No          | Sí         | No        |
| 7                  | No          | No         | Sí        |
| 8                  | No          | No         | No        |

## Corpus Mantarraya: Análisis de desaciertos...

Los siguientes ejemplos son parte de los datos aquí analizados. En cada desacuerdo se especifica el tipo de desacuerdo, la descripción de la información utilizada para realizar la inferencia, el folio del desacuerdo, las primeras tres letras del nombre del niño o niña, su grado escolar, el número de fragmento del texto al que pertenece el desacuerdo, el desacuerdo (marcado en negritas) y cómo está escrito en el texto.

| Tipo | Descrip.<br>F/si/se | Folio  | Nombre | Gr | Frag. | Desacuerdo                            | Debe ser                          |
|------|---------------------|--|--------|----|-------|---------------------------------------|-----------------------------------|
| 1    | s s s               | 207  | Ing    | 4  | 38    | Sé que ha-hay<br><b>diferentes</b>    | Sé que hay<br>delfines            |
| 2    | s s n               | 159  | Arl    | 2  | 30    | de que un <b>dibujo</b>               | de que un<br>tiburón              |
| 3    | s n s               | 78   | Irv    | 2  | 25    | para que los peces<br><b>quedaron</b> | para que los<br>peces<br>quedaran |
| 4    | n s s               | 82   | Irv    | 2  | 35    | de mi papá y<br><b>cuando</b>         | de mi papá y<br>nadó              |
| 5    | s n n               | 65   | Mar    | 2  | 29    | la cuerda jaló de<br><b>huevo</b>     | la cuerda<br>jaló de<br>nuevo     |
| 6    | n s n               | 15   | Ale    | 2  | 35    | se co:lo-có <b>junto</b> el           | justo en                          |
| 7    | n n s               | No se encontraron desaciertos de este tipo, en los datos analizados. |        |    |       |                                       |                                   |
| 8    | n n s               | 74   | Irv    | 2  | 12    | este puede ser<br><b>al</b>           | éste puede<br>ser un              |
| 9    | n n n               | 210  | AnaM   | 4  | 11    | oí que \ / un<br>banco                | oí qué hay<br>un banco            |

Los desaciertos no léxicos que, como ya se mencionó anteriormente, se caracterizan porque la unidad oralizada no es una palabra real del lenguaje o cuando el niño procesa unidades menores a la palabra, es decir, subléxico, lo leído no es la unidad subléxica efectivamente escrita (leer “chorri” sobre chori(z)o ).

Este tipo de desaciertos puede ser interpretado como producto de una deficiencia perceptual, ya que se omiten letras, se sustituyen letras parecidas, se invierten, etc. Lo que Vaca (1997) intenta demostrar es que también pueden ser explicados recurriendo a procedimientos inferenciales de sílabas y no de palabras. Es importante mencionar que las sílabas inferidas erróneamente, son por lo general sílabas posibles del español, lo cual reafirma que el niño está haciendo uso de su conocimiento del lenguaje. Vaca distingue en su propuesta 10 tipos de desaciertos no léxicos:

Tipo 1. *Código fonográfico*. Son las reglas de correspondencia entre lo gráfico y lo sonoro y que al no conocerlas, un niño comete este tipo de desaciertos.

Tipo 2, 3 y 4. *Sustitución del valor sonoro asignado a una grafía*. Se distinguen tres tipos de sustitución de acuerdo a la similitud gráfica entre la grafía a la que se asigna un valor sonoro y la grafía a la que convencionalmente se asigna el fonema que fue asignado a la primera. En el caso del tipo 2, la sustitución es entre grafías muy similares, por ejemplo asignar /p/ a q, /b/ a d. En el tipo 3 se contemplan aún similitudes gráficas, por ejemplo entre o, e, a; y, por último, en el caso 4 se sustituye el valor sonoro entre grafías claramente diferentes, por ejemplo a/i, g/n, etc.

Tipo 5, 6, 7 y 8. *Inserción, omisión o cambio de orden*. En el tipo 5, además de las grafías escritas, el niño oraliza un fonema no escrito. En el tipo 6 sucede lo contrario, no se oraliza una grafía y en el tipo 7 se altera aparentemente el orden de las grafías (y/o fonemas) dentro de una sílaba; en el tipo 8, la sílaba oralizada está muy alejada de la sílaba escrita; se cree que no se tomó suficiente información para realizar la inferencia.

Tipo 9. *Interferencia léxica*. En este tipo de desaciertos se pueden apreciar dos palabras que entran en juego durante la oralización de una, generando una palabra no existente.

Tipo 10. *Divergencia en el esquema acentual*. Este tipo de desaciertos se comete cuando se atribuye un esquema acentual diferente a las palabras leídas. Estos desaciertos no serán analizados en este trabajo, ya que están siendo estudiados desde la perspectiva de la ortografía.

## Corpus Mantarraya: Análisis de desaciertos...

Ejemplos de desaciertos no léxicos:

| <b>Tipo</b> | <b>Folio</b> | <b>Nombre</b> | <b>Gr</b> | <b>Fr</b> | <b>desacuerdo</b>        | <b>Debe ser</b> |
|-------------|--------------|---------------|-----------|-----------|--------------------------|-----------------|
| 1           | 142          | Arl           | 2         | 1         | chorri-chorrizo          | chorizo         |
| 2           | 151          | Arl           | 2         | 25        | /atra-bá-dos/            | atrapados       |
| 3           | 79           | Irv           | 2         | 25        | atropados                | atrapados       |
| 4           | 67           | Mar           | 2         | 31        | se fa--fo-zafó           | se zafó         |
| 5           | 63           | Mar           | 2         | 27        | se le habida-            | se le había     |
| 6           | 124          | Lau           | 2         | 27        | cuando de /proto/ sintió | de pronto       |
| 7           | 114          | Edu           | 2         | 35        | /ba-nó/                  | nadó            |
| 8           | 145          | Arl           | 2         | 19        | /páp-sa-be/-pásam-me     | pásame          |
| 9           | 139          | Arl           | 2         | 16        | rápido-mente             | rápidamente     |

La tercera clase de desaciertos propuesta para este modelo corresponde a las oralizaciones, que por su rápida corrección no brindan la información suficiente para incluirlas en una de las categorías anteriores; se supone que los procesos que las generan son los anteriormente descritos.

Como se mencionó, este trabajo se centrará en tres aspectos específicos: se aplica la clasificación propuesta por Vaca (1997), se analiza las omisiones de palabras y se realiza una comparación numérica entre las palabras leídas y las escritas.

### Método

Los datos aquí analizados corresponden a una parte del “Corpus mantarraya”, descrito más adelante; se seleccionaron específicamente los fragmentos que corresponden a desaciertos y se clasificaron de acuerdo a su Clase y Tipo con base en el modelo descrito anteriormente.

La descripción metodológica sobre la elaboración del Corpus es retomada de Vaca (2000):

Consta principalmente de una serie de entrevistas clínicas realizadas a 20 niños hispanófonos y 20 francófonos<sup>3</sup> (entre otros datos recabados y transcritos), que proporciona información sobre:

- la conceptualización de la escritura infantil de pares homófonos de palabras (ortografía en circunstancias de justificación);
- la conceptualización de la escritura de cuasihomófonos (ortografía en circunstancias de justificación);
- la lectura de frases homófonas;
- la lectura en voz alta de un cuento (procesamiento y comprensión);
- la redacción de un resumen;
- la ortografía de uso por parte de los niños (de palabras y de un texto).

Puesto que en este trabajo sólo analizaremos los desaciertos cometidos por la población hispanófona, sólo se describen las características metodológicas que son pertinentes.

### Sujetos

Los sujetos del corpus tienen las siguientes características:

|                  | 2º de primaria | 4º de primaria |
|------------------|----------------|----------------|
| Niños            | 6              | 3              |
| Niñas            | 4              | 7              |
| Promedio de edad | 8.4            | 9.8            |

Criterios de preselección de los niños:

1. Los niños debían ser de lengua materna española y francesa, para asegurar el dominio de la lengua oral correspondiente esperado en niños de esa edad. Para eso, nos aseguramos que los padres de los niños fueran igualmente francófonos e hispanófonos, y de que los niños hubieran nacido y vivido en Lyon (Francia) o Xalapa (Méjico).

## Corpus Mantarraya: Análisis de desaciertos...

2. Se decidió trabajar con niños de 7-8 años porque a esa edad el niño está recientemente alfabetizado y ya es capaz de leer un pequeño cuento. Además, se decidió trabajar con niños de 9-10 años para comparar las conceptualizaciones y los desempeños con dos años de diferencia con relación a la experiencia con la lengua escrita. Se trata pues de un estudio transversal.
3. Los niños debían seguir el nivel escolar correspondiente a su edad; lo anterior con el objeto de reducir la posibilidad de que los niños hubieran presentado problemas de desempeño escolar particulares.
4. Los niños debían provenir de grupos escolares diferentes, para intentar neutralizar la influencia que un solo maestro podría tener sobre los niños.

### El texto leído

El texto en su versión original para segundo año aparece en el anexo 2 y para tercero en el anexo 3 (Vaca, 2000).

### Procedimientos de entrevista

Las entrevistas fueron realizadas a medio camino entre la entrevista 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. Se realizaron individualmente en un salón de la escuela, registrándose con una grabadora. En primer lugar, el entrevistador se presentó con el niño, platicó libremente con él. Le explicó que las actividades que iban a realizar no tenían que ver con la escuela ni con su evaluación escolar y que sólo tenía interés en conocer su manera de trabajar y pensar.

En lo que respecta a la lectura del cuento, el procedimiento fue 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.

Si el niño no comprendía el título o si interpretaba “raya” de una manera diferente a la esperada, el entrevistador le explicaba que las rayas son peces a los que también se les llama “mantarrayas”.

Una vez leída cada parte del texto, simplemente se le decía: entonces dime qué comprendiste, o bien se le preguntaba ¿de qué te acuerdas?, ¿de qué se trata esta parte del cuento?. Si el niño decía muy pocas cosas, era animado preguntándole de qué otra cosa se acordaba.

### Análisis

De las transcripciones de las entrevistas se hizo una cuantificación y clasificación de los desaciertos, de acuerdo a las categorías propuestas en el modelo de Vaca (1997) señaladas en el apartado de antecedentes.

Se tomó como unidad de análisis a la palabra, con fines de cuantificación, sin que por eso creamos que el procesamiento del texto se da sólo en ese nivel. Los desaciertos fueron clasificados tratando de apegarse a los criterios marcados en la descripción de la clasificación; sin embargo, algunos desaciertos podían caber en dos categorías o muestran un procesamiento más complejo del que aparentan; en estos casos se tomó la decisión de clasificarlos en la categoría más cercana a sus características.

Los niños produjeron durante la lectura 476 desaciertos. 176 tienen que ver con aspectos ortográficos y no serán analizados en este trabajo, ya que implican otro tipo de análisis.

## Corpus Mantarraya: Análisis de desaciertos...

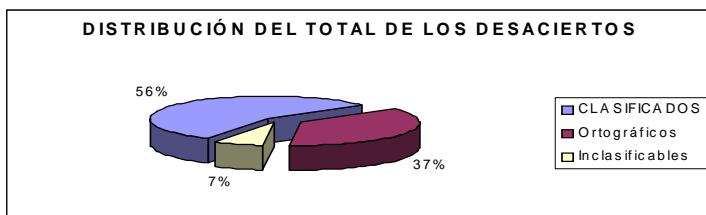
Por otro lado, se encuentran 32 desaciertos que por su naturaleza no fue posible clasificarlos en ninguna categoría; por esta razón no se toman en cuenta en el análisis. Por lo tanto, el número de desaciertos con los que se trabajó es 268. Lo anterior se representa en la gráfica 1.

**GRÁFICA 1: Distribución total de los desaciertos producidos por los 20 niños hispanófonos**

| Clase      | 2°                | 4°                | Total de desaciertos |
|------------|-------------------|-------------------|----------------------|
| Léxicos    | 108 48%<br>(75%)  | 115 52%<br>(92%)  | 223 83%              |
| no léxicos | 36 80%<br>(25%)   | 9 20%<br>(8%)     | 45 17%               |
| Total      | 144 53%<br>(100%) | 124 47%<br>(100%) | 268 (100%)           |

Los desaciertos clasificados se distribuyen como se observa en la tabla 2.

**TABLA 2: Distribución de los desaciertos analizados en este trabajo.**  
(Los porcentajes orientados horizontalmente corresponden a la clase de desaciertos y los orientados verticalmente al grado)



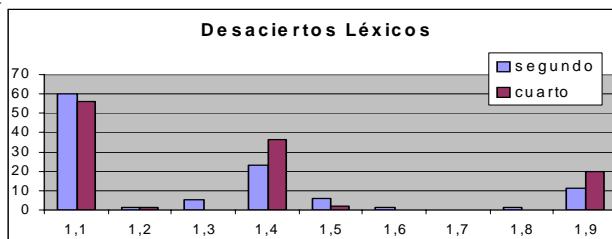
Como se muestra en la tabla, el 83% de los desaciertos son léxicos, es decir, es la sustitución de una palabra por otra; esto coincide con Kollers (1985). También se puede observar que los niños de 2º año produjeron sólo ligeramente más desaciertos –144; el 53%– que los niños de 4º –124; el 47%. Esto significa que la diferencia entre estos dos grupos se encuentra en la clase de desaciertos cometidos y no en la cantidad.

De los 268 desaciertos clasificados, 223 (83.2%) son léxicos, lo cual quiere decir que los niños principalmente utilizan, entre otras, la información sintáctica, semántica y/o fonográfica para acceder a las unidades del texto escrito.

Los desaciertos no léxicos son 45 y la mayoría (4/5 partes) son cometidos por los niños de menor edad. Esto nos indica que los niños de segundo grado aún procesan unidades subléxicas a diferencia de los niños mayores que, con mayor experiencia en la lectura, procesan el texto con unidades mayores como palabras o grupos de palabras. Este dato nos muestra un claro ejemplo de evolución.

Veamos lo que sucede dentro de los desaciertos léxicos (clase 1) en la gráfica 2.

**GRÁFICA 2:** Distribución de los desaciertos léxicos.



Los niños de 4º año produjeron el 52% (115) de los desaciertos léxicos, los de segundo el 48% (108), estos porcentajes parecen no variar mucho; sin embargo, en cuanto al tipo de desacierto se encontró lo siguiente: se puede observar que los desaciertos léxicos de tipo 1, los que toman en cuenta al menos las tres clases de información para realizar la inferencia (sintáctica, semántica y fonográfica), tienen la mayor frecuencia: obtienen un porcentaje de 52% (116 desaciertos) del cual 52% (60 desaciertos) corresponde a los niños de 2º y el 48% (56 desaciertos) a los de 4º, eso muestra que efectivamente los niños se basan en esta información para que la inferencia realizada concuerde con el texto leído previamente.

Pareciera paradójico que más niños pequeños manejen al menos las tres clases de información, pero no lo es, puesto que al leer privilegiando las sílabas como unidad de procesamiento reducen la velocidad de lectura, permitiéndoles coordinar mejor la información disponible; en cambio, los niños mayores deben sacrificar cierta información para ganar fluidez y velocidad, privilegiando, por ejemplo, la información semántica y sintáctica.

## Corpus Mantarraya: Análisis de desaciertos...

Por otro lado, hay una concentración de desaciertos en el tipo 4 con una frecuencia de 59, de los cuales 23 son de los niños de segundo y 36 de los niños de cuarto; este tipo de desacuerdo se refiere a que toman en cuenta la información sintáctica y semántica pero no toman en cuenta la información fonográfica. Congruente con lo anterior, creemos que los niños mayores produjeron más este tipo de desacuerdo porque se encuentran en un nivel evolutivo en lectura mayor que los de segundo, hacen uso de su experiencia y de su conocimiento realizando inferencias sintáctica y semánticamente posibles con lo leído previamente, descuidando el aspecto fonográfico, y de esta manera pueden leer rápida y fluidamente; pero en ocasiones estas inferencias no son acertadas, dando lugar a este tipo de desaciertos. Por otro lado, los de segundo utilizan predominantemente unidades pequeñas de procesamiento (véase Vaca, 1997); de esta manera centralizan su atención en la información fonográfica, reduciendo mucho la probabilidad de caer en este tipo de desaciertos; además, si tomamos en cuenta que al utilizar unidades pequeñas disminuye la velocidad de lectura, pueden manejar al mismo tiempo la información sintáctica y semántica.

En estos dos tipos de desaciertos (1 y 4) se concentra el 78 % de los desaciertos léxicos, lo cual concuerda con los resultados de Vaca (1997).

El otro tipo de desacuerdo donde se observa una frecuencia relativamente alta es el tipo 9, representando el 14% (31 desaciertos) de los desaciertos léxicos, y corresponden a los que no pueden clasificarse en alguno de los tipos anteriores debido a que no se cuenta con suficiente información, puesto que, según el modelo utilizado, con fines de una exploración cuantitativa se analizan los desaciertos con lo leído previamente por el niño, pero no se toma en cuenta lo que se leyó después del desacuerdo. Todos los desaciertos de los datos analizados que fueron clasificados en esta categoría corresponden a omisiones de palabras. Por lo tanto, se decidió que sólo para el caso de las omisiones se tomará en cuenta la palabra posterior a la omisión, para que en función de ésta pueda juzgarse la semantidicidad y sintacticidad de lo leído. Un ejemplo de este tipo es el siguiente:

| Tipo | Folio | Nombre | Gr | Fr | desacuerdo                  | Debe ser            |
|------|-------|--------|----|----|-----------------------------|---------------------|
| 9    | 210   | AnaM   | 4  | 11 | oí que <b>\hay</b> un banco | oí que hay un banco |

En este caso AnaM omitió la oralización de la palabra hay, pero no tenemos información previa que nos permita definir la sintacticidad, a menos que apliquemos el criterio antes expuesto. De esta forma ya se cuenta con información para decir que la manera como leyó este fragmento del cuento es sintáctica y semánticamente posible y podría ser colocado en el tipo 4. En el que, si observamos la gráfica 2, vemos que los niños de cuarto tienen una mayor frecuencia en este tipo de desacuerdo; este dato es congruente con la idea del desapego que los niños tienen respecto del texto cuando adquieren mayor habilidad, fluidez y velocidad. Realizamos un análisis detallado de este tipo al final de este trabajo.

Con menor frecuencia se encuentra el tipo 5, con 8 desaciertos (4% del total), de los cuales 6 corresponden a los niños de segundo y 2 a los de cuarto. Este tipo se refiere a que sólo toman en cuenta la información fonográfica para realizar la inferencia; este dato es coherente con el caso anterior, es decir, acerca de que los niños pequeños tienden a considerar más la información fonográfica que los niños mayores.

El tipo 3 obtuvo una frecuencia de 5 desaciertos (2.2%), todos producidos por los niños de segundo; este tipo expresa una centración en la información semántica y fonográfica, lo cual puede ser un indicativo de que la información sintáctica es de las últimas en ser coordinadas por los niños. Esto lo observamos nuevamente en la gráfica 2 en los tipos 8, 5 y 3, en los que la información sintáctica no es considerada y en donde los niños de segundo grado presentan una frecuencia mayor a los de cuarto.

Los desaciertos de tipo 2 tienen una frecuencia de 2 (0.5%) y corresponden a los que, para realizar la inferencia, utilizan la información fonográfica y la sintáctica pero no se toma en cuenta la información semántica. Esto puede significar que la información semántica es un factor indispensable en la formulación de las inferencias; esto es congruente con van Dijk y Kintsch (1983) en donde mencionan que la comprensión de textos es determinada por la macroestructura generada por el lector a partir del contexto semántico del texto. Sin embargo, en nuestros datos observamos que ningún niño produjo un

## Corpus Mantarraya: Análisis de desaciertos...

desacuerdo de tipo 7 (sólo toma en cuenta la información semántica); dada la importancia del aspecto semántico en la realización de inferencias, este dato parece indicar que esta información no es frecuentemente utilizada independientemente de otras clases de información.

El porcentaje que le sigue corresponde al tipo 6 con menos del 1%. En este tipo de desacuerdo hay una centración sólo en la información sintáctica. Nuevamente son pocos los desaciertos en los que la información semántica no es tomada en cuenta, lo cual enfatiza su importancia.

El mismo porcentaje corresponde al tipo 8, en el que no se toma en cuenta ninguno de los tres tipos de información. Este dato nos indica que es muy poco probable que los niños no utilicen al menos la información sintáctica, semántica o fonográfica.

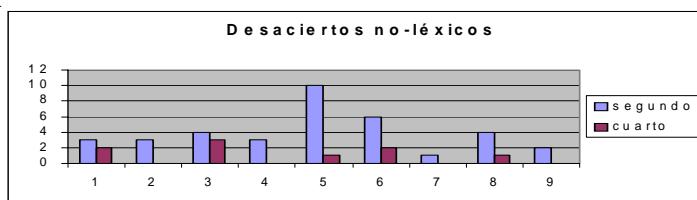
### Desaciertos no léxicos (clase 2)

El número de los desaciertos no léxicos es de 45, equivalente al 16.79% del total. De éstos, 36 (80%) son producidos por los niños de 2º año, lo cual es un dato interesante. El 20% restante es de los niños de 4º. Esta diferencia nos indica un salto evolutivo entre los niños de cuarto y los de segundo. La oralización de palabras inexistentes, producto de inferir unidades subléxicas como sílabas, es un fenómeno que tiende a disminuir en la medida que el niño adquiere mayor habilidad en la lectura, pasando después al privilegio de unidades mayores.

En el caso de los niños de 4º que producen esta clase de desaciertos, posiblemente sea debido a que, por cierta dificultad en el procesamiento, les fue necesario recurrir a los mecanismos anteriormente utilizados (tomar a la sílaba como unidad, por ejemplo), pero esto sólo en casos especiales. Es importante mencionar que algunos de los desaciertos no léxicos pueden ser producidos en el nivel de la oralidad, es decir, en la reproducción del texto y no en su procesamiento.

Veamos cómo se distribuyen las frecuencias en cuanto al tipo de desacuerdo en la gráfica 3:

GRÁFICO 3. Distribución de los desaciertos no léxicos.



En principio, podemos observar una alta frecuencia en el desacuerdo del tipo 5, con 11, y en el tipo 6 con una frecuencia de 8; estos desaciertos son producidos cuando se omite o se añade un fonema dando lugar a la oralización de una palabra inexistente; por ejemplo:

| Tipo | Folio | Nombre | Gr | Fr | desacuerdo              | Debe ser    |
|------|-------|--------|----|----|-------------------------|-------------|
| 5    | 185   | Car    | 4  | 32 | se <b>atravesió</b> -só | se atravesó |

Si reconstruimos la lectura, Car lee: “fue atacado porque tenía hambre en ese momento se ....” justo antes de oralizar el desacuerdo, muchas palabras pueden ser inferidas (espantó, desmayó, etc.), pero Car usa la información que tiene a su disposición, por ejemplo las letras que tiene la palabra, el tamaño de la misma, la información semántica y las reglas sintácticas. Entonces lee: /**atravesió**-so/. ¿Qué provocó que leyera así la última sílaba de esa palabra? Posiblemente fue influida por la evocación de la palabra *atravesía* que, al mantener presente el fonema /i/, lo integra en la oralización. Es posible también que sólo sea un desacuerdo en el nivel de la producción, algo frecuente en la producción oral del lenguaje (véase Rossi y Peter-Defare, 1998).

El siguiente tipo con mayor frecuencia es el 3, con 7 desaciertos, correspondiente a la oralización de un fonema con cierta similitud gráfica al escrito dentro de una palabra. En el siguiente ejemplo se muestra una de las dificultades al intentar atribuir una causa a este tipo de desaciertos:

| Tipo | Folio | Nombre | Gr | Fr | desacuerdo       | Debe ser  |
|------|-------|--------|----|----|------------------|-----------|
| 3    | 79    | Irv    | 2  | 25 | <b>atropados</b> | atrapados |

## Corpus Mantarraya: Análisis de desaciertos...

La similitud gráfica que tiene la grafía a con la o podría ser la causa de este desacuerdo; además, dentro de la misma palabra se encuentra una o, que como una anticipación a nivel de palabra también pudo haber causado esta interferencia.

Con respecto al tipo 8, que involucra un procesamiento complejo, se nota una diferencia entre segundo (80%) y cuarto año (20%); la producción de este tipo puede atribuirse, como se menciona en Vaca (1997), a que el lector toma muy poca información para inferir la sílaba; además creemos que este tipo de desacuerdo puede ser ocasionado por un procesamiento muy rápido, en donde según la clasificación intervendrían diferentes tipos de desaciertos, por ejemplo:

| Tipo | Folio | Nombre | Gr | Fr | desacuerdo           | Debe ser |
|------|-------|--------|----|----|----------------------|----------|
| 8    | 145   | Arl    | 2  | 1  | /páp-sa-be/-pásam-me | pásame   |

Puede ser que en este caso exista una interferencia léxica entre /papá/ y /sabe/, que correspondería al tipo 9, o también la inserción de fonemas, es decir, el tipo 5; otra posibilidad es una sustitución de fonemas diferentes correspondiente al tipo 4, o quizás, todo al mismo tiempo, dando como resultado esta producción.

También se encuentran con la misma frecuencia de 3 los tipos 2 y 4, *sustitución de una grafía muy similar y diferente*, respectivamente. Ejemplos:

| Tipo | Folio | Nombre | Gr | Fr | desacuerdo    | Debe ser  |
|------|-------|--------|----|----|---------------|-----------|
| 2    | 151   | Arl    | 2  | 25 | /atra-bá-dos/ | atrapados |

En el ejemplo se puede ver uno de los desaciertos más comunes en este tipo en donde la similitud de la /p/ con la /b/ provoca otorgarle un valor sonoro diferente. Es importante señalar que el oralizar este tipo de desaciertos no necesariamente tiene que ver con problemas viso-perceptuales puesto que, siguiendo el modelo en el que nos basamos, es posible generar una interpretación de estos desaciertos desde el punto de vista de la inferencia. En este caso, Arl decidió otorgarle un valor sonoro distinto a la p basándose en la información disponible y de la cual privilegia sólo una parte, la información fonográfica, dando como resultado una inferencia no acertada.

En el tipo 7 se altera el orden. Fue el de menor frecuencia, con sólo 1 caso y fue generado por un niño de 2°; se clasificó en esta categoría porque es a la que más se aproxima, pero consideramos que si cambió el orden fue de una manera muy especial, lo cual indica un procesamiento muy particular:

| Tipo | Folio | Nombre | Gr | Fr | desacuerdo | Debe ser |
|------|-------|--------|----|----|------------|----------|
| 7    | 114   | Edu    | 2  | 35 | /ba-nó/    | nadó     |

Si cambió el orden sería de la siguiente manera: cambió sólo las consonantes pero, además, invirtió la /d/ por la /b/, algo que, como ya dijimos, es muy común; lo especial es que se haya presentado todo junto y que no lo haya corregido.

El análisis de esta clase de desaciertos nos hace coincidir con la idea de que en la oralización de los desaciertos no léxicos interviene el mismo mecanismo que genera los desaciertos léxicos, es decir, en los que entran en consideración por lo menos las tres clases de información (fonográfica, sintáctica y semántica), sólo que por ser unidades de procesamiento más pequeñas requieren otro tipo clasificación y análisis.

### Omisiones

Al intentar aplicar el sistema de clasificación de los desaciertos en el que nos estamos basando para el análisis de estos datos, nos encontramos con el problema de las omisiones, a las que, por falta de información, era necesario colocar en el tipo 9, errores que no puedan ser clasificados en las otras categorías.

Al realizar un análisis detallado de ellas encontramos que el 64% de las omisiones fueron producidas por los niños de 4° año, y hubo 7 niños que no omitieron palabras en su lectura (4 de 2° y 3 de 4°). También hay fragmentos en los que varios niños omiten la misma palabra; esto puede indicarnos que ciertas estructuras sintácticas favorecen las omisiones.

## Corpus Mantarraya: Análisis de desaciertos...

Otro dato importante es que sólo el 16% de las omisiones fueron corregidas (esto es: 5 de 31), lo que significa que la mayoría de estas omisiones no causa ningún problema en el procesamiento, caso contrario en quienes sí corrigieron la omisión:

| Folio | Nombre | Gr | Fr | Desacuerdo                 | Debe ser    |
|-------|--------|----|----|----------------------------|-------------|
| 125   | Lau    | 2  | 27 | porque \se/ l:-se le había | se le había |

En este ejemplo la omisión de la palabra *\se/* rompe la semanticidad de la oración, por lo que al intentar oralizar la palabra *le* corrige inmediatamente.

| Folio | Nombre | Gr | Fr | Desacuerdo   | Debe ser  |
|-------|--------|----|----|--|-----------|
| 131   | Arl    | 2  | 6  | en-ton-ces- salió- corriendo- de-mi- cuarto--ly/ el-y el- di-j:e-di-dije | salí<br>y |

En este ejemplo vemos que primero lee “salió” en donde dice *salí*, además lee “el” en donde está escrito *le*. Esto se debe a que al leer /salió/ necesita estructurar la oración para que sea semánticamente correcta, lo cual no fue posible mantener al leer acertadamente la palabra /dije/.

Otra de las omisiones corregidas es:

| Folio | Nombre | Gr | Fr | Desacuerdo                      | Debe ser |
|-------|--------|----|----|---------------------------------|----------|
| 215   | AnaM   | 4  | 30 | se dió cuenta \de/ que:--de que | de que   |

La omisión de la palabra */de/* no afecta la sintaxis ni la semántica de la oración. Probablemente se produjo porque este tipo de estructura es muy común o es como debería ser escrita.

| Folio | Nombre | Gr | Fr | Desacuerdo                             | Debe ser       |
|-------|--------|----|----|--|----------------|
| 217   | AnaM   | 4  | 37 | cuando mi papá \se/ repu-se repuso del | papá se repuso |

En este ejemplo AnaM se ve obligada a corregir porque se rompe la sintaxis de la oración con la palabra */del/*, si suponemos que iba a leer /mi papá repuso.../, lo cual es sintácticamente válido, pero al leer */del/* requiere de una corrección.

Otra omisión corregida es:

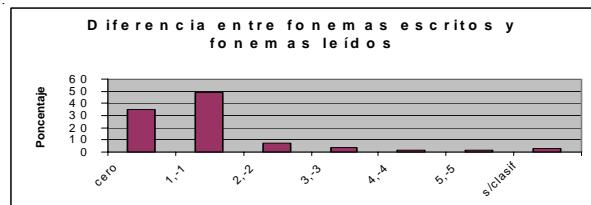
| Folio | Nombre | Gr | Fr | Desacuerdo  | Debe ser          |
|-------|--------|----|----|---|-------------------|
| 279   | Wen    | 4  | 32 | fue a atacarlo porque tenía-hambre- <del>en/</del><br>ese-en ese momento--se atravesó | en ese<br>momento |

Aquí nuevamente la sintaxis de la oración se ve afectada al omitir la palabra \en/ por lo que creemos que Wen corrigió.

### El tamaño de la palabra, información básica para realizar inferencias

Para comprobar si el tamaño de la palabra es un factor que determina la inferencia, se realizó una cuantificación de las letras en las palabras sustituidas y las comparamos con el número de letras de la palabra inferida; el resultado se muestra en la gráfica 4.

GRÁFICO 4. Distribución de la variación entre los fonemas leídos y los escritos.



El mayor porcentaje se concentra en las palabras de uno y cero fonemas de diferencia, lo cual significa que las palabras sustituidas son de tamaño muy similar a las escritas; esto es obvio, puesto que es difícil que se infiera una palabra pequeña en donde está escrita una grande y viceversa, al menos a esta edad. Con lo anterior queda confirmado que además de contar con la información semántica, sintáctica y fonográfica, los lectores pueden utilizar el tamaño de la palabra para realizar la inferencia, entre otras muchas clases de información.

Estos datos concuerdan con los reportados por Kollers (1985), quien afirma que las palabras sustitutas son aproximadamente iguales a las sustituidas.

## Corpus Mantarraya: Análisis de desaciertos...

### El tamaño de las palabras omitidas

Realizamos una cuantificación de letras de las palabras omitidas dando como resultado la tabla 3:

TABLA 3. Cuantificación de las letras de las palabras omitidas.

| Letras de la palabra | Frecuencia |
|----------------------|------------|
| 1                    | 5          |
| 2                    | 19         |
| 3                    | 5          |
| 4                    | 2          |
| total                | 31         |

Las palabras de dos letras son las que se omiten con mayor frecuencia, en menor cantidad las de una letra y las de tres. Ahora, si observamos los datos nos damos cuenta que son preposiciones, conjunciones, etc., cuya omisión generalmente no afecta ni la sintaxis ni la semántica de la oración.

### Conclusiones

En este trabajo se revisaron algunos aspectos correspondientes a los desaciertos de la lectura; a partir del análisis de cada uno de ellos es posible generar las siguientes conclusiones:

Tener desaciertos durante la lectura es una de las características de un procesamiento activo, en donde el lector aporta su conocimiento para poder acceder a lo que está escrito. Un dato interesante de este análisis es que todos los niños de la muestra presentan desaciertos, lo cual confirma que son inherentes al proceso de lectura –en promedio 22.7 los de segundo y 19.9 los de cuarto–, y las características de estos dependen, entre otras cosas, de su conocimiento del lenguaje, de su experiencia en la lectura y del tipo de texto leído.

El niño trata de acceder al texto empleando para ello todos los recursos que tiene a su disposición. Para adquirir fluidez en la lectura, es necesario inferir unidades mayores y realizar un cierto desapego del texto; por tal motivo, el lector requiere hacer mayor uso del contexto semántico, sacrificando un poco las otras clases de información.

Pudimos observar que los niños de segundo año realizan un procesamiento que favorece la aparición de desaciertos no léxicos, a diferencia de los de cuarto que, con un procesamiento más evolucionado, reducen considerablemente esta clase de desaciertos.

Generalmente, cuando se infiere una palabra distinta a la que está escrita, cambiando con ello el significado de la oración, nos vemos obligados a intentar acomodar las nuevas palabras leídas para que se mantenga la coherencia; cuando esto es posible generalmente el desacuerdo no se corrige.

El análisis de los datos del *corpus mantarraya* nos permitió verificar que el tamaño de la palabra también es una información básica que el lector usa para realizar sus inferencias.

Las palabras omitidas, no se corrigen generalmente cuando su ausencia no afecta el aspecto sintáctico y semántico de la oración.

### Notas

<sup>1</sup> Preferimos emplear la palabra desacuerdo para evitar la connotación negativa que generalmente tiene el término error.

<sup>2</sup> Se considera al “desacuerdo” como la variación entre lo escrito y lo leído por el niño.

<sup>3</sup> Datos que corresponden a la tesis doctoral elaborada por Jorge Vaca Uribe (2000)

## Corpus Mantarraya: Análisis de desaciertos...

### Bibliografía

- Carrasco (1999). "La aplicación incompleta de inferencias, una explicación a los problemas de comprensión de lectura." En *Memoria electrónica del V Congreso de Investigación Educativa*. Aguascalientes, 30, 31 de noviembre, y 1 y 2 de diciembre.
- Goodman, K. (1967). "Reading: A Psycholinguistic Guessing Game." In Singer, H., Ruddell, B. (eds.). 1976, *Theoretical models and processes of reading*. Newark: IRA.
- Kolers, P. (july, 1972). "Experiments in reading." *Scientific American*, vol. 227,1, 84-92.
- Kolers, P. (1985). "Skill in reading and memory." *Canadian Journal of Psychology*. 39(2), 232-239.
- Rossi, Mario & Peter-Defare, E. (1998). *Les lapsus*. Paris: PUF.
- Smith, F. (1983). *Comprensión de la lectura. Análisis psicolingüístico de la lectura y su aprendizaje*. México: Trillas.
- Vaca, J. (1986). *Lo no alfabetico en el sistema de escritura: ¿qué piensa el escolar?* Tesis de maestría. México: DIE-CINVESTAV-IPN
- \_\_\_\_\_ (1997). *El niño y la escritura*. Xalapa: U.V.
- \_\_\_\_\_ (2000). *Morfogramas, logogramas de la lectura infantil*. Tesis de doctorado. Lyon: Université Lumière Lyon 2.
- van Dijk, T. & Kintsch, W. (1983). *Strategies of discourse comprehension*, New York: Academic Press.