

SISTEMA GRAFICO Y SISTEMA ORTOGRAFICO (O BIEN: NO TODO ES ORTOGRAFICO EN LA ADQUISICION DE LA ORTOGRAFIA)

Emilia Ferreiro

El modo clásico de estudiar las producciones infantiles para evaluar el desempeño ortográfico consiste en focalizar los errores y clasificarlos en tres o cuatro categorías: omisiones, sustituciones, adiciones y transposiciones o permutaciones. Estamos tan habituados a utilizar esta clasificación que es difícil pensar en otra (a lo sumo, se crean sub-categorías dentro de esas grandes categorías). El problema con este modo de clasificar las desviaciones es que sólo permite evaluarlas en términos de la distancia con respecto a la norma actual. Pone en relieve los desaciertos y no permite focalizar los aciertos. Además, no nos dice nada con respecto a los conocimientos efectivos que se van adquiriendo y dice muy poco sobre las motivaciones que podrían justificar los desaciertos.

La distinción entre *sistema gráfico* y *sistema ortográfico* ya fue utilizada con éxito por C.Pontecorvo y N.Ribeiro Moreira (cap. in Ferreiro, Pontecorvo & al, 1996). Utilizando la extensa base de datos constituida para analizar narrativas infantiles (re-escrituras de la historia tradicional de Cappuccetto Rosso), estas autoras contrastaron las formas de graficar el nombre de la protagonista en italiano y portugués. En el presente análisis trataremos de extender a todas las palabras utilizadas en el corpus una distinción similar, aunque con algunas modificaciones por las razones que expondremos a continuación.

Los textos de nuestro corpus presentan múltiples desaciertos ortográficos. Razón de más para intentar cambiar la mirada y tratar de ver cuáles son los desaciertos que NO aparecen. En efecto, todo dato relativo a combinaciones de letras que no se registran (en un corpus donde los errores ortográficos abundan)

puede ser prueba de la apropiación (inconciente) de principios combinatorios que son propios de la escritura de la lengua que estos escritores debutantes están tratando de aprender.

En las condiciones de producción de los textos de nuestro corpus, la atención de los niños no estaba dirigida a los aspectos gráficos ni ortográficos sino a los aspectos textuales (cómo contar la historia, qué elecciones léxicas tomar, cómo nombrar a los personajes en momentos sucesivos de la historia, cómo encadenar los episodios, cómo dar la voz a los personajes, etc.). Es posible que, en las edades estudiadas, la corrección ortográfica corresponda a una etapa posterior, que podemos llamar de "revisión del texto" (etapa que pocas propuestas didácticas toman en consideración). Precisamente por esas mismas restricciones en las condiciones de producción, los datos de nuestro corpus pueden considerarse ejemplares, más por lo que NO aparece en ellos que por lo que sí aparece.

Uno de los problemas que enfrenta el investigador es que no hay una terminología consensuada para designar los distintos fenómenos involucrados. El termino "ortográfico" puede cubrirlos a todos pero pierde especificidad. Gak (1976) hizo una contribución importante al proponer distinguir entre "sistema gráfico" y "sistema ortográfico". Según Gak, el sistema gráfico se refiere a "los medios de que dispone una lengua para expresar los sonidos", entendiendo por tal el conjunto de letras y diacríticos. El sistema ortográfico se relaciona con las reglas que determinan "el empleo de las letras según las circunstancias". Las posibilidades de elección entre grafías alternativas para expresar un mismo sonido están determinadas, según Gak, por el sistema gráfico, pero los casos en que una determinada grafía se impone, impidiendo la ocurrencia de otras igualmente posibles, es del dominio ortográfico. En consecuencia, si una escritura alfabética fuera totalmente regular, o sea, si tuviera una correspondencia biunívoca entre letras y sonidos, esa escritura tendría un sistema gráfico pero no se podría hablar de ortografía.

En un trabajo contemporáneo al de Gak pero independiente -también tomando como problema central la ortografía del francés- Blanche-Benveniste y Chervel (1974: 117-120) sostienen una distinción similar. Estos autores hablan de la necesidad de describir el "material gráfico" (*matériau graphique*) antes de tratar las relaciones fonema-grafema. Sostienen que "nombre des questions que l'on a l'habitude de traiter dans l'étude de la phono-graphie relèvent exclusivement de la description du matériau graphique" (p.122)¹. Sin embargo, dejan sin resolver el problema de cómo hacer la descripción del material gráfico y dónde ubicar la frontera entre lo gráfico y lo ortográfico.

Desde un punto de vista psicolingüístico, Cassar y Treiman (1997), en un estudio sobre las consonantes o vocales que pueden duplicarse en inglés, y las restricciones posicionales de las mismas, sostienen que:

"..much evidence now points to the early emergence of orthographic knowledge. This knowledge about the written language is to some degree separate from phonological knowledge. Thus, orthographic knowledge may play a role in the letter choices that young children make when spelling." (p.632)

Estas autoras hablan de "orthographic knowledge" allí donde creemos más conveniente hablar de "graphic knowledge". Los niños pequeños pueden expresar en sus producciones y en sus elecciones interpretativas conocimiento de las combinaciones posibles o imposibles, sin tener aún un conocimiento de las alternancias en un punto dado de una cadena gráfica (condición limitativa de lo ortográfico, según Gak). En otras palabras, puede haber un conocimiento de las secuencias como tales (eje sintagmático) sin que las correspondencias grafemas-fonemas estén consolidadas y, sobre todo, sin que las alternativas en el eje paradigmático estén organizadas (o sea, los lugares de alternancia gráfica en una secuencia dada). Ese conocimiento gráfico seguramente es reorganizado a medida que progresa el conocimiento ortográfico, pero no hay razón para pensar que desaparece. Hay que tener en

cuenta que los requerimientos de cantidad mínima y variedad interna (Ferreiro y Teberosky, 1979; Ferreiro,) están sólidamente establecidos mucho antes de las correspondencias fonema-grafema, lo cual sugiere una atención precoz a las secuencias de letras como tales.

Por todas las razones antedichas, proponemos considerar tres niveles de organización:

1) EL SISTEMA GRAFICO.

Escrituras históricamente establecidas que comparten el mismo alfabeto latino difieren en la distribución de los caracteres. Las diferencias más ostensibles son las siguientes: longitud de las palabras gráficas, cantidad y tipo de diacríticos, distribución de mayúsculas y minúsculas, secuencias de letras en posiciones iniciales o finales, secuencias de una a tres letras que constituyen palabras gráficas de uso frecuente. Estas propiedades gráficas tienen cierta relación con propiedades fonológicas de la lengua en cuestión pero no están totalmente determinadas por ella. Por ejemplo, la conjunción <y> en español se podría graficar <i>²; lo que en español se grafica <ñ> corresponde a <nh> en portugués y a <gn> en italiano (por ejemplo: *gnocchi*, *ñoqui*, *nhoqui*).

Dentro del sistema gráfico es preciso distinguir:

1.a) EL REPERTORIO GRÁFICO, o sea, el conjunto de formas-letras con sus variantes mayúsculas, y quizás las variantes tipográficas; es preciso incluir en el repertorio gráfico, además, los diacríticos, los signos de puntuación, las marcas de fin de renglón que indican palabra incompleta, y quizás otros elementos gráficos que no acostumbran aparecer como parte del "repertorio gráfico", como los espacios en blanco. En todo caso, es preciso tener en cuenta desde el inicio que el repertorio gráfico no se agota en la pregunta: "¿cuántas letras tiene el alfabeto de la lengua X?".

1.b) LAS POSIBILIDADES COMBINATORIAS: combinaciones prohibidas, obligatorias o permisibles.

1.c) LAS RESTRICCIONES POSICIONALES A LAS POSIBILIDADES COMBINATORIAS, o sea, la toma en consideración de los espacios que delimitan palabras gráficas, y que determinan que una cierta combinación de grafías pueda ocurrir o resulte imposible inmediatamente después de un espacio o cuando es seguida por un espacio.

2) LAS REGLAS DE CORRESPONDENCIA ENTRE SECUENCIAS GRÁFICAS Y SECUENCIAS SONORAS. Gak ubica este nivel de análisis dentro de los principios gráficos, pero nos parece más conveniente considerarlo como un nivel independiente, que reorganiza lo gráfico con respecto a otro sistema (el de las unidades sonoras). Es probablemente este nuevo nivel de organización del material gráfico el que da acceso a lo ortográfico, en sentido estricto. Nótese que preferimos hablar de correspondencia entre secuencias y no entre elementos aislados; de hecho, preferimos analizar la correspondencia entre elementos como un caso límite de las secuencias: secuencias de un único elemento.

3) EL NIVEL ORTOGRAFICO que, provisoriamente, proponemos considerar, siguiendo a Gak, como un nivel de alternancias paradigmáticas, que rigen la alternancia de grafías potencialmente equivalentes, pero obligatorias "según las circunstancias".

En lo que sigue, nuestra atención estará focalizada en el primer nivel (las variantes 1.b y 1.c de organización del sistema gráfico), y por dicha razón no avanzamos aquí en la elaboración conceptual de los otros dos niveles.

-.--.-.-.-.-.-.-

ANALISIS DE LOS DATOS.

La mayor parte de los análisis que aquí presentamos han sido realizados sobre el corpus de español (412 textos, producidos por niños de 2º y 3er año de escuela primaria de tres países diferentes: México, Argentina y Uruguay). Sin embargo, hemos podido hacer algunos análisis comparativos con el corpus de portugués (275 textos de 2º y 3º, recogidos en Brasil) y con el corpus italiano (450 textos de 1º a 3º).³

Para tratar de mostrar el interés de tomar en cuenta el conocimiento del sistema gráfico tomaremos algunos casos que nos parecen privilegiados para la reflexión y que podemos explorar en nuestro corpus: letras imposibles en posición final; dígrafos; repetición de vocales; secuencias gráficas que corresponden a reglas ortográficas explícitas.⁴

1) Letras imposibles en posición final.⁵

En las tres lenguas consideradas, algunas consonantes son imposibles en posición final. Es, por ejemplo, el caso de <-c>.

El cuadro 1 muestra que casi el 100% de las ocurrencias de <-c> son onomatopeyas: *toc /tac /toc* en español; *toc /tuc /doc / nhac*, en portugués; exclusivamente *toc* en italiano. Estas onomatopeyas se constituyen generalmente en pares o tríos, con o sin separadores (comas, guiones o inclusive signos de exclamación o puntos suspensivos como en este caso italiano: *toc...toc*).

En todos los casos estas onomatopeyas corresponden al sonido que se hace al golpear la puerta (llegada del lobo o de Caperucita a la casa de la abuela), excepto la onomatopeya *nhac*, en portugués, utilizada 4 veces en 2 textos diferentes, para simular el ruido de las fauces del lobo devorando a la protagonista o a su abuela. Esta onomatopeya es particularmente interesante, porque viola al mismo tiempo restricciones gráficas iniciales y finales: un final <-c> y un inicio <nh->, ambos imposibles en la escritura de la lengua. Señalemos, además, que las únicas dos ocurrencias que, en portugués, no son onomatopeyas, corresponden a la abreviatura etc

(= etcétera). (Las abreviaturas, tanto como las onomatopeyas, pueden violar intencionalmente las reglas de combinación de letras de la lengua).

Otra consonante imposible en posición final, en italiano y español, es <-m>⁶. El cuadro 2 muestra los resultados, donde aparece nuevamente el predominio de las onomatopeyas.

Las onomatopeyas registradas son las siguientes: en español, *pum* /*pom* /*tam* ; en italiano, *pum* / *bum* / *bam*. Estas onomatopeyas corresponden a alguien que golpea la puerta, o a los tiros de los cazadores; también se organizan en pares o tríos, con o sin separadores. Pero una buena parte de estas onomatopeyas (la mayoría en italiano) son de otro tipo, e intentan representar el momento en que el lobo traga a la abuela o a Caperucita. En italiano registramos: *am* /*amm* / *aum* / *om* / *aaam* / *Hamm* / *aaamm* / *Haaamm* / *gnam* / *gniam* / *sgnm*; en español registramos: *amm* / *Amm* / *mmmm* (con cantidad variable de letras repetidas). Las dos únicas ocurrencias que no corresponden a onomatopeyas, en español, son producidas por un niño de e2b quien escribe dos veces *com* en lugar de la preposición *con*. En italiano, una de las ocurrencias es producida por un niño de i2b que escribe: *nom mi fermero*, pero la otra corresponde a un intento, seguramente intencional, de dar énfasis a la expresión crucial del lobo, dejando una <m> aislada: "*Per m ma mangiarti meglio!!*"

El uso casi exclusivo de un final <-c> o <-m> para las onomatopeyas plantea un problema interesante: las onomatopeyas no son palabras; por lo tanto, es posible pensar que sólo cuando se escribe algo que no es una palabra de la lengua se pueden violar las restricciones gráficas.

2) El caso de los dígrafos compuestos por dos letras diferentes.

2.1) *Los dígrafos con <h> en portugués.* Conviene empezar por el análisis del portugués porque tiene mayor cantidad de dígrafos. Hay tres dígrafos que se construyen con una consonante seguida de <h>: <ch>, <lh> y <nh>. Lo interesante es que estos tres dígrafos aparecen en el nombre de la protagonista de la historia: **Chapeuzinho Vermelho**. Por lo tanto, su frecuencia es muy alta en el corpus.

Lo primero que podemos preguntarnos es en cuántos casos observamos alguna permutación en el orden de estos dígrafos. La respuesta es contundente: las secuencias <hn> y <hl> no fueron **nunca** producidas. Sólo registramos seis permutaciones <hc> en un único texto de p2b donde se lo emplea seis veces en tres palabras: *hcapeuzion* (en lugar de *Chapeuzinho*), *hcega* (en lugar de *chegar*) y *hcamado* (en lugar de *chamado*).

2.2) *El dígrafo <ch> en español.* Este dígrafo puede considerarse sólidamente establecido ya que en 398 utilizaciones **nunca** apareció con permutación (o sea, ningún niño escribió <hc>) y apenas en una única ocurrencia fue sustituido por <sh>, inexistente en la ortografía del español: *escusho* (en lugar de *escuchó*). La categoría gramatical de las palabras en las cuales este dígrafo fue utilizado es muy variable: sustantivos (*leche*, *cuchillo*), verbos (*escuchar*, *aprovechar*), adjetivos (*chico*, *muchos*), expresiones fijas de despedida (*chau*).

2.3) *El dígrafo <ch> en italiano.* Este dígrafo es extremadamente frecuente: aparece utilizado 3520 veces en el corpus completo (de 1° a 3°), y **nunca** apareció con permutación (o sea, ningún niño escribió <hc>). Incluso cuando aparece utilizado en contextos desviantes, las letras fueron escritas en el orden obligatorio, como en esta escritura: *boscho* (= *bosco*). La categoría gramatical de las palabras en las cuales aparece este dígrafo es muy variable: sustantivos (*occhi*), verbos (*chiama*, *chiedi*), expresiones interrogativas (*perché*), expresiones coordinativas (*anche*), pronombres (*chi*), etc.

2.4) *La secuencia <qu> en español y portugués.* En la escritura de estas lenguas la letra <q> no puede aparecer nunca sola, razón por la cual la secuencia <qu> debe considerarse obligatoria y ser tratada como un dígrafo (con una <u> "muda"). En el total de la muestra de español, la frecuencia de esta secuencia gráfica es muy alta: en 3343 ocurrencias no se observó **ningún** caso de permutación de las letras de esta secuencia obligatoria. Es decir, ningún niño escribió <uq>. Al igual que en español, esta secuencia puede considerarse sólidamente establecida en portugués, sin que se hayan registrado permutaciones en un total de 1456 ocurrencias.

Las tablas 3 y 4 muestran los altísimos resultados correctos en el uso de <qu> en ambas lenguas (98% en español y 99% en portugués). Hay pues, poquísimas producciones desviantes (que no por ello pierden interés) pero esas desviaciones no consisten nunca en el uso de la combinación <uq>.

Ortográficamente, <qu> sólo puede ser seguida por dos vocales: la secuencia <que>, que corresponde a /ke/, y la secuencia <qui> que corresponde a /ki/. Veremos por separado estos dos casos.

2.4.1) La secuencia <que>, en español.

Esta secuencia puede constituir palabra autónoma y funcionar, según los contextos, como pronombre relativo, conjunción, expresión interrogativa o exclamativa, y asociarse con otras palabras para constituir expresiones complejas, conservando a veces su autonomía gráfica (por ej. *para que*, *por que*). En el total del corpus español hubo 2135 ocurrencias de <que> como palabra autónoma, y 99% fueron escritas correctamente.

La misma secuencia puede constituir el inicio de una palabra gráfica (es el caso de diversas formas conjugadas de los verbos *querer* y *quedar*, frecuentes en este corpus), puede encontrarse al final de una palabra (*bosque* es una palabra de muy alta

frecuencia) o bien ubicarse dentro de la palabra (por ej, *pequeña*, y otras de baja frecuencia).

La gran mayoría de las utilizaciones de la secuencia <que> son correctas en todos los grupos, inclusive en los más pequeños de grupos desfavorecidos (2b).

2.4.2) La secuencia <qui>, en español.

Esta secuencia no puede constituir palabra gráfica autónoma. La encontramos en posición inicial en palabras frecuentes como el pronombre interrogativo *quién* y formas del verbo *querer* (*quiere*, *quiso*) o del verbo *quitar*. Es final de palabra en el pronombre de lugar *aquí*, y está en el interior de palabras como *izquierda*, *mantequilla*, y formas conjugadas del verbo *equivocar(se)*.

Los porcentajes de graficaciones correctas de la secuencia <qui>, sobre la muestra total (92.5%) son inferiores a los de la secuencia <que> (98%), pero de todos modos suficientemente altos como para afirmar que, desde el inicio, esta secuencia gráfica está firmemente establecida.

2.4.3) Las producciones desviantes

Varias de las producciones desviantes consisten en graficar el exclamativo *qué* como *ce*. Hay en total 13 producciones de este tipo, pero corresponden apenas a 4 textos ya que tres niños acumulan varias ocurrencias en el diálogo canónico (*ce ojos* / *ce orejas* / *ce boca*). Solo en otras cuatro ocurrencias aparece algo similar en otras palabras: *bosce* (= *bosque*), *porce* (= *porque*), *ceria* (= *quería*), *acel* (= *aquel*).⁷

La utilización aislada de <q>, sin su "vocal muda" obligatoria, se registró apenas 11 veces, en ocurrencias aisladas en todos los grupos de la muestra, principalmente como la secuencia gráfica aislada *ge* (= *que*).

En otros 6 casos aparece lo que a primera vista podría considerarse como una omisión de la vocal: *porqu*, *bosqu*, *qu*, (en lugar de *porque*, *bosque*, *que*). Sin embargo, la omisión aparente podría quizás corresponder a otra interpretación, según la cual la secuencia <qu> correspondería a /ke/. En otras palabras, un contexto en el cual la vocal <u> cambia de valor sonoro, en lugar de perderlo.

El mantenimiento del valor sonoro de la vocal <u> determina la aparición de 3 producciones desviantes: *quidado* (= cuidado), *quida* (= cuida) y *es qucharte* (= escucharte, pero dejando aislado el segmento inicial *es*, que corresponde a una forma del verbo *ser*). Finalmente, puede ocurrir que la secuencia <qu> aparezca con su valor /k/, pero sin respetar las restricciones ortográficas de las vocales que pueden seguirla. Esto dió lugar a 13 ocurrencias, todas con <qua> o <quo> (ya que ningún niño intentó <quu>, lo que constituiría una violación grave a la distribución gráfica de las vocales en español). Así observamos casos como los siguientes: *quaperusita* (= caperucita), *boqua* (= boca), *toquo* (= tocó), *saquo* (= sacó).

En el caso de la secuencia <qui-> las producciones desviantes se concentran en el pronombre interrogativo *quién*, que presenta la dificultad de un diptongo en el que intervienen precisamente las únicas dos vocales que pueden seguir al dígrafo <qu>. En efecto, de las 20 graficaciones desviantes para este pronombre interrogativo, apenas 7 (repartidas en todos los grupos) corresponden a la solución alternativa *cien* (una solución no-ortográfica, pero coherente con las correspondencias grafo-fónicas). Todas las otras producciones desviantes involucran el uso del dígrafo <qu>, pero no solucionan el diptongo (cuyo análisis sonoro se ve complicado por la nasalización de la última vocal). Las soluciones observadas son las siguientes (con su frecuencia entre paréntesis): *quen* (4); *quín* (2) *quie* (3), *quí* (2), *queien* (1), *quaiien* (1), *cien* (7).

Las respuestas desviantes en portugués son enteramente similares a las del español:

- utilización de la alternativa fono-gráfica: *ci, ce, cem* (en lugar de *qui, que, quem*);
- desaparición del dígrafo con preservación de <q>: *qe* (= *que*);
- utilización de <qu> con valor /ke/.

La única respuesta desviante que no tiene equivalente en español es una del grupo p2b que utiliza la alternativa fono-gráfica <k>: *ke* (= *que*).

2.5) El dígrafo <sc> en italiano.

El italiano, la secuencia <sc> funciona como dígrafo sólo cuando es seguida por las vocales <i> o <e>: *esce, uscire, conosce, capisce, riesce, finisce, ubbidisce, scendere, lasciare, scegliere* (todas ellas palabras registradas en el corpus).

Nos preguntamos si habría ocurrencias de esta secuencia permutada (o sea, <cs>. La respuesta es clara: sólo registramos tres ocurrencias de dicha secuencia, en tres textos diferentes: *uccícse* (= *uccise*), en un texto de 2º; *csi* (= *cosí*) y *ucsci* (= *usci*) en dos textos diferentes de 3º.

En conclusión, el análisis de dígrafos compuestos por letras diferentes nos ha mostrado que, en las tres lenguas, las permutaciones son prácticamente inexistentes (totalmente inexistentes en la mayoría de los casos), un hecho pleno de consecuencias, que comentaremos en las conclusiones.

3) Las vocales repetidas en posición contigua.

Tanto en italiano como en español, hay severas restricciones a la repetición de la misma vocal en posición contigua. En español, las únicas vocales que pueden repetirse son <e> y <o>⁸ En italiano, la

duplicación de vocales es mucho menos frecuente que la duplicación de consonantes. Sólo pueden aparecer, y en raros casos, <ee>, <ii> y <oo>.⁹

Veamos si los niños respetan estas restricciones gráficas. La tabla 5 muestra, en apariencia, que no es así, ya que la vocal más repetida, en ambas lenguas, es <aa>, precisamente la que no puede repetirse. De todas las vocales repetidas, <aa> representa aproximadamente la mitad (el 66% en español y el 48% en italiano). Sin embargo, el análisis de las expresiones donde se utiliza esta vocal repetida nos muestra que se trata en muchos casos de onomatopeyas (en español, 15/25 = 60%; en italiano, 7/15 = 47%). Al menos un ejemplo de onomatopeya aparece con cada una de las vocales en español: *aay* / *eeeeee* / *riiiin* / *ooo* / *Buuuu*. (Como puede apreciarse, pueden aparecer hasta seis vocales repetidas en posición contigua). En italiano ocurre lo mismo, excepto para la vocal <e>, que no aparece repetida: *Aaamm* / *ii* / *Hoooo* / *uuu*.

Por lo tanto, volvemos a encontrar aquí el fenómeno antes señalado: pueden violarse las restricciones gráficas cuando se escriben expresiones que NO son palabras de la lengua. Sin embargo, estamos lejos de los porcentajes de los cuadros 1 y 2, donde prácticamente el 100% de las expresiones con consonantes imposibles en posición final eran onomatopeyas. ¿En qué otros casos se utilizan vocales repetidas?

En los textos de español un lugar privilegiado para ver aparecer vocales repetidas son los vocativos. Veamos dos ejemplos:

- Un texto de e2b inicia con un vocativo producido por la madre, quien desde lejos llama a Caperucita así: *Kaaaperusita*. En el mismo texto, los turnos de Caperucita durante el diálogo canónico también inician con vocativos, escritos así: *aaabuelita* / *aabuelita*. Tenemos derecho a considerar que estas repeticiones de vocales cumplen una función expresiva, porque en el mismo texto el nombre de la protagonista y de su abuela aparecen escritos sin repeticiones vocálicas, cuando no se los usa como vocativos.

- Otro texto de e2b utiliza un vocativo cuando Caperucita llega a casa de su abuela y la llama así: *ABUELITAA* (escrito con doble A final y todo con mayúsculas). El efecto expresivo se obtiene por el doble uso de mayúsculas y repetición de la vocal final, recursos gráficos que permiten aludir a la vez al aumento de volumen de voz y al alargamiento de la emisión. (Nótese que varias onomatopeyas también utilizan mayúsculas).

En los textos de italiano registramos apenas 3 casos de repetición vocálica en los que podemos atribuir, con fundamento, una función expresiva:

- Un texto de 3er año presenta un *sí* enfático (*sí entra*) escrito: *siiii*.
- Otro texto del mismo nivel escolar presenta una doble expresión de despedida, la segunda de las cuales con repetición vocálica: *ciao ciaooo*.

En total, las repeticiones que podemos atribuir, con base al análisis de cada texto completo, a una intención expresiva (énfasis, sorpresa, llamado desde un lugar distante, aumento del volumen o prolongación del sonido emitido) constituyen 8 ocurrencias en el corpus de español y apenas 3 en el italiano. A pesar de su escaso número, nos parece útil poner de relieve que algunos niños se atreven a apartarse de las normas gráficas propias a la escritura de su lengua en momentos específicos, buscando efectos particulares similares a los de las onomatopeyas. Así como las onomatopeyas intentan representar lo que la escritura no pretende representar (los ruidos, por oposición a las palabras), las repeticiones indebidas dentro de palabras de la lengua que aquí estamos considerando intentan representar énfasis, alargamientos, y otras características de la emisión sonora que van más allá de las diferencias entre fonemas, único foco de atención de las escrituras alfabéticas. De hecho, se trata de recursos icónicos que han sido ampliamente explotados por "les bandes dessinéés" (¿i fumentí??) y que los niños han aprendido a ubicar en los límites del discurso directo.

En italiano, algunas de las repeticiones vocálicas parecen obedecer a razones puramente gráficas. Por ejemplo, la escritura *straaa*, en lugar de *strada*, o la escritura *aal*, en lugar de *dal*, podrían explicarse por razones puramente gráficas vinculadas al trazado de las letras en cursiva: una <a> cursiva sólo difiere de una <d> cursiva en la altura del trazo vertical. (Ambos ejemplos aparecen en textos de primer año, o sea, textos producidos por ejecutantes inexpertos).

También en italiano, algunas de las repeticiones vocálicas parecen obedecer a una desorientación acerca de cuáles y cuántas *doppie* deben aparecer en cada palabra. Hemos registrado <aa> y <ee> en el nombre del personaje principal (que contiene tres *doppie*) en textos de niños de 3er año, quizás particularmente atentos a las *doppie*: *caapuccietto* / *cappeecietto* / *cappuceetto*. Como puede observarse, las tres escrituras desviantes presentan tres *doppie* (aunque no exactamente las tres esperadas).

Finalmente, en ambas lenguas aparecen algunas duplicaciones vocálicas cuya motivación se nos escapa, y que quizás hubieran sido corregidas en caso de proceder a una revisión posterior. Señalemos únicamente una repetición intrigante, registrada una única vez en español y una única vez en italiano: *pieedi* (italiano) y *piees* (español).

4) Secuencias graficas (y ortograficas) que no corresponden a un fonema de la lengua.

Hasta ahora hemos analizado casos de dígrafos que constituyen a la vez secuencias gráficas obligatorias y "casos particulares" de las correspondencias entre letras y grafemas (ya que, por definición, un dígrafo es una secuencia de dos letras que corresponde a un único fonema).

Es útil confrontar estos casos con otros que son, en apariencia, puramente ortográficos pero que, dado su carácter obligatorio (sin excepciones) podrían analizarse como fenómenos gráficos. Contrariamente a los casos anteriormente comentados, que se referían (de preferencia) a combinaciones iniciales o finales de una secuencia gráfica separada por espacios en blanco, el caso que vamos a comentar concierne a secuencias gráficas que ocurren *dentro* de una cadena gráfica.

Una de las reglas ortográficas del español que las instrucciones escolares repiten sin cesar (y una de las pocas que los usuarios adultos pueden recitar sin reflexión) es la siguiente: "se escribe eme antes de be y de pe". La frontera silábica de estas palabras se ubica obligatoriamente entre esas dos letras (por ej. *cam-po*, *tam-bién*).

La tabla 6 muestra, del total de palabras que debían escribirse con <-mp->, cuántas resultaron escritas correctamente y cuántas fueron escritas con la secuencia <-np-> (ortográficamente desviante).

Los datos son muy claros, sobre todo si tenemos en cuenta que 21 de las 35 ocurrencias de graficación desviante (60%) son palabras que inician con la secuencia <en>, la cual constituye una secuencia gráfica autónoma (preposición). Podría tratarse en este caso de un conflicto entre dos secuencias gráficas: <en> y <-mp->. En el orden secuencial de la escritura <n> se impondría sobre <m> porque sucede a <e>. Del total de 162 palabras escritas correctamente apenas 50 (31%) comienzan con <em->, lo cual refuerza nuestra hipótesis.

En el corpus italiano encontramos algo similar: en total hay apenas 6 palabras escritas con <-np-> en lugar de la secuencia normativa <-mp->, de las cuales 4 comienzan con la secuencia <in>, la cual constituye, por sí misma, una secuencia gráfica autónoma

(preposición): *inproviso, inprovisamente, inpauritta, inproblia* (= imbroglia). (Las restantes ocurrencias son: *senpre, zanpe*.)

En el corpus de español aparecen muchas menos palabras que debieran graficarse con <-mb->. En este caso existe, además, una alternativa ortográfica (<v> en lugar de) que no corresponde a ninguna diferencia sonora. El análisis debe entonces tomar en cuenta cuatro alternativas: <-mb-> y <-nv-> (ambas correctas); <-nb-> y <-mv-> (ambas incorrectas). El escaso número de ocurrencias hace riesgosa la comparación. A reserva de otros datos confirmatorios y a título puramente tentativo observamos las siguientes oposiciones:

TOTAL <-mb-> (correcto) = 71 (92%) vs TOTAL <-nb-> (incorrecto) = 6

TOTAL <-nv-> (correcto) = 15 (75%) vs TOTAL <-mv-> (incorrecto) = 5

En el corpus de italiano encontramos un número mayor de secuencias (incorrectas) <nb>, debido quizás a la alta frecuencia de la palabra *bambina*, escrita 14 veces como *banbina*. Con respecto a la secuencia imposible <mv>, se registra una única ocurrencia, producto de una omisión (probablemente involuntaria): *chiamva* (= chiamava).

CONCLUSIONES.

Hemos podido mostrar que durante la ejecución de una tarea de redacción que no focalizaba la atención de los niños hacia los aspectos ortográficos, la producción escrita constituye una valiosa fuente de datos acerca del conocimiento (seguramente inconciente) de las restricciones gráficas vinculadas con secuencias gráficas imposibles y secuencias gráficas obligatorias. Incluso los niños de medios menos favorecidos (i1b, i2b e i3b) o francamente desfavorecidos (p2b y p3b; e2b y e3b) son sensibles a las restricciones gráficas, como puede verse en todas las tablas.

Es importante recordar que ninguno de los textos incorporados al corpus fue aceptado o rechazado por razones ortográficas. Los criterios para no incorporarlos fueron de naturaleza narrativa (ausencia de alguno(s) de los personajes principales, ausencia de episodios claves de la historia o mezcla con otras historias tradicionales) o bien de naturaleza gráfica, pero relativos al tipo de trazado (ausencia casi completa de límites de palabras gráficas, por ejemplo). Por lo tanto, y dadas las dimensiones del corpus, la ausencia total de ciertos fenómenos nos parece un dato crucial.

Ciertas consonantes imposibles en posición final, tales como <-c> y <-m> se reservan para las onomatopeyas. Dado que las onomatopeyas no son palabras de la lengua, los niños parecieran expresar, a través de estas elecciones gráficas, un principio fundamental: sólo cuando se escribe algo que no es una palabra se pueden violar las restricciones gráficas. El análisis de las vocales repetidas parece evidenciar un principio similar, con matices vinculados con la posibilidad de expresar, en el discurso directo, propiedades de la emisión oral a las cuales la escritura alfabética no presta atención: alargamientos, aumento de volumen, énfasis, etc.

El análisis de los dígrafos compuestos por dos letras diferentes ha sido sumamente instructivo. La ausencia total de permutaciones en el orden de estos dígrafos nos parece sugerir que ellos no han sido aprendidos como adición de letras. Si los niños comenzaran por las correspondencias grafo-fónicas que definen al sistema alfabético, letra por letra, los dígrafos serían particularmente difíciles de asimilar, precisamente porque violan los principios de correspondencia bi-unívoca entre letras y fonemas. Nuestros datos sugieren que los dígrafos no son tratados como excepciones a una regla general, sino como casos fáciles de asimilar (quizás porque no son excepciones a una regla de los niños, sino excepciones a reglas formuladas por analistas del lenguaje y por pedagogos de la lengua). Es probable que, desde el punto de vista

de los niños, las primeras correspondencias no se establezcan entre elementos aislados sino entre secuencias gráficas y secuencias sonoras y que, por esa razón, los dígrafos no aparezcan como excepciones sino como una de las tantas secuencias que tratan de comprender. La atención precoz que los niños prestan a las secuencias (en períodos previos a las correspondencias silábicas), y que los llevan a elaborar principios organizativos muy poderosos como el de cantidad mínima y el de variación interna, nos hace pensar que el trabajo sobre secuencias está ya sólidamente instalado cuando comienza el trabajo sobre las correspondencias grafo-fónicas.

La confrontación de los datos anteriores con algunas secuencias gráficas obligatorias -habitualmente descriptas como reglas ortográficas- tales como <-mp-> y <-mb-> nos lleva a pensar que quizás los niños puedan aprender como restricciones gráficas lo que generalmente se enseña como reglas ortográficas.

En ninguno de los países donde fueron recogidos nuestros textos (Brasil, Argentina, México, Uruguay e Italia) los manuales escolares hacen mención de combinaciones posibles o imposibles de letras. Y quizás es bueno que así sea. El ambiente pedagógico tiene una tendencia persistente: cuando se reconoce que algo es psicológicamente importante, se lo transforma de inmediato en un contenido que debe ser enseñado. Creo que estamos en un terreno que es proclive a tales deslizamientos, ya que la ortografía ha sido tradicionalmente concebida como "un contenido a ser explícitamente enseñado". Si fuera cierto, como estos datos lo sugieren, que hay un conocimiento inconciente de la organización del material gráfico, en términos de sus combinaciones obligatorias y/o imposibles, quizás se trate de un conocimiento que hay que dejar que se constituya precisamente de manera inconciente, sin tratar de elevarlo precozmente al rango de las "reglas" (más a menudo recitadas que obedecidas). Es muy posible que se trate de un conocimiento que es eficiente, precisamente porque es inconciente. En todo caso, mejor tratemos de seguir

comprendiendo antes de pasar apresuradamente a las "consecuencias pedagógicas" de los resultados de la investigación.

Sin embargo, algo de lo dicho antes tiene inmediatas consecuencias pedagógicas, pero no para fabricar lecciones sino para cambiar la mirada de los docentes: en textos plagados de desviaciones, irregularidades y errores ortográficos, hemos podido ver "lo que de veras tienen" en lugar de ver "lo que no tienen". En otras palabras, hemos podido iniciar una descripción, en términos positivos, de aspectos clásicamente considerados como ortográficos de estos escritores debutantes. Y estamos sugiriendo que quizás los niños aprendan como restricciones gráficas las combinaciones obligatorias que la escuela enseña habitualmente como reglas ortográficas.

¹ "muchas cuestiones habitualmente tratadas como fono-gráficas están, de hecho, relacionadas exclusivamente con la descripción del material gráfico".

² De hecho, así se graficó la conjunción hasta fecha reciente. La oscilación entre <i> / <y> es uno de los problemas más interesantes en la historia de la ortografía del español.

³ Agradecemos a Nadja Ribeiro Moreira los análisis proporcionados sobre el corpus portugués. Agradecemos a Daniela Fabbretti su colaboración para obtener algunos datos comparativos sobre el corpus italiano.

⁴ La lengua de origen, el grado y grupo de procedencia de cada texto serán indicados de la siguiente manera: i2b = italiano, 2º, clase baja; e3m = español, 3º, clase media; p2b = portugués, 2º, clase baja; etc. Daremos particular relieve en nuestro análisis a los textos de 2º y 3er año producidos por niños de clase baja, ya que estos textos tienen el máximo de problemas ortográficos y se trata de niños con pocas ocasiones de interactuar con material impreso (libros) fuera de la escuela, particularmente en el corpus de portugués y español.

⁵ En lo que sigue utilizaremos el modo de notación que es usado por Sampson (1985:90): un guión corto que precede a una letra indica letra final; un guión corto que sigue a la letra indica letra inicial; un guión a ambos lados indica letra o secuencia intermedia.

⁶ Rarísimas excepciones, como *album*. En portugués el final (casi)imposible no es <-m> sino <-n>, pero su situación no es similar a la de las otras lenguas porque el portugués tiene un número considerable de palabras terminadas en <-m> cuyo plural es <-ns>, por ejemplo: *homem* / *homens*; *bom* / *bons*; *um* / *uns*. Si tenemos en cuenta que el caso más general para la formación del plural en portugués es el agregado del morfema <-s>, sin modificación de la consonante precedente, es fácil imaginar que los

finales <-m> con plural <-ns> constituyen una zona de incertidumbre para los niños en proceso de alfabetización.

⁷ El caso más claro de duda ortográfica aparece en un texto de e3b donde la palabra que aparece graficada *cque*, sin borrar la letra inicial. Es evidente la oscilación de ese niño entre dos variantes de las correspondencias grafo-fónicas: <c>, con valor /k/ delante de todas las vocales excepto <e,i>, y <qu>, también con valor /k/ pero sólo en el caso de esas dos vocales.

⁸ Aparece <ee> en algunos verbos como *leer*, *creer* y en palabras precedidas del prefijo <pre-> e iniciadas con <e>, tales como *preescolar*; sin embargo, en el uso actual del español escrito ésta es una zona de incertidumbre ortográfica ya que la doble <ee> tiende a ser reducida a una sola, *prescolar*, o bien puede introducirse un guión para indicar la naturaleza compuesta de la palabra: *pre-escolar*; todas esas formas de graficar son aceptables, en el presente estado de la ortografía del español. La secuencia <oo> aparece en una palabra frecuente para los niños: *zoológico*, y en palabras con el prefijo <co-> que inician con la misma vocal, por ejemplo *cooperar*.

⁹ La secuencia <oo> aparece en palabras derivadas del griego: *zoo*, *zoologo*, y en el léxico científico. La secuencia <ee> corresponde al plural de algunos nombres y adjetivos femeninos tales como *idee*, *aeree*. La secuencia <ii> corresponde al plural de algunos pocos nombres masculinos: *zii*. (Ferreiro, Pontecorvo & Zuccheromaglio, 1987:25)

TABLA 1
Occurrencias de <-c> en textos de Español, Portugués e Italiano

	# Textos	# Occurrencias	Onomatopeyas
e2b + e3b (N=207)	07	17	16 (94%)
e2m + e3m (N=205)	28	77	76 (99%)
p2b + p3b (N=119)	08	18	18 (100%)
p2m + p3m (N=156)	19	59	57 (97%)
i2b + i3b (N= 152)	21	57	57 (100%)
i2m + i3m (N= 148)	19	47	46 (98%)

TABLA 2
Occurrencias de <-m> en textos de Español e Italiano

	#Textos	# Occurrencias	Onomatopeyas
e2b (N=103)	05	08	06 *
e3b (N=104)	06	14	14 (100%)
e2m (N=102)	02	03	02 *
e3m (N=103)	01	02	02 (100%)
TOTAL (N=412)	14	27	24 (89%)

	#Textos	# Occurrencias	Onomatopeyas
i2b (N= 78)	10	19	18 (95%)
i3b (N= 74)	09	15	15 (100%)
i2m (N= 71)	09	09	09 (100%)
i3m (N= 77)	03	04	03 *
TOTAL (N=300)	31	47	45 (96%)

* = no se calculan % por ocurrencias muy bajas.

TABLA 3
Occurrencias de <qu> en textos de Español

	TOTAL	CORRECTAS	DESVIANTES
e2b	814	777 (95.5%)	37 (4.5%)
e3b	924	905 (98%)	19 (2%)
e2m	668	654 (98%)	14 (2%)
e3m	937	927 (99%)	10 (1%)
TOTAL	3343	3263 (98%)	80 (2%)

TABLA 4

Occurrencias de <qu> en textos de Portugués

	TOTAL	CORRECTAS	DESVIANTES
p2b	230	223 (97%)	7 (3%)
p3b	411	409 (99.5%)	2 (0.5%)
p2m	300	300 (100%)	-
p3m	515	512 (99%)	3 (1%)
TOTAL	1456	1444 (99%)	12 (1%)

TABLA 5
Occurrencias de vocales duplicadas en textos de Español e Italiano.

	Español (N=412)	Italiano (N=450)
<aa>	25	15
<ee>	06	06
<ii>	02	05
<oo>	02	02
<uu>	03	03
TOTALES	38	31

TABLA 6
Occurrencias de <-mp-> y <-np-> en textos de Español

	<-mp-> (correcto)	<-np-> (incorrecto)
e2b (N=103)	32 (73%)	12 (27%)
e3b (N=104)	46 (75%)	15 (25%)
e2m (N=102)	36 (86%)	06 (14%)
e3m (N=103)	48 (96%)	02 (4%)
TOTAL (N= 412)	162 (82%)	35 (18%)