

EFFE-Escreves como falas – falas como escreves?

Lourenço-Gomes, MC, C. Rodrigues & I. Alves (draft accepted to be published) Revue Romane, John Benjamins.

Abstract: This paper introduces the project EFFE that applies and develops a methodology for data collection primarily intended for the study of the correlation between oral and written production in European Portuguese, and that also aims to serve a broad range of studies in different areas. It comprehends (i) a brief review of two scientific contexts occupied with learning to read and write, which are at the basis of the methodological planning of future *corpora*; (ii) the methodology adopted for the collection of speech and writing material of 48 children of a private school in Lisbon, and a summary of the quantitative data already observed; (iii) some considerations on unconventional forms (UF) observed in written production that reflect the speech of those children.

Keywords: spelling development; alphabetic spelling; syllable complexity; written and spoken language

1. Introdução

Uma professora de alfabetização pediu a seus alunos que escrevessem uma frase com a palavra *formiga*, um dos temas da aula. Uma criança escreveu *A formiga é* . A professora, sem entender, disse que se ela não soubesse escrever a palavra, poderia desenhá-la. Surpresa com a dificuldade de compreensão da professora, a criança explicou que a frase dizia: "a formiga é invisível". (Lourenço-Gomes, 1999)

Todas as crianças, mesmo ainda muito novas, demonstram um grande interesse pela leitura e a escrita. Querem saber o que está escrito aqui e ali e tentam escrever tudo, mesmo o que ainda não sabem. Aquelas que dizem não gostar de ler ou escrever, em geral, passaram ou passam por uma experiência de insucesso no decorrer do processo de aprendizagem. Elas parecem saber desde muito cedo que essas atividades são parte de suas vidas e demonstram uma vontade intrínseca por seu aprendizado (Lourenço-Gomes, 1999). Suas primeiras produções revelam conhecimentos sobre o material impresso, sobre a sua língua e sobre as relações entre fala e escrita. O interesse dos investigadores acerca de como esses conhecimentos se desenvolvem tem feito do desempenho em leitura e em escrita tópicos de investigação produtivos em diferentes áreas.

Para além da imaginação e criatividade das crianças, refletidas na *história da formiga*, de uma delas, acima, a leitura e a escrita são estudadas sob diferentes ângulos. Esses dois sistemas funcionais da linguagem abarcam processos neuropsicológicos e linguísticos altamente complexos, alguns que são próprios a cada um dos sistemas e outros que demonstram uma estreita relação entre eles. Um desses ângulos refere-se aos erros observados na escrita inicial das crianças, sendo *erros* aqui interpretados como formas gráficas que não atendem às convenções estabelecidas para a língua mas que, em vez disso, refletem um aprendizado em desenvolvimento. Deste ponto em diante, serão chamados *formas não-convencionais* (FN-C).

A literatura dedicada à classificação e análise de erros ortográficos nos anos iniciais do ensino formal concorda, embora com diferentes designações e interpretações, que entre os tipos de FN-C mais frequentes nesta fase estão aqueles que refletem relações entre fala e escrita (Cagliari, 1984; Pinto, 1997, entre muitos outros). Algumas formas gráficas são facilmente associadas a formas da língua falada, quer na sua face fonológica, quer fonética. Este processo decorre da necessidade de fazer corresponder unidade fonológica e grafema para representar as palavras de grafia não estabilizada e mantém-se até que um léxico ortográfico se torne autónomo (Vellutino et al., 2004; Berninger, Abbott & Jones, 2006). Isto não é espontâneo apenas no início da aprendizagem, mas é também um recurso natural usado pelo adulto letrado quando este é solicitado a representar uma forma ortográfica desconhecida, pouco frequente ou uma não-palavra (Morais, 1995). Cagliari (1994), por exemplo, mostra algumas coincidências entre formas encontradas em textos arcaicos e aquelas frequentemente vistas nos textos das crianças, em português brasileiro, e que também podem revelar *intenções de escrita ou hipóteses sobre formas ortográficas*, tais como *coraçõ, pegarom, fizeru, nam, pam e coelio (coração, pegaram, fizeram, não, pão, coelho)*. Ou ainda, *tudos e aquiles (todos, aqueles)* em que <u> e <i>, em contraste com <o> e <e> poderiam, alternativamente, indicar uma pronúncia diferente da época ou um caso de hipercorreção ortográfica, isto é, uma generalização do fato de muitas palavras serem grafadas com <o> e <e>, mas pronunciadas como [u] e [i], conduzindo à confusão no uso dessas letras.

As propostas tipológicas descritas na literatura para as FN-C têm contribuído para um entendimento de aspetos importantes sobre o percurso de desenvolvimento da aprendizagem da ortografia (Pinto, 1997; Sim-Sim, Duarte & Ferraz, 1997; Mateus, 2002; Horta, 2001; Horta & Santos, 2013; cf., p. ex., Guimarães, 2005 e Moojen, 2009 para uma revisão de tipologias em português brasileiro). Embora essas propostas resultem de estudos oriundos de diferentes campos disciplinares, com objetivos e critérios metodológicos distintos, é ponto comum que as FN-C relacionadas com a fala diminuem em frequência à medida que a aprendizagem evolui. Esta observação sugere, à primeira vista, que o grau de domínio ortográfico pode ser, em parte, aferido a partir da frequência de ocorrência dessas formas, o que possibilitaria, em fases precoces, a implementação de estratégias pedagógicas para otimizar este domínio e de estratégias de intervenção para minimizar eventuais dificuldades das crianças.

De um ponto de vista mais teórico, a descrição do desempenho ortográfico em sua correlação com a fala fornece elementos de estudo para análise linguística e experimentação psicolinguística. Veloso (2010), por exemplo, sublinha que as FN-C relacionadas com a fala "(...) constituem pistas *reveladoras* do conhecimento fonológico intuitivo dos sujeitos que as produzem, devendo, por isso, ser merecedoras de atenção científica por parte dos estudos linguísticos e psicolinguísticos". Indo além do conhecimento fonológico, Sim-Sim, Duarte & Ferraz (1997) notam que o conhecimento meramente implícito da língua materna se converte em conhecimento explícito à medida que as crianças se desenvolvem e que as primeiras manifestações disso se revelam em autocorreções que elas fazem, principalmente a nível da articulação, da especificação lexical e de concordâncias sintáticas.

Deste modo, além de *merecedoras de atenção científica por parte dos estudos linguísticos e psicolinguísticos*, as questões examinadas, e ainda por examinar, são manifestamente importantes em termos de aplicação no ensino e na clínica.

No entanto, o domínio ortográfico e as relações que se estabelecem entre fala e ortografia nos anos iniciais da escola ainda carecem, por um lado, de uma observação mais ampla e, por outro, as próprias FN-Cs carecem de uma definição mais assertiva. Os processos cognitivos subjacentes a estas competências são demasiado complexos e os sistemas funcionais da linguagem (expressão escrita, expressão oral, compreensão da fala e compreensão da leitura), apesar de interagirem, são neuropsicologicamente distintos e têm trajetórias de desenvolvimento diferentes (Berninger, Abbott & Jones, 2006).

Com o principal propósito de ilustrar esta complexidade, na secção 2 fazemos uma apresentação sumária de alguns estudos sobre o desempenho de leitura e escrita no domínio das ciências cognitivas. Muito embora não estejam diretamente relacionados aos dados que serão

apresentados neste artigo, esses estudos sinalizam intenções de alargamento desses dados para investigação futura. Na secção 3, são enumerados os principais resultados de alguns estudos da área do ensino. Eles reforçam a convicção de que o conhecimento detalhado das estratégias utilizadas pelas crianças na sua fase de desenvolvimento da escrita é fundamental para traçar estratégias pedagógicas adequadas e de intervenção precoce que reduzam o insucesso escolar. Na secção 4, é apresentada a metodologia de recolha e tratamento adotada no âmbito do EFFE, até ao momento. Na secção 5 é feita uma avaliação global do projeto e, finalmente, na secção 6 são extraídas algumas considerações finais, com o propósito de orientar o trabalho futuro.

2. Leitura e escrita no cenário das ciências cognitivas

Desempenho em leitura

Muitos avanços foram feitos nas últimas quatro décadas acerca do conhecimento sobre o desempenho em leitura, marcadamente no que diz respeito à mudança de visão que considerava a leitura como uma capacidade essencialmente visoperceptiva (Maslow et al., 1964; Barret, 1965) para uma concepção de que ela é absolutamente linguística, mantendo relações com outras funções cognitivas¹. Esses avanços foram marcados por um esforço multidisciplinar, teórico e experimental, em especial da Psicologia Cognitiva, da Psicolinguística e da Neurociência Cognitiva no sentido de se perceberem as bases psicológicas e biológicas da leitura, no leitor iniciante e no leitor avançado, em termos de desenvolvimento normal e atípico (cf., p. ex., Morais, 1995; Dehaene, 2012)². Entre os tópicos de investigação nesta área, uma atenção especial foi dada aos processos relacionados ao reconhecimento de palavras, nomeadamente os processos de codificação visual e linguística. Diferentes modelos foram propostos mas, em geral, assume-se neles que no leitor hábil uma forma ortográfica ativa diretamente a representação fonológica correspondente e também o seu significado (Vellutino et al., 2004). Esta competência começa a desenvolver-se nos primeiros meses de escola, ou mesmo de instrução não formal da leitura, quando a criança é levada a prestar atenção às unidades gráficas (letras e sequências de letras, como os dígrafos em português), a aprender a fazer a correspondência com os sons de sua língua e a reuni-los para formar as palavras. Mas a consciência da correspondência grafema-fonema não é automática, precisa antes ser explicitamente aprendida (Morais, 1995).

Aprender a ler implica, porém, a aquisição de diferentes tipos de conhecimentos, os quais dependem, por sua vez, do desenvolvimento normal de competências cognitivas linguísticas e não-linguísticas que, em conjunto, irão permitir ao leitor estabelecer associações e contrastes entre as palavras faladas e escritas. De maneira mais direta, o leitor principiante exposto a um sistema de escrita alfabético precisa compreender as correspondências que se estabelecem entre letra e som (princípio alfabético) de maneira tal que lhe permita desempenhar com proficiência análises e manipulações das estruturas da língua em um nível sublexical. A compreensão conceitual e a consciência explícita de que as palavras são compostas por unidades segmentais (fonemas) e por combinações dessas unidades (sílabas) – consciência fonológica – e a compreensão e manipulação da organização das letras nas palavras – consciência ortográfica – são capacidades que se relacionam reciprocamente e facilitam a identificação das regularidades e redundâncias do sistema de escrita (cf. Wagner, Piasta & Torgesen, 2006; Vellutino et al., 2004; Rayner et al., 2012).

É verdade que a abertura de novas perspectivas para o conhecimento que temos hoje sobre como aprendemos a ler foi beneficiada por progressos paralelos de métodos de mapeamento cerebral, com ótima resolução temporal (EEG-ERP, MEG)³ e espacial (fMRI, PET-*scan*, TMS)⁴. Isto

¹ Para uma visão mais alargada sobre a aprendizagem da leitura e de seus transtornos, considerando a mudança da abordagem tradicional, veja, por exemplo, os trabalhos seminais de Adams (1990) e Snow, Burns & Griffin (1998).

² Cf. Morais (1995) e Dehaene (2012), dois importantes livros científicos com uma linguagem simples, para uma descrição detalhada deste processo.

³ Eletroencefalografia – Potenciais Relacionados a Eventos, Magnetoencefalografia.

alargou o foco das investigações mais recentes, que não se limita a observações da influência da língua falada na aprendizagem da leitura e da escrita, dirigindo-se também à influência desta última sobre o desempenho da primeira.

Num estudo usando fMRI com adultos de variados graus de literacia (iletrados e letrados na idade adulta e letrados na infância), falantes do português europeu e brasileiro, Dehaene et al. (2010), demonstraram correlações entre o grau de literacia e uma melhoria nas respostas cerebrais, que se manifesta de três modos: impulsionando a organização do córtex visual, permitindo a ativação de quase toda a rede neuronal da língua falada a partir de estímulos escritos e refinando o processamento da linguagem falada através do aperfeiçoamento de uma região fonológica específica. Por exemplo, eles observaram que estímulos falados provocavam aumento da ativação de uma área cerebral envolvida no processamento de material escrito em adultos letrados na infância e em ex-iletrados, comparativamente a adultos não letrados. Em termos simples, os autores concluem que a literacia, tanto na infância como na idade adulta, pode refinar profundamente a organização cortical. Investigações semelhantes são discutidas em Kolinsky, Pattamadilok & Morais (2012). Os autores apresentam evidências de uma série de estudos sobre o envolvimento de representações ortográficas no reconhecimento da palavra falada e no processamento metafonológico. Estudos comportamentais e funcionais envolvendo falantes do português europeu, letrados e iletrados, reportados por Castro-Caldas & Reis (2003), também indicam o impacto da escolarização na organização cerebral.

A utilização de métodos altamente sofisticados (e de alto custo) com abordagem da linguagem como uma das funções do cérebro não implica, no entanto, como bem sublinhado por Morais & Kolinsky (2001), a aceitação reducionista de que compreender a linguagem é compreender como o cérebro a produz.

Além dos métodos de investigação que permitem descrever fenómenos eletroquímicos da atividade cerebral, os métodos comportamentais descrevem as funções cognitivas em termos de processamento da informação. Eles têm fornecido as bases para os influentes modelos psicolinguísticos de processamento da leitura, no nível da palavra, da sentença e do texto (cf. Rayner & Reichle, 2010; Rayner et al., 2012). Boa parte do que sabemos hoje sobre leitura e escrita vem de estudos comportamentais prévios aos estudos de mapeamento cerebral. Estes últimos, mais sofisticados, têm, então, trazido novos conhecimentos, mas também corroborado resultados obtidos com métodos comportamentais.

A influência do conhecimento ortográfico sobre a percepção e o processamento da fala demonstrada em estudos de imagem foi também evidenciada em uma ampla variedade de estudos comportamentais (cf. Kolinsky, Pattamadilok & Morais, 2012). Por exemplo, Ziegler & Ferrand (1998) mostraram que a apresentação aos ouvintes de palavras com rimas fonológicas que podem ser escritas de várias maneiras acarretava tempos de latência auditiva mais longos e mais erros do que palavras com rimas que são grafadas apenas de uma maneira.

No entanto, as relações que se estabelecem entre os dois sistemas funcionais – escrita e fala – podem ser ainda mais complexas. Taft (2006) sugere que a representação fonológica envolvida na criação de um *output* de som (pronúncia) a partir de um *input* ortográfico é mais abstrata em natureza do que a forma fonémica da palavra. Ele considera que uma representação fonológica pode não corresponder à sua pronúncia de superfície mas a uma representação ainda mais abstrata, ou modificada, desta. Em outras palavras, as representações criadas durante o processo de aprendizagem da leitura seriam diferentes da própria representação fonológica que a palavra tem. Ou seja, quando lemos, não associaríamos somente a representação fonológica da palavra na língua com a sua pronúncia, mas mais propriamente a associaríamos a uma representação construída a partir do *input* visual e da forma fonológica com a cadeia de sons realizada na fala.

O autor conduziu uma experiência com falantes do inglês australiano, uma variedade em que o /r/ pós-vocálico não é pronunciado. Por exemplo, *car* /ka:r/ é pronunciado /ka:/ e *court* e *caught*

⁴ Imagem por Ressonância Magnética Funcional, Tomografia por Emissão de Pósitrons, Estimulação Magnética Transcraniana.

são homófonas, sendo pronunciadas /kɔ:t/. Os itens de teste incluíam não-palavras cuja pronúncia poderia ser homófona a uma palavra real – pseudo-homófonos – e eram de dois tipos: (i) pseudo-homófonos apresentando um conflito com a palavra real, a qual poderia ser gerada a partir deles em termos de presença ou ausência de um *r* pós-vocálico (p. ex., a não-palavra *cawn* poderia, na variedade testada, ser homófona da palavra *corn* e, do mesmo modo, *forl* ser homófona da palavra *fall*) – a condição experimental que incluía esses itens era chamada *R*; (ii) pseudo-homófonos que não apresentavam conflito com a suposta representação fonológica subjacente da palavra real (p. ex., a não-palavra *soke* poderia ser homófona da palavra *soak* e *troo* ser homófona da palavra *true*) – condição *não-R*. Os participantes eram instruídos a ler cada pseudo-homófono para si próprios (sem ler em voz alta, apenas subvocalmente) e então decidir se o item poderia ou não ser pronunciado de maneira idêntica a uma palavra real. Ao final desta tarefa, os itens eram apresentados novamente e os participantes deveriam proceder da mesma maneira que na tarefa anterior, mas, desta vez, pronunciando em voz alta os itens e ouvindo-os cuidadosamente antes de darem suas respostas. Os resultados mostraram que os participantes cometiam menos erros de julgamento nos itens da condição *não-R* (*forl- fall*) do que na condição *R* (*cawn- corn*) e também menos erros quando os itens eram lidos em voz alta do que quando eram processados silenciosamente.

Para explicar esses achados, Taft traça o seguinte raciocínio (p. 73-4): com respeito à diferença entre as condições *não-R* e *R*, ele argumenta que enquanto a representação fonológica gerada a partir de *soke* corresponde à representação fonológica armazenada para *soak*, a representação fonológica gerada a partir de *cawn* não corresponde àquela armazenada para *corn*. A sugestão é que *corn* seja representado fonologicamente algo como /korn/, refletindo o *r* que é evidente na grafia da palavra e, portanto, é difícil ativá-lo como um homófono de *cawn*. Com respeito à maior precisão nas respostas na tarefa de leitura em voz alta, comparativamente à tarefa de leitura silenciosa na condição *R*, ele diz que o melhor desempenho da primeira demonstra que os itens *R* eram verdadeiramente homofônicos em relação às palavras-base (as palavras reais) e que o desempenho inferior na tarefa de leitura silenciosa indica que era mais difícil para os falantes estabelecer tal homofonia a partir apenas da representação fonológica. Pode-se considerar que o desempenho inferior observado na tarefa de leitura silenciosa, em relação à leitura em voz alta, tenha ocorrido também porque esta tarefa seguia a primeira e, por conseguinte, representava uma segunda oportunidade para que os participantes julgassem adequadamente as suas respostas. Mas os resultados não podem, em princípio, ser explicados apenas por isto, uma vez que a diferença de respostas afirmativas para os itens *R* e *não-R* foi significativa na tarefa de leitura silenciosa, mas não significativa na tarefa de leitura em voz alta.

Nos modelos correntes de reconhecimento de palavras considera-se, em geral, que a representação fonológica ativada a partir do *input* escrito é idêntica à pronúncia de superfície. No entanto, os resultados de Taft (2006) sugerem, em alternativa, como mencionámos antes, que a aprendizagem da leitura também pode modificar a própria natureza das representações lexicais, seja pela incorporação de informação ortográfica ou por as refinar (Kolinsky, Pattamadilok & Morais, 2012). Uma compreensão mais apurada sobre essas observações é importante porque nos remete a outros horizontes de investigação (Taft, 2006, p. 75):

(i) é possível que o desempenho de um leitor iniciante esteja baseado diretamente nas características do *input-output* da língua mas o desempenho do leitor mais avançado pode refletir um sistema fonológico mais refinado, com representações mais abstratas que medeiam entre ortografia e pronúncia;

(ii) essas representações mais abstratas podem melhorar a eficiência do sistema de processamento lexical, talvez não apenas na fluência da leitura e da escrita mas também para o processamento da linguagem falada;

(iii) se a ortografia influencia a natureza da representação fonológica, então o processamento fonológico de um leitor avançado, de um leitor iniciante e de uma pessoa iletrada serão qualitativamente diferentes.

Desempenho em escrita

Comparativamente à leitura e audição, os estudos sobre o desempenho em escrita são menos numerosos nas ciências cognitivas que se ocupam desses tópicos (Psicologia Cognitiva, Psicolinguística, Neurolinguística). A razão mais frequentemente reportada é a dificuldade em aceder ao mecanismo de processamento da escrita por meio de técnicas experimentais tradicionalmente empregadas para a investigação de processos receptivos. É mais difícil recorrer a uma medida de tempos de reação para examinar as operações do sistema de escrita do que as dos sistemas de leitura e audição e, portanto, também é mais difícil fazer inferências sobre o primeiro do que sobre os últimos (mas cf. Guinet & Kandel, 2010; Shen, Damian & Stadthagen-Gonzalez, 2013; Kandel, Peereman & Ghimenton, 2014). No entanto, é possível prever a sua arquitetura e interpretar qualitativamente dados de produção para compreender a natureza das informações que constituem o conhecimento ortográfico

No que respeita aos processos cognitivos, leitura e escrita têm em comum o fato de envolverem a conversão de representações mentais – conversão grafema-fonema, na leitura, e fonema-grafema, na escrita – em uma forma manifesta (pronúncia ou escrita). Vistos isoladamente, ou em suas interações, os dois sistemas funcionais são igualmente complexos, como ilustrado em Morais (1995, p. 133-34) no relato de caso descrito por Bramwell, em 1897, sobre uma jovem de 26 anos que sofrera um acidente vascular encefálico:

Após algumas semanas de recuperação, ela era capaz de ler e escrever, e falava corretamente; mas era totalmente incapaz de compreender o que lhe diziam, embora não tivesse ficado surda (ela própria dizia: "Não é estranho que eu possa ouvir o tique-taque do relógio e não ouvir quando vocês falam?"). Felizmente para ela, essa mulher encontrou um meio de entender seus interlocutores. Depois de ouvir uma frase, ela era capaz de repeti-la, em seguida a escrevia, e depois relia, o que lhe permitia compreender.

A paciente de Bramwell seguia uma trajetória que começava pela análise acústico-fonética do estímulo ouvido e, portanto, reconhecia a forma acústica das palavras (mas não o seu significado). Através de processos de conversão fonema-grafema, ela escrevia as palavras ouvidas e, uma vez que o seu sistema de reconhecimento de palavras escritas estava intacto, a paciente era capaz de aceder às representações semânticas do estímulo através dele. Ela compreendia, assim, o que lhe era dito graças às inter-relações entre os sistemas de compreensão de fala, leitura e escrita. Isto também é muito interessante porque revela que esta paciente associava o significado da palavra escrita à sua imagem visual. Isso significa que há dois mecanismos de estabelecimento de relação com o significado das palavras: um por via da sua forma gráfica e outro da sua forma fonológica ou fonética.

A habilidade para formar palavras corretamente a partir de letras individuais, a que chamamos *ortografia*⁵, exige a computação das representações do grafema, entendido não como sinónimo de *letra* mas, em um sentido mais abstrato, como uma unidade contrastiva mínima, incluindo, em línguas como o português, qualquer letra ou sequência de letras que representam um único som (como os dígrafos).

De modo simplificado, a arquitetura do sistema de escrita é composta pelos seguintes componentes:

- (i) um componente que analisa o sinal de fala em unidades funcionais sublexicais (fonemas, sílabas e outras unidades relevantes, tais como bigramas);
- (ii) um léxico fonológico onde são armazenadas palavras familiares e que é acedido a partir, por exemplo, de tarefas ditadas e na compreensão da fala;

⁵ Usamos os termos *escrita* e *ortografia* de modo intercambiável no contexto deste artigo, mas guardando mais o primeiro para nos referirmos especificamente à habilidade geral que envolve operações não apenas no nível da palavra mas também da sentença (sintáticas e semânticas), embora estas últimas não sejam o foco aqui.

- (iii) um componente semântico, ativado pelo léxico fonológico, que armazena o significado das palavras;
- (iv) um léxico ortográfico, ativado pelo sistema semântico, ou diretamente pelo léxico fonológico;

As representações ortográficas têm especificada a ordem dos grafemas da palavra a qual, supõe-se, seja recuperada da memória a partir da sua forma global, no caso de palavras familiares, ou a partir da junção dos sons em uma forma sonora plausível na língua, no caso de palavras desconhecidas ou de pseudopalavras (Ellis, 1988; Caramazza & Miceli, 1990)⁶. Até este ponto, essas representações consistem em letras abstratas cujo formato (escrita à mão, digitação, soletração em voz alta) apenas será especificado em etapa posterior (a dos processos periféricos), envolvendo assim processos de conversão das representações abstratas do grafema em seus respectivos formatos e nomes que serão expressos na modalidade escrita (conversão grafema-formato) ou oral (conversão grafema-nome). Acredita-se que para manter a sequência correta de grafemas durante a transição dos processos centrais para os processos periféricos exista um componente (*buffer* grafêmico) que opera até que a tarefa de conversão fonema-grafema da palavra esteja concluída (Tainturier & Rapp, 2001).

A autonomia e as relações entre os componentes do sistema de escrita e deste com os sistemas de leitura e fala têm sido propostas com base na análise e interpretação de dados especialmente obtidos nas produções de pacientes com disgrafias adquiridas e, em menor proporção, nas de sujeitos normais em seus *deslizes de transcrição – slips of the pen* (cf. Tainturier & Rapp, 2001, para uma revisão). Nos diferentes estudos, a forma de obter os dados difere conforme o *input* oferecido (p. ex., ditado de palavras isoladas, sentenças ou textos; nomeação e/ou composição a partir de imagens, etc.) e o *output* solicitado (escrita à mão, digitação, soletração oral).

Em um estudo amplamente conhecido e que inspirou muitas das propostas sobre a estrutura das representações gráficas, Caramazza & Miceli (1990) analisaram exaustivamente a produção de 5089 palavras, ditadas no decorrer de cinco meses a um paciente de 64 anos, falante nativo do italiano, com nível superior de escolaridade, o qual sofrera um acidente vascular encefálico (AVE). A condição provocou lesões parietais do hemisfério esquerdo, causando prejuízo grave em seu desempenho de leitura e escrita que persistiram após um ano desde o AVE. Os erros foram divididos em diferentes tipos, incluindo substituições, inserções, omissões e transposições. Os erros podiam ser simples (substituição ou transposição envolvendo apenas uma letra) ou mistos (palavras contendo mais de um tipo de erro).

Uma primeira análise quantitativa mostrou uma influência marcante do tamanho da palavra sobre o desempenho (a taxa de respostas incorretas aumentava em função do tamanho da palavra (número de letras) nos 36.3% de erros observados). A maioria dos erros consistia em substituições, omissões e transposições de letras, havendo poucos erros de inserção. Os dados também mostraram uma interação significativa entre tipo de erro e tamanho da palavra, com uma marcada diminuição na proporção de erros envolvendo apenas uma letra em palavras curtas e um correspondente aumento na proporção de erros envolvendo mais do que uma letra em palavras longas. Dada esta influência do tamanho da palavra no desempenho do paciente, os autores realizaram uma análise quantitativa limitada às palavras de seis letras (3300 das 5089 palavras ditadas), procedendo ainda a uma análise qualitativa envolvendo todos os estímulos. Os tipos de erro foram detalhadamente descritos e interpretados, considerando-se, entre outros fatores, a estrutura silábica (simples vs. complexa) dos estímulos:

⁶ Esta parte do processo de escrita pode ser chamada, seguindo Ellis (1988), de *processos centrais*, distinguindo-se da parte que envolve *processos periféricos*, isto é, a tradução das representações ortográficas (*output*) em uma imagem gráfica – seja por meio de escrita à mão ou digitação – ou sonora, por meio de soletração em voz alta.

(i) Sequências-simples (CV): por exemplo, *tavolo* (mesa) - CVCVCV;

(ii) Sequências-complexas (com sílabas complexas): por exemplo, *premio* - CCVCVV, *nastro* (fita) - CVCCCV, *attore* (ator) - VCCVCV.

Entre os achados sobre a estrutura silábica, destacamos os seguintes: palavras com sequências simples foram corretamente grafadas com mais frequência do que palavras com sequências complexas. Erros de substituição sempre envolviam a substituição de uma consoante por outra consoante e uma vogal por outra vogal, tanto em palavras simples como complexas. Os erros de transposição de letras não adjacentes também respeitavam a distinção consoante-vogal, ou seja, consoantes eram transpostas por consoantes e vogais por vogais. Erros de omissão, inserção, substituição e transposição de letras adjacentes foram frequentemente produzidos em palavras complexas, mas praticamente não produzidos em palavras simples. O paciente produzia mais erros e mais tipos de erros em grupos CC do que em sequências CV. O desempenho com as consoantes geminadas era radicalmente diferente dos outros grupos CC – elas eram mais resistentes ao erro e não estavam envolvidas em alguns tipos de erros normalmente presentes para outros grupos CC (por exemplo, erros de substituição). Provavelmente, o grau de complexidade das geminadas é menor para o paciente do que o das estruturas mais complexas (como as estruturas *Cr*, *Cl*, em português ou de estruturas ainda mais complexas como as sequências consonânticas de palavras como *afta*, *pneu*, *psicologia*, por terem núcleos vazios). Um grupo de dados, visto em conjunto, sugeria que a estrutura grafêmica e não a fonológica determinava a distribuição de erros no desempenho do paciente. Especificamente, os erros violavam o princípio da sonoridade e os grupos CC e CV correspondendo a fonemas simples comportavam-se como dois elementos grafêmicos e não como uma unidade fonológica simples.

Esses dados sugerem que as representações grafêmicas especificam mais do que a identidade dos grafemas (a especificação dos grafemas individuais que compõem uma palavra) e sua ordem linear na palavra, consistindo também em informações sobre a estrutura silábica, com a distinção consoante-vogal. Além desses efeitos silábicos, os autores ainda consideram que as representações grafêmicas tenham um caráter grafotático mais do que fonotático. Os erros do paciente mostravam, de modo consistente, que muitas restrições fonotáticas eram violadas (p. ex., *sfondo* (fundo) → **fsondo*; *scalda* (aquece) → **slcada*).

Quanto à diferenciação entre consoantes e vogais, esta também se refletiu em dados de dois pacientes de Cubelli (1991). Um deles omitia todas as vogais das palavras e o outro cometia quase só erros envolvendo vogais, sugerindo que o estatuto de consoantes e vogais é especificado diferentemente no processo de execução da escrita, podendo ser afetado de forma seletiva após lesão encefálica.

É curioso verificar que os dados de Kandel et al. (2009), com crianças do 3º, 4º, e 5º anos de escolaridade, falantes nativas do francês, sugeriram que a sílaba ortográfica, e não a sílaba fonológica, era psicologicamente real como unidade de processamento na escrita.

Em um estudo comportamental simples, porém muito revelador, Veloso (2007) também apontou indicações da influência do conhecimento ortográfico sobre o conhecimento fonológico. Ele aplicou uma tarefa oral de segmentação silábica de palavras contendo sequências mediais de /S/+obstruente (*mosca*, *ginástica*, *floresta*, *rasga* e *cesto*) a 42 crianças em dois momentos diferentes: no último mês do primeiro ano e no último mês do segundo ano de escolaridade. Em português europeu, essas sequências são geralmente consideradas heterossilábicas (C.C) em termos fonotáticos (Mateus et al., 2003, p. 996), embora alguns autores argumentem em favor de uma interpretação fonológica homossilábica – .CC (Cf. Veloso, 2007 e suas referências). Em termos de regras ortográficas, elas são sempre consideradas heterossilábicas. A recolha dos dados capturava então observações anteriores (1º ano) e posteriores (2º ano) à exposição/familiarização explícita das regras ortográficas de segmentação silábica. Para o 1º ano, os resultados mostraram uma maior frequência de segmentações homossilábicas (p. ex., [‘mo.jke]) – 53,1% – do que heterossilábicas ([‘moj.ke]) – 46,9%. No 2º ano, em contrapartida, a frequência de segmentações homossilábicas foi

significativamente inferior à frequência de segmentações heterossilábicas (17,5% e 82,5%, respetivamente). Note que, enquanto no 1º ano o julgamento das crianças sobre a segmentação homossilábica/heterossilábica das palavras-alvo testadas não era de todo consistente, embora significativamente diferentes, no 2º ano as mesmas crianças as julgavam como heterossilábicas de modo bem mais consistente. Resultados como estes indicam que a escolarização parece ter efeitos não apenas sobre o desempenho ortográfico dos sujeitos, estendendo-se também ao próprio conhecimento fonológico destes (p. 9).

Alves (2012) examinou, em português europeu, o efeito de propriedades segmentais, na emergência da consciência segmental e no desempenho de leitura e escrita de palavras e pseudopalavras, em crianças ainda com pouco domínio do princípio alfabético. Entre os resultados do estudo, a autora conclui que (i) há efeitos de propriedades segmentais, particularmente das relativas ao modo de articulação, isoladamente e em sua relação com o ponto de articulação ou com o vozeamento, na capacidade de deteção segmental e (ii) no desempenho de leitura e escrita de palavras e pseudopalavras, com maior significância quanto ao modo de articulação.

Santos (2013) analisou, também no português europeu, diferenças entre a aquisição de ataques ramificados e de sequências consonânticas problemáticas, em dados de nomeação oral e de escrita de um mesmo conjunto de palavras isoladas produzidas por dois grupos de crianças, 27 do 1º ano e 29 do 4º ano de escolaridade. Nos estímulos do 1º ano foram incluídas oito palavras dissilábicas com ataques ramificados compostos por oclusiva+vibrante (*bruxa, prato*), fricativa+vibrante (*fruta, frasco*), oclusiva+lateral (*globo, planta*) e fricativa+lateral (*flauta, flores*). As sequências consonânticas incluíam cinco palavras contendo oclusiva+nasal (*gnomo, pneu, pigmeu, submarino*) e fricativa+oclusiva (*afta*). Nos estímulos do 4º ano foram incluídas oito palavras polissilábicas contendo os mesmos ataques ramificados descritos acima (*astrónomo e radiografia; africana e confronto; atlântico e biblioteca; afluente e insuflável*). As sequências consonânticas problemáticas incluíam sete palavras contendo oclusiva+oclusiva (*pictograma, neptuno*), oclusiva+fricativa (*fricção, tsunami*), oclusiva+nasal (*enigma, algoritmo*) e fricativa+oclusiva (*oftalmologista*). Entre os estímulos foram ainda incluídos sete distratores no 1º ano e 11 no 4º. As tarefas foram aplicadas em sessões individuais e pedia-se às crianças que nomeassem o mais rapidamente possível as imagens que eram apresentadas no ecrã do computador, sendo suas respostas gravadas. Imediatamente após a nomeação de cada palavra, elas deveriam escrever a palavra em uma folha de papel contendo as imagens apresentadas. Entre os achados deste trabalho, destacamos os seguintes: foram observados mais desvios na escrita das palavras examinadas do que na sua produção oral, sendo esta discrepância maior no 1º do que no 4º ano, não se verificando, assim, um paralelismo direto entre desenvolvimento fonológico e aprendizagem da escrita; as crianças do 1º ano recorreram às mesmas estratégias de reconstrução das estruturas examinadas na oralidade e na escrita e as crianças do 4º ano, por sua vez, demonstraram domínio completo do sistema fonológico, embora ainda com desvios na produção escrita nas estruturas silábicas examinadas. A autora considera que estes últimos resultados sugerem que as estratégias usadas no processo de aquisição parecem ficar disponíveis cognitivamente para serem reativadas mais tarde na escrita (p. 84); no 1º ano, os dados de oralidade mostraram mais desvios nos grupos consonânticos problemáticos do que nos ataques ramificados e os dados de escrita seguiram a mesma direção, porém com um percentual de desvios bem maior; no 4º ano não foram observados desvios na produção oral de ataques ramificados e apenas três desvios em grupos consonânticos problemáticos; na produção escrita deste grupo, por outro lado, um percentual maior de desvios foi observado, sendo, como no 1º ano, maior nos grupos consonânticos problemáticos do que nos ataques ramificados.

Em resumo, no campo dos estudos sobre a leitura, assistimos a uma mudança radical no conceito de dislexia nas décadas de 1980 e 1990, em parte resultante dos avanços marcantes nas investigações sobre os processos cognitivos envolvidos na aprendizagem da leitura, sobre os meios de prevenir o insucesso escolar nos anos iniciais e sobre os meios de otimizar o aprendizado da leitura (cf., p. ex., Adams, 1990; Snow, Burns & Griffin, 1998). A década de 1990, conhecida como a década do cérebro, trouxe um grande número de importantes estudos, apoiados em métodos de

mapeamento cerebral, sobre as diversas facetas da habilidade da leitura e alargaram o foco das investigações. Os estudos de Taft (2006) e de Ziegler & Ferrand (1998), por exemplo, são particularmente interessantes porque demonstram as consequências da complexidade estrutural de unidades fonológicas e ortográficas sobre o processamento da leitura e, em última análise, ilustram algumas das importantes questões que foram, e ainda são, tratadas nos diversos estudos sobre o reconhecimento de palavras escritas. O cenário sobre a investigação em leitura, no entanto, é bastante diferente do cenário das investigações sobre a escrita. Entre as inúmeras razões desta discrepância, como observado por Perfetti (1997, p. 21), pode-se dizer que o privilégio científico, herdado da linguística, foi dado à linguagem falada e que a ortografia costuma ser vista menos como um problema científico do uso da língua do que como uma convenção da alfabetização ou um assunto escolar. Ele ainda nota que, talvez, a simplicidade da ortografia tenha mascarado suas intrigantes questões, tornando-a menos desafiadora do que os processos mentais superiores que têm preocupado as ciências cognitivas (p. 23). É importante sublinhar que a investigação dos mecanismos e processos envolvidos na produção escrita é importante e necessária nas diferentes línguas. O estudo da ortografia (e, em certa medida, o da leitura também) tem sido dominado por investigações de um tipo de sistema de escrita, a alfabética, e o de uma ortografia, a do Inglês (uma língua sabidamente opaca em termos de forma fonológica e forma ortográfica), o que limita conclusões que precisamos compreender (Perfetti, 1997). É claro que os investigadores preocuparam-se no contraste entre diferentes ortografias, com estudos como o de Caramazza & Miceli (1990), por exemplo, que inspirou muitas das propostas que temos hoje sobre a estrutura das representações grafêmicas. No entanto, esses estudos ainda são escassos.

3. Ortografia no contexto do ensino da língua

Le Bouch (1982), quase antecipando os reflexos dos avanços científicos na formação dos professores, que viriam de diferentes disciplinas nas décadas subsequentes, nota que o professor precisa ter um conhecimento amplo sobre o desenvolvimento da criança e que isto não se trata de uma *medicalização* ou *psiquiatrização* da escola, mas de se criar condições para que o professor possa conhecer a criança e aplicar o que é mais adequado à sua fase de desenvolvimento e prevenir/intervir nas suas dificuldades. Com a mesma visão, Goodman (1995) considera que o conhecimento do professor sobre a natureza da linguagem oral e escrita e sobre o desenvolvimento da criança o capacita a escolher e aplicar estratégias apropriadas a cada momento deste desenvolvimento. Zunino & Pizani (1995) acrescentam que conhecer o processo de aprendizagem, do ponto de vista geral e do ponto de vista de cada aluno, permite ao professor entender as produções e interpretações das crianças e, dessa maneira, saber o que há por detrás das suas perguntas e respostas. Além disso, permite antecipar as hipóteses das crianças sobre as questões que para elas são relevantes em cada fase da aprendizagem.

Muito embora esses autores não estivessem se referindo diretamente a uma formação do professor que incorpore os conhecimentos advindos das ciências cognitivas, como alguns dos que foram apresentados na secção anterior, as considerações são perfeitamente aplicáveis aos conhecimentos que deveriam fazer parte desta formação. Para além da componente de formação dos professores, a articulação da atividade dos terapeutas da fala com a dos professores na fase inicial de aprendizagem da escrita deve ser estrategicamente pensada.

Numa investigação sobre a ação do terapeuta de fala na sala de aula, em colaboração com o professor, Lourenço-Gomes (1999) acompanhou dois grupos de crianças de segundo ano em duas escolas públicas do Estado do Rio de Janeiro. Após identificar algumas dificuldades quanto à ortografia, leitura e consciência fonológica, a autora elaborou estratégias de otimização e intervenção que foram aplicadas, alternadamente, a um grupo experimental e a um grupo de controlo. O desempenho dos dois grupos foi comparado numa avaliação no início de seu programa e outra no final, tendo sido observado melhor desempenho do grupo experimental, em comparação com o grupo de controlo, na última avaliação. Com base nos seus dados a autora reforça que a

intervenção em sala de aula pode trazer benefícios imediatos às crianças com dificuldades e pode prevenir essas mesmas dificuldades. Além disso, rentabiliza o desenvolvimento das outras crianças e contribui para a minimização das consequências individuais e sociais decorrentes do insucesso escolar.

Rodrigues (2013) também avaliou o impacto de um programa de intervenção de sensibilização à linguagem escrita, mas em crianças em idade pré-escolar, com um nível de escrita pré-silábico e que possuíam conhecimento sobre o nome das letras, controlando ainda as variáveis idade e inteligência. O estudo envolveu um grupo experimental e um de controlo e consistiu na aplicação de um pré e um pós-teste, intermediado pela realização do seu programa. A intervenção ocorreu em 45 sessões de aproximadamente 30 minutos divididas conforme o plano de atividades programadas. Os resultados mostraram um progresso significativamente superior do grupo experimental em relação ao grupo de controlo no que respeita às seguintes tarefas: (i) número de correspondências grafofônicas estabelecidas para palavras ditadas. Por exemplo, para a palavra *pato* na sequência de letras *BRUN* escrita, nenhuma correspondência era cotada, para *PMRI* era dada cotação 1, para *PTQ* cotação 3, e assim por diante; (ii) classificação de sílaba e de fonema inicial de palavras. As crianças deveriam apontar em um conjunto de imagens representativas de palavras, sendo solicitado à criança que indicasse aquelas que começavam com a mesma sílaba ou som, conforme a prova; (iii) análise silábica, em que a criança era solicitada a pronunciar as sílabas que compunham palavras representadas por imagem, embora neste item o grupo de controlo tenha apresentado também uma evolução entre o pré e o pós-teste, porém em menor proporção. Além disso, a autora observou um aumento ligeiramente superior, embora não significativo, no grupo experimental em relação ao grupo de controlo, do número de *razões funcionais para aprender a ler* apontadas pelas crianças como, por exemplo, *Vou ler sozinho o que está escrito nos meus desenhos*, *Ler uma carta importante* (razões não-funcionais incluíam respostas como (...) *porque não quero fazer tudo mal; quero fazer bem; Ficar em casa com os manos*).

Muitos trabalhos têm dirigido a atenção para o desenvolvimento de competências de crianças e jovens em contexto escolar, propondo estratégias de otimização e facilitação da aprendizagem da leitura e escrita. Apenas para ilustrar, Scliar-Cabral (2003a,b) reuniu uma série de conceitos sobre fala e ortografia, baseados no sistema de sons do português do Brasil, e propôs uma variedade de estratégias pedagógicas em duas obras que se complementam, diretamente dirigidas a educadores dos primeiros anos de escolaridade. Com um enfoque centrado no ensino e aprendizagem da ortografia, destaca-se aqui também a série de artigos apresentada em Silva, Morais e Melo (2005) e, ainda, a edição 35, organizada por Miranda & Cunha (2010) do *Cadernos de Educação*, que reúne estudos principalmente desenvolvidos no âmbito de grupos brasileiros de investigação sobre a aquisição da escrita. Em Freitas, Gonçalves & Duarte (2010) encontra-se uma série de artigos que revelam o carácter incremental e multidimensional do conhecimento implícito da língua, e de seu impacto no aprendizado da leitura e escrita, e que sublinham a importância da experiência linguística da criança no decorrer do seu desenvolvimento, considerando-se que *é tarefa da escola promovê-la, adotando um conjunto de práticas diversificadas, adequadas ao nível etário e de ensino* (p. 15). Com uma abordagem semelhante de interface da Linguística e do Ensino da Língua, Freitas, Rodrigues, Costa & Castelo (2012) apresentam uma descrição do sistema de sons do português europeu, considerado na forma padrão e discutido em termos de diversidade dialetal. Destaca-se, ainda, a exposição reflexiva de Sousa & Cardoso (2005) especificamente centrada no desenvolvimento de competências metalinguísticas no Ensino da Língua Portuguesa nos 1º e 2º ciclos, e as propostas didáticas de Castelo (2012) para a promoção da competência metafonológica nos ensino básico e secundário após um minucioso trabalho que examinou o desempenho desta competência em falantes jovens, nativos do português europeu, no 5º, 7º e 10º anos e no 1º do ensino superior.

Outros estudos concentraram-se na descrição e análise das FN-Cs produzidas pelas crianças e/ou na evolução do domínio ortográfico nos anos iniciais de escolaridade. Suas abordagens são diversificadas e, por isso, também o são as categorias e as designações atribuídas a cada uma delas. Os estudos também diferem em seus procedimentos metodológicos, seja quanto à população

examinada (perfil socioeconómico e cultural, método de alfabetização a que as crianças são expostas, etc.); ao tipo de *input* empregado (ditado de palavras isoladas, frases ou textos; nomeação ou narrativa a partir de imagens; escrita espontânea ou semi-dirigida por temas, etc.); ao período de recolha de dados (1º. ou 2º. semestre académico, com recolhas pontuais ou contínuas), ao tipo de estudo (transversal ou longitudinal), além dos diferentes critérios de inclusão na amostra. Em termos gerais, eles apresentam um carácter que varia entre o mais descritivo, mais analítico (do ponto de vista linguístico) ou mais cognitivo (do ponto de vista dos processos subjacentes às FN-Cs), conforme o campo disciplinar dos autores e os objetivos de aplicação das propostas de classificação (teóricos, clínicos e/ou pedagógicos). Dado o grande número de estudos sobre o tema, nos limitamos aqui a sugerir os trabalhos de Guimarães (2005) e Moojen (2009), para uma revisão das propostas em Português do Brasil e os estudos de Perrolas (1996), Pinto (1997), Horta (2001), Mateus (2002) e Horta & Martins (2004) para dados do Português Europeu.

A despeito da diversidade dos estudos, existe consenso sobre algumas observações gerais das FN-Cs nos anos iniciais de escolaridade, por exemplo:

(i) O tipo de FN-Cs e o número de ocorrências variam em função do ano de escolaridade.

(ii) Há um conjunto de FN-Cs de aparecimento consistente nas produções escritas das crianças, que é menos dependente de fatores como perfil socioeconómico e cultural e método de alfabetização, e mais dependente do sistema de escrita ou das normas ortográficas da língua. No caso do português, as FN-Cs são mais consistentes: FN-Cs relacionadas à estabilidade de correspondência entre o fonema e o grafema, sendo menos frequentes, no que respeita às consoantes, os fonemas que apresentam uma correspondência estável com os grafemas que os representam (/p/, /b/, /f/, /v/, /t/, /d/)⁷; relacionadas às regras contextuais, as quais estabelecem o valor de um determinado grafema consoante o contexto em que aparecem (p. ex., <c> antes de <e> e <i> ou antes de <a>, <o> ou <u>); relacionadas a irregularidades, cuja grafia é unicamente dependente da norma ortográfica fixada (p. ex., <s>, <x> e <z> para representar o som /z/, em *asa*, *exame*, *fazer*).

(iii) Em comparação com FN-Cs relacionadas às regras contextuais ou a irregularidades, as FN-Cs motivadas pela fala são particularmente evidentes no início da aprendizagem da leitura e escrita, sendo, porém, transitórias e diminuindo ou mesmo desaparecendo no decorrer da instrução formal.

Dado que o enfoque das investigações sobre o desempenho ortográfico é multifacetado, o tema é tratado por diferentes áreas, nomeadamente a da Pedagogia e a da Linguística. No entanto, uma abordagem mais descritiva do que propriamente cognitiva tem sido mais frequentemente observada. Apesar do interesse manifesto por parte da psicolinguística e da neurolinguística sobre o mecanismo de processamento ortográfico, nomeadamente em finais da década de 1990, o que sabemos hoje sobre os processos subjacentes ao aprendizado da ortografia ainda é, notoriamente, pouco. Mas, nos últimos anos, outras possibilidades de investigação apontam novas perspectivas para a compreensão desta componente do sistema de linguagem em estudos de processamento ortográfico aferido com técnicas *online*.

Tendo em conta a complexidade do estudo e da interpretação dos dados de escrita inicial, que fica bem patente na revisão acima apresentada, e com o propósito de construir uma base de dados que possibilitasse o estudo das FN-Cs mais especificamente relacionadas com a fala em português europeu, desenhamos um plano de recolha de dados. Para além de uma análise de correspondências entre fala e ortografia, esta base de dados irá servir à determinação de parâmetros para a elaboração de estudos do processamento ortográfico em PE.

⁷ FN-Cs envolvendo esses fonemas geralmente decorrem do traço [+/- vozeado], mas são mais observados em algumas crianças em particular do que de uma maneira consistente entre todas as crianças.

Neste artigo, limitamo-nos a descrever sumariamente os dados até agora recolhidos no âmbito do EFFE. Nosso objetivo, vale ressaltar, foi apresentar uma proposta que intenciona estudar a produção escrita de crianças portuguesas nos primeiros anos de escolaridade, tendo em vista duas principais metas: o estudo das correspondências que podem ser efetivamente apontadas entre a fala e a ortografia e o levantamento de indicadores para a elaboração de estudos sobre o processamento *online* da produção escrita e que permitam compreender em um nível mais abstrato essas correspondências. Embora metodologias *online* tenham sido exploradas na última década na investigação da produção escrita, o seu número é marcadamente pequeno. Por exemplo, Bonin & Meot (2002) mostraram que tanto unidades lexicais quanto sublexicais influenciam o desempenho de escrita em adultos, falantes do francês. Eles examinaram tempos de latência para a escrita de nomes, usando uma mesa gráfica e caneta de contacto, com apresentação auditiva dos estímulos. Análises de regressão múltipla demonstraram influências significativas da duração acústica dos estímulos e de outras variáveis, já bem conhecidas em estudos de leitura, tais como idade de aquisição, frequência da palavra, consistência entre forma ortográfica e fonológica e comprimento ortográfico (número de letras). Este estudo ilustra a diversidade de fatores que precisam ser examinados em diferentes línguas, de modo que os resultados de línguas com diferentes ortografias possam ser contrastados e os processos cognitivos envolvidos melhor compreendidos.

Uma análise quantitativa mais pormenorizada e uma interpretação linguística de FN-C específicas, observadas nas crianças do 2º ano de escolaridade, juntamente com dados das mesmas crianças, cursando o 4º ano na época da preparação deste artigo, serão apresentadas em estudo independente (Rodrigues, Lourenço-Gomes & Alves, em preparação). A ênfase sobre revisão da literatura procurou mostrar o caráter multifacetado dos estudos sobre o desenvolvimento da produção escrita e a discrepância entre o volume de estudos sobre esta habilidade e a habilidade de ler. Assim, deixamos expressa, também, a opinião sobre a necessidade de um maior empenho por parte dos investigadores, nas diversas línguas, no sentido de examinar, com uma abordagem francamente cognitiva, o que há por detrás das palavras que as crianças escrevem, tal como tem sido feito em relação ao reconhecimento de palavras na leitura.

4. Metodologia para as primeiras observações de *corpora* de escrita e fala correspondentes

Participantes

Participaram do estudo 48 crianças de duas turmas (A e B) do segundo ano de uma escola particular de Lisboa, de ambos os sexos, todas com sete anos e sem historial de repetência escolar ou queixa de dificuldades relatada pelos professores. Na composição da amostra foi considerado o relato por escrito de um dos pais ou responsáveis quanto aos seguintes critérios de inclusão: as crianças são falantes nativas do português europeu; não foram até à época da recolha dos dados; acompanhadas por otorrinolaringologista por motivos relacionados à audição e/ou doenças do aparelho auditivo, e por psicólogo, neurologista ou terapeuta da fala por qualquer razão; não falam mais de uma língua em casa; não viveram em outro país e não estudaram em escola bilingue. A amostra inicial foi composta por 58 crianças, sendo que destas cinco não cumpriam pelo menos um dos critérios de inclusão, quatro estavam ausentes em mais de uma data da recolha dos dados e uma não tinha o Termo de Consentimento Informado assinado pelos responsáveis. Com exceção desta última, que recebeu tarefas da mesma natureza durante a recolha, todas as outras crianças não incluídas na análise dos dados realizaram as mesmas tarefas com os materiais destinados ao presente estudo. A Tabela 1 mostra a distribuição da amostra, por sexo, em cada turma.

Tabela 1 – Distribuição da amostra, por sexo, em cada turma

Turma	Masculino	Feminino	Total
A	11	12	23

B	12	13	25
Total	23	25	48

Materials

Para a tarefa principal de produção escrita e para a de produção oral, foram usadas cinco imagens temáticas adaptadas por Guerreiro (2007) para o português europeu (PE), originalmente criadas para o português do Brasil (PB) por Yavas, Hernandorena & Lamprecht (1991) para utilização na avaliação fonológica de crianças a partir dos três anos de idade. Os temas das imagens são *Floresta, Cidade, Sala, Cozinha e Casa de banho*. Em cada imagem aparecem figuras que servem de estímulos para a nomeação de palavras-alvo e para a descrição da cena. Embora o material tenha sido criado com propósito distinto do proposto neste estudo, a sua escolha deveu-se principalmente ao fato das imagens já terem sido avaliadas, em duas variedades do português, e consideradas adequadas para a evocação dos alvos e descrição das cenas. Isto era importante pelas seguintes razões: (i) o conjunto de palavras-alvo do instrumento mencionado está balanceado com base em critérios fonéticos, fonológicos, lexicais e de frequência que serviam também aos propósitos deste estudo-piloto; (ii) esses critérios tomam por base duas variedades do português que poderiam ser contrastadas em estudo posterior com materiais análogos; (iii) a tarefa de produção escrita deste estudo foi aplicada em grupo e o investigador limitava-se a fornecer as instruções às crianças apenas no início da tarefa. Esperava-se, como já mencionado, a produção do maior número possível de produções escritas contendo palavras-alvo, representativas do sistema de sons do português europeu e arranjadas de tal modo a servirem à descrição das cenas e à criação de narrativas. Ao mesmo tempo, era desejável que as produções escritas correspondessem, do modo mais semelhante possível, às produções orais que seriam realizadas posteriormente. Era então importante que qualquer resultado que contrariasse as expectativas deste estudo-piloto não fosse atribuído a aspetos relacionados às próprias imagens.

Nas cinco imagens temáticas do instrumento adaptado para o PE, prevê-se a evocação de 127 palavras-alvo, dentre as quais foram mantidas 85 do instrumento original, eliminadas 40 e incluídas 42 para a aproximação de características fonéticas, fonológicas, lexicais e de frequência do PE (cf. Guerreiro, 2007, p. 98-113, sobre a seleção das palavras alvo em PE). Para o presente estudo, foram selecionadas 43 palavras, acrescidas de outras 14 que estavam representadas nas imagens mas não faziam parte dos alvos do instrumento original como, por exemplo, *voar, vassoura, prateleira, vermelho*. A inclusão de determinadas palavras foi possível porque as imagens do instrumento adaptado são coloridas, enquanto as do instrumento original em PB são em preto e branco. A lista de palavras-alvo utilizadas encontra-se em Anexo.

Para a primeira sessão de trabalho, que visava familiarizar as crianças com as atividades propostas posteriormente, foi utilizada a história *O chapéu*, extraída do livro *A Bruxinha Atrapalhada* (Furnari, 1991).

Procedimentos

O estudo foi dividido em três etapas⁸. Na primeira, foi realizada uma sessão de trabalho em grupo, em cada turma, com a presença dos professores, os quais haviam sido previamente informados sobre os procedimentos gerais do estudo. Nesta sessão, o investigador propunha às crianças a produção de uma composição a partir das cenas encadeadas da história *O chapéu* (Furnari, 1991). Além de estabelecer uma interação com os professores e com as crianças, esta sessão também tinha como objetivo familiarizar as crianças com as tarefas que seriam propostas em sessões subsequentes, procurando dar-se ênfase à expressão livre, criativa e detalhada nas narrativas

⁸ A recolha de dados foi realizada pelo mesmo investigador em todas as etapas do estudo.

escritas e orais a partir de imagens, ou seja, o que se pretendia estava além da nomeação e descrição dos elementos que compunham as imagens-alvo. Apenas uma criança não estava presente nesta etapa. A segunda etapa consistiu na recolha dos dados de escrita com as imagens-alvo. Em uma sessão de trabalho, também em grupo e com a presença dos professores, as crianças eram solicitadas a criar suas histórias do mesmo modo como haviam feito com a história da bruxinha. Uma vez que um dos propósitos deste estudo era a avaliação qualitativa de recursos para a elaboração de um banco de dados de fala e escrita correspondentes, não estando em causa a análise de dados das crianças individualmente neste primeiro momento, elas foram divididas em grupos e realizaram a tarefa a partir de uma única imagem. Em concordância com os nossos propósitos, era importante, como mencionado, obter dados de formas de palavras escritas e das formas orais correspondentes a partir de um conjunto controlado de palavras-alvo, considerando-se o sistema de sons do PE, e de características fonéticas, fonológicas, lexicais e de frequência dos estímulos. Cada criança, em cada grupo, recebeu uma imagem-alvo impressa em formato A4 e uma folha com linhas para a escrita das composições. Para esta secção foram despendidas cerca de duas horas, incluindo o tempo de interação do investigador com as crianças. A Tabela 2 mostra a distribuição das imagens nos grupos, nas duas turmas. Nesta sessão, e nas seguintes, as instruções que o investigador forneceu às crianças foram gravadas, permitindo a elaboração de um relatório fiel deste momento, o qual é importante do ponto de vista metodológico.

Tabela 2 – Distribuição das imagens temáticas nos grupos, em cada turma

Turma	Imagens Temáticas					Total
	FL	CI	SA	CZ	CB	
A	4	5	4	5	5	23
B	3	6	6	6	4	25
Total	7	11	10	11	9	48

FL = "Floresta"; CI = "Cidade"; SA = "Sala"; CZ = "Cozinha"; CB = "Casa de banho"

A terceira, e última, etapa do estudo consistiu na recolha de dados de produção oral. Em sessões individuais com as crianças, foi proposto que contassem novamente uma história sobre a mesma imagem que haviam recebido para a composição escrita. Durante a realização da tarefa as crianças tinham a imagem em mãos. As gravações foram feitas em uma sala reservada na própria escola, previamente visitada pelo investigador para observação de sua adequação quanto à qualidade das gravações e ao conforto das crianças. As sessões de gravação decorreram em quatro dias, dois para cada turma, e o tempo para a realização da tarefa foi de cerca de 15 minutos. As sessões foram gravadas na íntegra, de modo a incluir diálogos entre o investigador e a criança. Foi utilizado um gravador *Roland R26* com microfone interno.

Os textos das crianças foram digitalizados em *formato.jpg* e transliterados em *formato.docx*. Os dados de fala foram transcritos foneticamente (IPA) na íntegra. Foi criada uma base de dados em *formato.xlsx* com os registos das produções escritas e, no caso das palavras-alvo, das respectivas produções orais de cada criança. Desta forma, temos uma base de dados que possibilita uma correlação efectiva entre fala e escrita de cada uma das crianças. Os dados são quantitativamente referidos na secção seguinte.

5. Avaliação global dos primeiros dados do Projeto EFFE

Os primeiros dados do EFFE demonstram que os procedimentos empregados, de modo geral, produziram resultados em conformidade com o primeiro objetivo do NICEF, designadamente a composição de um banco de dados de fala e escrita em PE nos primeiros anos de escolaridade, que sirva um leque alargado de estudos em diferentes áreas e com diferentes abordagens. Foi possível

obter um conjunto de dados de fala e escrita correlacionável para análise linguística minuciosa de FN-Cs motivadas pela fala, em sua face fonética e fonológica⁹.

A Tabela 3 mostra o total de palavras escritas, com formas convencionais e não-convencionais, observado na tarefa de produção de texto a partir das imagens temáticas. Elas estão separadas em *alvos*, correspondendo às palavras-alvo do instrumento utilizado, e *não-alvos*, correspondendo a qualquer outra palavra não constante da lista de alvos, incluindo palavras funcionais. A inclusão dessas últimas deveu-se ao fato de também nelas aparecerem grafias não-convencionais. A produção de narrativas (e não apenas a nomeação das palavras-alvo) permitiu, assim, a obtenção de um número considerável de FN-Cs para análise (724 – 84 alvos e 640 não-alvos).

Tabela 3 – Total de ocorrência de formas convencionais e não-convencionais observado nas duas turmas (Imagens).

Turma	Alvos		Não-alvos		Total palavras**
	FC	FN-C	FC*	FN-C	
A	106	22	1269	282	1681
B	184	62	1851	358	2453
Total	290	84	3120	640	4134

* Inclui palavras de função. ** São consideradas as palavras que envolvem desvios, independentemente do número destes no interior de cada palavra. FC = formas convencionais; FN-C = formas não-convencionais

Como foi exposto na descrição dos procedimentos, em concordância com os nossos propósitos, era importante obter dados de formas de palavras escritas e das formas orais correspondentes a partir de um conjunto controlado de palavras-alvo, considerando-se o sistema de sons do PE, suas características fonéticas, fonológicas, lexicais e de frequência dos estímulos. Neste primeiro momento, estivemos mais concentradas no exame das formas orais e escritas que fossem características do grupo que compunha a amostra e não em uma avaliação das crianças em particular, muito embora algumas observações características de algumas crianças tenham sido possíveis. Por esta razão, as imagens temáticas foram divididas, de modo equilibrado, entre grupos distintos de crianças (cf. Tabela 2, acima). Com isto, esperávamos que, em conjunto, as produções de fala e de escrita de palavras-alvo garantissem a observação de formas características da amostra nas duas modalidades. A Tabela 4 mostra o total de correspondências que foi obtido entre fala e escrita a partir das palavras-alvo em cada uma das cinco imagens temáticas.

Tabela 4 – Correspondências de fala e de escrita obtidas para as palavras-alvo

Modalidade	Imagens Temáticas					Total
	FL	CI	SA	CZ	CB	
Escrita	64	69	102	55	41	331
Fala	136	275	318	121	180	1030

FL = "Floresta"; CI = "Cidade"; SA = "Sala"; CZ = "Cozinha"; CB = "Casa de banho"

⁹ Cf. os primeiros resultados desta análise, mais pormenorizada, em Rodrigues, Lourenço-Gomes e Alves (em prep.).

Com base apenas na tarefa de produção escrita a partir das imagens temáticas, vê-se que a frequência de FN-C excedeu em muito a da produção de palavras-alvo, revelando a vantagem da utilização do instrumento para a produção de narrativas e não apenas para a nomeação, uma vez que algumas estruturas presentes em palavras não-alvo eram comparáveis às das palavras-alvo.

Dentre as FN-Cs, verificamos diferentes tipos que podem ser interpretados como relacionados, direta ou indiretamente, com as produções fonéticas observadas entre as crianças mas também outros que revelam instabilidades estritamente ortográficas, sem nenhuma correspondência com a fala e gerados por diferentes razões. A título de ilustração, destacamos alguns exemplos, como segue.

Apercebemo-nos da existência de FN-Cs que estabelecem uma relação com a fala de determinadas crianças. Elas ocorrem, evidentemente, em estruturas fonológicas de vários tipos. Por exemplo, as estruturas fonológicas mais complexas são mais suscetíveis a grafias não convencionais. Existem diferentes graus de complexidade, no entanto, dependentes do tipo de constituinte silábico e do tipo morfosintático da palavra. O tipo de relação existente entre as FN-Cs e a fala pode variar, como a comparação com as pronúncias das crianças demonstra. Há casos que só indiretamente apresentam uma relação com a fala.

(1) Estruturas fonológicas com complexidade no ataque revelam FN-Cs:

- (i) Com inserção de vogal (a que for menos marcada em termos gráficos e que puder representar um vazio acústico - em PE será preferencialmente um <e>) entre as duas consoantes (*<esterelas> sendo realizada sistematicamente como [ʃtréle]) (criança 48_CZ_2A_CM, por exemplo) por <estrelas>; *<peresente> por <presente>; *<afelitinho> por <aflitinho>; *<deseliguar> por <desligar>). Embora os nossos dados demonstrem que essas FN-Cs não estejam diretamente relacionadas com a fala, é possível que, para lidar com certas complexidades, as crianças realizem silabação (ainda que mentalmente ou de forma sussurrada) como apoio à escrita, o que resultaria na inserção da vogal.
- (ii) Em ataque com segmento africado <táxi> a criança poderia pensar que haveria [k] seguido de [s] como unidades separadas e, portanto, a palavra pode aparecer como *<taquesi>, por exemplo. Também, aqui, não foram observados casos de inserção nas realizações das crianças, podendo ser a inserção na escrita decorrente de silabação na fase de aprendizagem.
- (iii) Em ataque com consoante labializada, como /k^w/ ou /g^w/ palavras como <quando> podem aparecer grafadas como *<coando>. Os dados de fala demonstram realizações sistemáticas do /k^w/ como ataque da sílaba inicial de que é núcleo a vogal nasalizada, não justificando directamente as FN-Cs observadas. Mas vale notar que esses são casos em que <o> muitas vezes pode representar na escrita o som [w], tal como <u>.
- (iv) Sílabas ultra-complexas como a de *transformou* geram muitos erros na escrita, especialmente se forem seguidas de outras consoantes (*<estranformo> [trẽʃfurmó], por exemplo).

(2) Estruturas fonológicas com complexidade na coda revelam FN-Cs:

- (i) /r/ parece ter fraca associação à estrutura da sílaba, uma vez que é frequentemente ressilabificado (ma.ra.zul por *mar azul*), elidido (*amo∅ real*; *faze∅ muito*), mas também pode estar apoiado em vogal paragógica antes de uma pausa (*Eu não quero*

fazer[i].), contrariamente ao que se poderia esperar. Assim, as crianças incluem com frequência <e> em final de palavra após /r/.

- (ii) A coda fricativa em PE corresponde a dois sons apenas, [ʃ] ou [ʒ], ocorrendo este último sempre antes de consoante vozeada, mas podendo ser escrito como <s> ou como <z> (*pés, fiz*), as crianças, em palavras que devem levar <z> grafam, por vezes, um <s> (*eu fiz* → *<eu fis> ou mesmo *<eu fix>); em palavras que devem levar <s>, escrevem <z> - se já souberem algo da ortografia das sibilantes (*os pés* → *<os péz>). As crianças não apresentaram nenhuma dificuldade nas realizações orais relacionadas a este tipo de estruturas e, portanto, as FN-Cs podem ser consideradas de natureza essencialmente ortográfica.
- (iii) Em PE, a lateral /l/ em coda normalmente não gera erros no interior da palavra como em *fal~~t~~a* (em PB pode ser substituída por <u> por causa de sua pronúncia [w]). No entanto, em posição final, observam-se FN-Cs com inserção de <e> (note-se que este grafema, por vezes, pode representar na escrita o vazio acústico), tal como observado em (1 i) e (2 i), acima. Nem sempre, no entanto, a inserção do <e> corresponde à inserção da vogal [ɨ].

(3) Estruturas fonológicas com complexidade no núcleo revelam FN-Cs:

- (i) Em estruturas não especificadas no núcleo inicial (núcleo vazio) poderiam ocorrer FN-Cs como *<scova> por escova [ʃkóve] em PE. No entanto, embora as realizações orais de palavras com esta estrutura tenham confirmado a pronúncia padrão do PE (com a vogal não realizada), não foram observadas FN-Cs relacionadas sem a vogal, revelando que esta estrutura não constituía um problema para o grupo examinado, ou seja, as crianças já demonstravam familiaridade com a forma gráfica das palavras começadas por <esC> que são vulgarmente pronunciadas sem vogal inicial. Em PB observa-se com frequência grafias como *<iscova>, revelando, pelo contrário, a maior saliência fónica do núcleo neste tipo de estrutura silábica, podendo o <e> ser representado como <i>, mas nunca omitido.
- (ii) Em núcleos complexos habitualmente produzidos com monotongação (/oU/ → [o]: *tesoura* → *<tesora> [tizóre] (53_SA_2A_CM)) ou em que haja uma forma de pronunciar que seja diferente da que a ortografia contempla (*queijo* → * <caijo> por haver centralização [kėjzu] (48_CZ_2A_CM)). Esses são, assim, casos típicos de FN-Cs relacionadas com a fala, isto é, diferentemente de alguns outros referidos acima, sobre os quais pode-se dizer que apenas indiretamente com ela se relaciona.

Portanto, o que poderíamos pensar ser relacionado com a fala nem sempre é. Consideramos importante procurar seguir o raciocínio da criança e validar os achados com base na fala real da criança e não com base no conhecimento fonológico e ortográfico do adulto (como muitas vezes sucede).

As estruturas fonológicas mais complexas geram muitos erros, mas não constituem sempre razão de dificuldades. Por exemplo, na palavra <bicicleta> a grafia convencional do som [s] é dependente de regra contextual (antes de <e> ou <i>) e o ataque complexo Cl também pode ser alvo de alguma dificuldade para grafar (veja-se *<afelinho> por <aflitinho> e *<gelobo> por <globo>, por exemplo). No entanto, na amostra, a palavra foi grafada corretamente por 8 das 11 crianças que a tinham como palavra-alvo e não-convencionalmente por apenas 2 (*<bissicleta> e *<biciqueta>). O mesmo foi observado para a palavra <dinheiro> sobre a qual poder-se-ia esperar alguma dificuldade

na representação gráfica da nasalidade com <nh> seguido por ditongo <ei> e terminando em <o>. Em PE, na pronúncia da região examinada, a palavra é realizada como [dijéjru]. Apesar, então, de existirem pelo menos três possibilidades para o aparecimento de FN-Cs, a palavra foi grafada por todas as crianças que a tinham como alvo e das 15 ocorrências observadas, apenas as de duas crianças não eram convencionais (*<dilhairo> e *<dinhairo>). Ainda assim, na primeira forma, vemos uma confusão típica entre as grafias de <nh> e <lh> e em ambas, o reflexo da centralização do ditongo /ei/ realizado como [éj], observado nas 32 realizações desta palavra, bem como em outras em que ele aparece. Na palavra <queijo>, novamente se verifica um reflexo das 13 realizações observadas para este ditongo ([kéjzu]). Esta pronúncia, consistente entre as crianças da amostra, não explica assim, a omissão do <i> em *<quejo> ([kézu]) na grafia produzida por duas crianças, as quais exibiram a mesma pronúncia das outras crianças. Esta última grafia seria facilmente entendida como uma grafia decorrente de um reflexo da fala na variedade do PB, por exemplo, cuja pronúncia ([kézu]) é comumente observada em vários dialetos (cf. Cagliari, 1989). Este exemplo mostra como a classificação das FN-C não deve ser feita em abstracto numa língua, mas sim atendendo à variedade linguística falada pela criança.

Apesar de apenas nos termos detido em um conjunto de FN-Cs, mais especificamente no que respeita à complexidade silábica, os dados permitem dizer que em sílabas simples não surgem muitas FN-Cs, por exemplo, em posição final de palavra: <gato> [gátu] (mas não *<gatu>). Isso é interessante porque nos mostra que a posição de fronteira de palavra final, tal como a posição inicial (cf. caso de <escova>), não é tão susceptível de causar erros como as restantes. Por outro lado, este tipo de FN-C é frequentemente relatada em estudos do PB (Zorzi, 1998; Cagliari, 1989, por exemplo). Novas análises serão, portanto, relevantes para maior compreensão das relações entre fala e escrita nas duas variedades.

Também é importante notar que nem tudo o que se esperaria se revela produtivo nas FN-Cs das crianças, porque nem sempre as crianças parecem levar em linha de conta a representação fonológica, levando antes em consideração a representação fonética ou uma representação intermédia fonética-ortográfica.

As observações que foram feitas entre as narrativas orais e escritas das crianças, a partir da tarefa com as imagens temáticas, permitiram uma análise mais consistente das FN-Cs observadas em palavras não-alvo desta tarefa e daquela que foi proposta anteriormente com a *história da bruxinha*. A tabela 5 mostra o total de desvios observados nas duas turmas, nas duas tarefas.

Tabela 5 – Total de ocorrências de desvios das formas convencionais das palavras nas duas turmas

Turma	“Imagens”	“História da Bruxinha”	Total de desvios*
	FN-C	FN-C	
A	395	625	1020
B	402	678	1080
Total	797	1303	2100

* Inclui todos os desvios da forma convencional dentro das palavras. FN-C = formas não-convencionais.

Os resultados apresentados até aqui mostram que a partir de um pequeno conjunto de dados de fala foi possível interpretar uma grande quantidade de FN-Cs que refletiam estratégias das crianças para representar formas gráficas com base na fala, seja em sua face fonética ou fonológica e, ao mesmo tempo, distingui-las de FN-Cs cuja natureza revela o uso de estratégias baseadas unicamente em fatos inerentes ao sistema ortográfico da língua (regras contextuais, irregularidades).

Essas últimas podem, ainda, estar relacionadas a outros fatores tais como complexidade silábica e número de caracteres da palavra, por exemplo. Além disso, as tarefas de produção de narrativas propostas, tanto nas *Imagens* como na *História da Bruxinha*, permitiram corroborar ou generalizar a interpretação de muitas das FN-Cs a palavras que não eram alvos, contrariamente ao que se poderia observar apenas em tarefa de nomeação.

6. Considerações finais

A metodologia utilizada para a recolha permitiu obter grande riqueza lexical - muito maior do que se tivéssemos optado pela técnica de nomeação de palavras. Obtivemos essa diversidade de formas lexicais na escrita e na fala de forma semi-dirigida, uma vez que a escrita foi livre (ainda que tendo como guia uma imagem que permitiu a evocação de um conjunto coerente de palavras) e a descrição oral da imagem foi uma conversa natural entre a criança e a entrevistadora, apoiada pela visualização da mesma imagem.

Os dados demonstram que no processo de aprendizagem do sistema ortográfico é necessário identificar aquelas FN-C motivadas pela fala que devem ser alvo de atenção, tanto para propósitos clínicos como do ensino. Algumas ocorrem com maior frequência do que outras e são mais consistentes em uma ou outra criança, ou são recorrentes em qualquer criança.

A tomada de consciência da correspondência grafema-fonema, como foi visto, não é automática e precisa ser explicitamente aprendida (Morais, 1995; Dehaene, 2012). No entanto, isto não é suficiente para as crianças aprenderem a ortografia convencional das palavras. Muitos fatores irão influenciar a evolução deste novo conhecimento e o seu domínio. Diante de segmentos de grafia irregular ou dependentes do contexto em que aparecem, ou de estruturas silábicas que ainda não dominam, mesmo as mais simples, as crianças tentam representar a grafia dessas formas recorrendo a características da fala e, como é sabido, esta estratégia é típica do início da aprendizagem. Visto assim, a frequência de ocorrência de FN-Cs, que refletem o uso desta estratégia, representa um bom indicador do grau de domínio das crianças quanto à apropriação do sistema ortográfico: de modo geral, quanto maior a sua frequência menos estáveis estão as representações ortográficas *per se*. No entanto, como foi ilustrado com algumas FN-Cs das crianças deste estudo, diferentes fatores, e em diversos graus, podem motivar o aparecimento dessas formas, como a complexidade silábica, as regras contextuais e as irregularidades. Adicionados a esses fatores inerentes ao próprio sistema ortográfico da língua, encontram-se outros igualmente importantes, tais como a familiaridade das crianças com o material impresso e as orientações metodológicas da escola para o ensino da escrita.

Ao mesmo tempo, é importante procurar por indicadores na escrita das crianças que revelam o seu conhecimento ortográfico. Elas sabem que existe uma relação entre fala e escrita, e muitas FN-Cs exibem as estratégias adotadas para resolver os problemas que resultam desta relação, que nem sempre é, como sabemos, transparente.

Dois tipos de FN-Cs, em particular, refletem o conhecimento por parte das crianças da distinção entre fala e escrita: a segmentação não-convencional e a hipercorreção. A primeira, cuja motivação, por vezes, está na interpretação das fronteiras de palavras ortográficas com base na cadeia da fala é um indício de uma fase muito inicial da aprendizagem. A segunda, por outro lado, reflete uma fase mais avançada do conhecimento das relações entre fala e escrita. Ela consiste numa generalização do fato de muitas palavras serem grafadas com uma dada letra mas pronunciadas com o som de outra, como <o> e <e> que podem ser pronunciados como [u] e [i], respetivamente, conduzindo à confusão no uso dessas letras. Evidentemente, outros indícios, como o uso de maiúsculas, de pontuação, e a grafia de unidades irregulares (por exemplo, do fonema /s/ que pode ser representado pelos grafemas <s; ss; c; ç; sc; sç; xc; x ou z> ou do grafema <r>, que pode representar os fonemas /R/ e /r/), são igualmente reveladores do grau de conhecimento do sistema de escrita.

Ao longo deste artigo expusemos as razões que motivam o projeto EFFE no âmbito do NICEF, nomeadamente o fato de diversas características dos processos envolvidos na escrita inicial de crianças ainda não terem sido alvo de trabalhos que exploram algumas das vertentes sobre o tema.

Ainda que numa fase inicial, este projeto pretende ver discutidas algumas questões essenciais entroncadas na revisão bibliográfica apresentadas nas seções 2 e 3.

A base de dados acima descrita permite a comparação das FN-Cs com a produção oral de cada criança. Desta comparação foi possível observar que nem todas as FN-Cs que poderiam parecer relacionadas com a fala na realidade com ela se relacionam. Vimos, por outro lado, que com base em pequenas amostras de fala pode-se obter evidências suficientes para traçar um perfil das dificuldades pelas quais as crianças passam em cada momento da aprendizagem. Assim, o professor poderá identificar estratégias adequadas para a resolução das dificuldades das crianças (cf. por exemplo, Cagliari 1989, Morais, 2007). Exige-se é que os professores saibam reconhecer adequadamente as características da variedade falada na região em que trabalham; caso contrário, poderão adotar estratégias que não se adequam às reais dificuldades das crianças.

Em Portugal ainda existem poucos dados de produção escrita nos primeiros anos de escolaridade acessíveis para a investigação. A base de dados continua a ser ampliada e possibilitará, para além da discussão criteriosa das FN-Cs relacionadas com a fala desta amostra, a descrição e interpretação de outras FN-Cs em amostras que nela serão subsequentemente incorporadas.

São ainda necessários estudos acerca da leitura em voz alta e da compreensão da leitura, associados a dados de produção escrita (em diferentes tarefas).

Referências

Adams, M. J. (1990): *Beginning to read: Thinking and learning about print*. MA: MIT Press, Cambridge.

Alves, D. C. (2012) *Efeito das propriedades segmentais em tarefas de consciência segmental, de leitura e de escrita*. Tese [Doutoramento em Linguística]. Faculdade de Letras, Universidade de Lisboa. Lisboa.

Barret, T. C. (1965): The relationship between measures of pre-reading visual discrimination and first grade reading achievement: a review of the literature, in *Reading Research Quarterly*. 1(1), pp. 51-76.

Bonin, P. & Meót, A. (2002) Writing to dictation in real time in adults: what are the determinants of written latencies. In Shohov, S. P. (Editor) *Advances in Psychology Research*, vol. 16. NY: Nova Science Publishers, p. 139-165.

Berninger, V.; Abbott, R. D.; Jones, J. (2006): Early Development of Language by Hand: Composing, Reading, Listening, and Speaking Connections; Three Letter-Writing Modes; and Fast Mapping in Spelling, in *Developmental Neuropsychology*. 29(1), pp. 61–92.

Cagliari, L. C. (1989): *Alfabetização e Linguística*. Scipione, Rio de Janeiro.

Cagliari, L. C. (1994): Algumas reflexões sobre o início da ortografia da língua portuguesa, in *Cad. Est. Ling.* 27, pp. 103-111.

Caramazza, A.; Miceli, R. (1990): The structure of graphemic representations, in *Cognition*. 37, pp. 243-297.

Castelo, A. (2012) *Competência Metafonológica e Sistema Não Consonântico no Português Europeu: Descrição, Implicações e Aplicações para o Ensino do Português como Língua Materna*. Tese [Doutoramento em Linguística]. Faculdade de Letras, Universidade de Lisboa. Lisboa.

Castro-Caldas, A.; Reis, A. (2003): The knowledge of orthography is a revolution in the brain. *Reading and Writing: An Interdisciplinary Journal*. 16, pp. 81–97.

Cubelli, R. (1991): *A selective deficit for writing vowels in acquired dysgraphia*. *Nature*, 353, pp. 258-260.

Dehaene, S.; Pegado, F.; Braga, L. W. et al. (2010): *How learning to read changes the cortical*

- networks for vision and language*. 330, *Science*. Versão online. Disponível em www.sciencemag.org.
- Dehaene, S. (2012): *Os neurônios da leitura: como a ciência explica a nossa capacidade de ler*. Penso, São Paulo. Tradução de Leonor Scliar-Cabral [do original *Reading in the Brain*, 2009].
- Ellis, A. W. (1988): Normal writing process and peripheral acquired dysgraphias, in *Language and Cognitive Process*. 3(2), pp. 99-127.
- Freitas, M. J.; Rodrigues, C.; Costa, T.; Castelo, A. (2012): *Os sons que estão dentro das palavras*. Cadernos Língua 5, Colibri-APP, Lisboa.
- Freitas, M. J.; Gonçalves, A.; Duarte, I. (coord.) (2010): *Avaliação da consciência linguística: aspectos fonológicos e sintáticos do português*. Colibri, Lisboa.
- Goodman, Y. M. (org.) (1995): Conhecimento das crianças sobre a alfabetização: um pós-fácio, in *Como as crianças constroem a leitura e a escrita: perspectivas piagetianas*. Artes Médicas, Porto Alegre.
- Guerreiro, H. W. Z. M. R. (2007): *Processos fonológicos na fala da criança de cinco anos*. Tese (Mestrado em Ciência da Fala), Instituto de Ciências da Saúde da Universidade Católica Portuguesa-Escola Superior de Saúde do Alcoitão da Santa Casa da Misericórdia de Lisboa, Lisboa.
- Guimarães, M. R. (2005): *Um estudo sobre a aquisição da ortografia nas séries iniciais*. Tese (Mestrado em Educação), Faculdade de Educação da Universidade Católica de Pelotas, Pelotas.
- Guinet, E. & Kandel, S. (2010) Ductus: a software package for the study of handwriting production. *Behavior Research Methods*, 42 (1), 326-332.
- Horta, I. V. (2001): *Pensar a ortografia: s, ss, c ou ç? Competências ortográficas no 3º. e 4º Anos de escolaridade do ensino básico*. Dissertação [Mestrado em Psicologia: Psicologia Educacional], Instituto Superior de Psicologia Aplicada, Lisboa.
- Horta, I. V. & Martins, M. A. (2004): Desenvolvimento e aprendizagem da ortografia: implicações educacionais, in *Análise Psicológica*. 1(XXII), pp. 213-223.
- Kandel, S.; Peereman, R. & Ghimenton (2014) How do we code the letters of a word when we have to write it? Investigating double letter representation in French. *Acta Psychologica*, 148, 56-62.
- Kolinsky, R.; Pattamadilok, C. & Morais, J. (2012): The impact of orthographic knowledge on speech processing, in *Ilha do Desterro*. Florianópolis, 63, pp. 161-186.
- Lourenço-Gomes, M. C. (1999): *Leitura e Escrita e Consciência Fonológica: intervenção em sala de aula*. Monografia (Especialização em Distúrbios da Comunicação Humana). Universidade Federal do Estado de São Paulo/ Universidade Católica de Petrópolis, Petrópolis. 132 p.
- Maslow, P.; Frostig, M.; Lefever, W.; Whittlesey, J. R. B. (1964): The Marianne Frostig Developmental Test of Visual Perception, in *Perceptual and Motor Skills*, 1963 Standardization, 19, pp. 463-499.
- Mateus, M. H. M. (2002/1962): O problema da ortografia. Métodos de ensino, in *A Face Exposta da Língua*. Imprensa Nacional Casa da Moeda, Lisboa, pp. 91-135.
- Mateus, M. H. M.; Brito, A. M.; Duarte, I. et al. (2003): *Gramática da Língua Portuguesa*. 7.^a ed., Caminho, Lisboa.
- Miranda, A. R. M. & Cunha, A. P. N. da (Orgs.) (2010): *Cadernos de Educação: A aquisição e o ensino da língua escrita*. 35.
- Moojen, S. M. P. (2009): *A escrita ortográfica na escola e na clínica: teoria, avaliação e tratamento*. Casa do Psicólogo, São Paulo.
- Morais, A. G. (2007): O diagnóstico como instrumento para o planeamento do ensino de ortografia,

in Silva, A.; Morais, A. G. de; Melo, K. L. R. de (Orgs.) (2007): *Ortografia na sala de aula*. Autêntica, Belo Horizonte, pp. 45-60.

Morais, J. (1995): *A arte de ler*. UNESP, São Paulo.

Morais, J. & Kolinsky, R. (2001): Neurociência cognitiva e Psicolinguística: A linguagem escrita, in Corrêa, L. M. S. Processamento linguístico e aquisição da linguagem, *Palavra*. Nº 6, v. II, Trarepa, Rio de Janeiro, pp. 77-93.

Perfetti, C. A. (1997) The Psycholinguistics of spelling and reading. In Perfetti, C. A.; Rieben, L. & Fayol, M. *Learning to spell: research, theory, and practice*. NJ: Lawrence Erlbaum. p. 21-38

Perrolas, M. P. C. R. (1996): Leitura e escrita no 1º Ano de escolaridade. Dissertação [Mestrado em Psicologia: Psicologia Educacional], Instituto Superior de Psicologia Aplicada, Lisboa.

Pinto, M. G. L. C. (1997): A ortografia e a escrita em crianças portuguesas nos primeiros anos de escolaridade: até que ponto dependem estas habilidades de um bom domínio do oral e de métodos adequados de leitura?, in *Revista da Faculdade de Letras: Línguas e Literatura*. XIV, Porto, pp. 7-58.

Rayner, K.; Reichle, E. D. (2010): Models of the reading process, in *Wiley Interdiscip Rev Cogn Sci*. 1(6), pp. 787–799.

Rayner, K.; Pollatsek, A.; Ashby, J.; Clifton Jr., C. (2012): *Psychology of reading*. 2nd ed., Psychology Press, NY.

Rodrigues, J. F. B. N. D. (2013): *Avaliação do impacto de um programa de intervenção de sensibilização à linguagem escrita*. Dissertação [Mestrado em Psicologia: Psicologia Educacional], Instituto Superior de Psicologia Aplicada, Lisboa.

Rodrigues, C.; Lourenço-Gomes, M. C. & Alves, I. (em prep.) Spelling abilities in 2nd and 4th graders - a longitudinal study in European Portuguese.

Santos, R. N. (2013): *Aquisição de grupos consonânticos e seu impacto nos desempenhos escritos no 1.º ciclo do ensino básico*. Tese (Mestrado em Linguística), Faculdade de Letras da Universidade de Lisboa, Lisboa.

Scliar-Cabral, L. (2003a): *Princípios do sistema alfabético do português do Brasil*. Contexto, São Paulo.

Scliar-Cabral, L. (2003b): *Guia prático de alfabetização*. Contexto, São Paulo.

Shen, X. R., Damian, M. F. & Stadthagen-Gonzalez (2013). Abstract graphemic representations support preparation of handwritten responses. *Journal of Memory and Language*, 68, 69-84.

Silva, A.; Morais, A. G. de; Melo, K. L. R. de (Orgs.) (2007): *Ortografia na sala de aula*. Autêntica, Belo Horizonte.

Sim-Sim, I.; Duarte, I. & Ferraz, M. J. (1997): *A língua materna na educação básica*. Ministério da Educação, DEB, Lisboa.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds). (1998): *Preventing reading difficulties in young children*. National Academy Press, Washington DC.

Sousa, O.; Cardoso, A. (2005): Da língua em funcionamento ao funcionamento da língua, in *Palavra*. 27, pp. 61-69.

Taft, M. (2006): Orthographically influenced abstract phonological representation: evidence from non-rhotic speakers, in *Journal of Psycholinguistic Research*. 35(1), pp. 67-78.

Vellutino, F. R.; Fletcher, J. M.; Snowling, M. J.; Scanlon, D. M. (2004): Specific reading disability (dyslexia): What have we learned in the past four decades?, in *Journal of Child Psychology and Psychiatry*. 45:1, pp. 2-40.

Veloso, J. (2007): Can orthographic knowledge modify phonological knowledge? Data from a study of Portuguese children's syllabifications and their relationship with schooling, in Ilieva-Baltova, P.; Petrova, K. (Eds.) *Psycholinguistic Studies. Papers in Honour of Prof. Dr. Tatiana Slama-Cazacu*. Sofia (Bulgaria): Riva Publishers, pp. 215-236.

Veloso, J. (2010): Primeiras produções escritas e operações metafonológicas explícitas como pistas para a caracterização inferencial do conhecimento fonológico, in *Cadernos de Educação – A aquisição e o ensino da língua escrita*. 35, pp. 19-49.

Wagner, R. K.; Piasta, S. B.; Torgesen, J. K. (2006): Learning to read, in Traxler, M. J.; Gernsbacher, M. A. (Ed.) *Handbook of Psycholinguistics*. 2nd ed., Elsevier, NY.

Yavas, M.; Hernandorena, C. L. M. & Lamprecht, R. R. (1991): *Avaliação fonológica da criança*. Artes Médicas, Porto Alegre.

Ziegler, J. & Ferrand, L. (1998): Orthography shapes the perception of speech: The consistency effect in auditory word recognition, in *Psychonomic Bulletin & Review*. 5 (4), pp. 683-689.

Zunino, D. L. de & Pizani, A. P. de (1995) *A aprendizagem da língua escrita na escola: reflexões sobre a proposta construtivista*. 2ª ed., Artes Médicas, Porto Alegre.

Anexo

Palavras-alvo utilizadas no estudo						
Imagem	SA	CI	CB	FL	CZ	
Fonte						
Guerreiro (2007)	Almofada			Nuvem	Estrela	
	Caixa			Sol	Fogão	
	Dinheiro		Menina	Trovoada	Gelado	
	Desenho	Carro	Espelho	Voar	Comer	
	Lápis	Grande	Escova	Dragão	Rato	
	Jornal	Andar	Torneira	Fogo		
	Tesoura	Bicicleta	Toalha	Floresta		
	Rádio	Estrada	Orelha	Verde		
	Televisão	Futebol	Pescoço	Peixe		
	Quadro		Botão	Borboleta		
	Igreja			Caracol		
	Bruxa			Flor		
	Acrescentadas no 2ºA e 2ºB*	Prateleira	Vermelho	Espuma	Elefante	Queijo
		Voar	Piscina	Transbordar		Parabéns
		Vassoura	Touca	Fechar		
		Telhas	Chave			

* Palavras acrescentadas pelas autoras de modo a adaptar o instrumento o mais possível ao propósito, nomeadamente, através da escolha de vocábulos pertinentes para o presente estudo na variedade em questão (variedade padrão do PE).