

Um gradiente da habilidade com a escrita: cartas privadas no contexto da imigração portuguesa para o Brasil no século XX

A scale for measuring writing skills:
private letters in the context of portuguese
immigration to Brazil in the 20th century

Rafael Rodrigues da Silva Cardoso¹ 

E-mail: rfcardoso86@gmail.com

Célia Regina dos Santos Lopes² 

E-mail: celiar.s.lopes@letras.ufrj.br

¹Instituto Nacional da Propriedade Industrial, Rio de Janeiro, RJ, Brasil.

²Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.

Editores-chefes

Marcus Dores
Célia Lopes

Editores convidados

Huda Santiago
Pedro Daniel Souza

Dossiê

Diálogos entre a
Sócio-História do
Português e a História
Social da Cultura Escrita

Recebido: 07/11/2022

Aceito: 08/09/2023

Como citar:

CARDOSO, Rafael
Rodrigues da Silva;
LOPES, Célia Regina dos
Santos. Um gradiente
da habilidade com a
escrita: cartas privadas
no contexto da imigra-
ção portuguesa para
o Brasil no século XX.
Revista LaborHistórico, v.9,
n.2, e55433, 2023. doi:
[https://doi.org/10.24206/
lh.v9i2.55433](https://doi.org/10.24206/lh.v9i2.55433)

RESUMO

O artigo apresenta uma proposta para a análise do grau de habilidade com a escrita de redatores de 99 cartas privadas escritas no contexto da imigração portuguesa para o Brasil de 1948 a 1960. Baseada em Barbosa (2017) e Santiago (2019), a análise é feita a partir do levantamento de desvios gráficos em relação ao referencial modelar escolhido, separando-se os redatores em um gradiente de habilidade. A proposta permitiu categorizar o *corpus* de cartas em função de critérios objetivos, evidenciando que os remetentes produziram textos mais ou menos influenciados pelas práticas correntes da escrita. Ademais, alguns desvios ocorreram apenas nos remetentes de menor grau de habilidade e são possivelmente mais indicativos de uma baixa habilidade com a escrita do que aqueles observados nos remetentes mais hábeis.

Palavras-chave

Migração Portuguesa. Sociolinguística Histórica. Corpus diacrônico. História Social da Escrita.

ABSTRACT

The article presents a proposal to analyze the writing skills in 99 private letters produced in the context of Portuguese Immigration to Brazil from 1948 to 1960. Following Barbosa (2017) and Santiago (2019), we chose a spelling model and collected, in each letter, deviations from the model, which enabled us to classify the writers in three groups according to their writing skills. It also allowed us to categorize the *corpus* based on objective criteria, proving that the writers produce texts more or less influenced by regular practices of writing. Furthermore, some deviations only appear in the text of low skilled writers, which indicate they may be “heavier” or signify lower skill than those found solely in the text of high skilled writers.

Keywords

Portuguese Migration. Historical Sociolinguistics. Diachronic Database. Social History of Written Culture.

Introdução

Um dos principais dilemas da reconstrução histórica de uma língua, no âmbito da (Socio)linguística Histórica (SH), é a natureza da fonte documental que serve de base para os estudos de mudança linguística. O problema é que, diferentemente das análises baseadas em dados reais de fala, a documentação escrita encontrada nos acervos públicos fornece dados parciais, representativos de grupos alfabetizados da comunidade, predominantemente produzidos por homens de status social médio ou alto, dada a falta secular de acesso à educação formal das mulheres. A parcela social que não escrevia ou cujo material escrito não sobreviveu ao tempo não deixou rastros da língua que utilizavam. Além disso, os materiais escritos do passado são esparsos, fragmentários e pouco representativos das comunidades pretéritas (Conde Silvestre, 2007), o que dificulta a reconstrução do nosso passado linguístico.

Sabe-se que a oralidade do passado era tão multifacetada quanto a de hoje (Romaine, 1982) e que também estava sujeita à variação, mas constatar em que consistia essa variação depende dos materiais disponíveis para pesquisas e de sua devida contextualização sócio-histórica. Assim, os estudos históricos buscam identificar materiais escritos mais propensos a refletir essa variação inerente à oralidade, uma vez que o domínio das práticas da escrita tende a apagá-la.

Diversos estudos têm buscado por fontes de redatores menos hábeis, pois em materiais, menos influenciados pelos modelos convencionais da escrita, emergem mais facilmente aspectos da oralidade. Esses *corpora* ajudariam a preencher lacunas na reconstituição linguística do passado e a ter um quadro um pouco mais completo do funcionamento da língua no tocante à variação, pela maior espontaneidade da

língua em uso. Entretanto, os dados sociais dos autores desses materiais são muitas vezes desconhecidos e de difícil rastreamento, daí surge a necessidade de propor métodos alternativos para esboçar o perfil social dos escreventes a partir da tipificação/caracterização do próprio material de análise, como fizeram Santiago (2012; 2019), Silva (2012), entre outros pesquisadores. A descrição e contextualização histórica desse material pode nos mostrar registros de uma “oralidade histórica” a que Elspass (2012) se refere.

Com tal propósito, o nosso artigo visa a apresentar uma proposta para a análise do grau de habilidade com a escrita de redatores de 99 cartas privadas escritas no contexto da imigração portuguesa para o Brasil de 1948 a 1960.

Com base na proposta de Barbosa e Lima (2019) e Barbosa (2017), coletam-se dessas cartas dados de desvios gráficos em relação a um referencial modelar escolhido: o *Dicionário contemporâneo da língua portuguesa Francisco Julio de Caldas Aulete* (Caldas Aulete, 1948).

Como método de trabalho, os dados de desvios ortográficos encontrados nas cartas são classificados tendo em vista dois aspectos: (i) *escrita fonética*, na forma de índices grafofonéticos; (ii) *escriptualidade*, na forma de inobservância a grafismos normatizados e da grafia irregular para sílabas complexas.

A hipótese é de que a frequência das dimensões analisadas pode ajudar a medir o grau de habilidade de escrita que cada missivista possui, caracterizando-os em um *continuum* para diferenciá-los por um fator social externo não conhecido previamente. A ideia é mostrar a eficácia do protocolo metodológico proposto para comprovar que os textos, em análise, realmente se aproximariam, pelo desconhecimento de práticas da cultura escrita, de uma concepção oral, no que se refere exclusivamente a tal parâmetro. Além disso, pretende-se responder às seguintes questões:

1. É possível encontrar dados empíricos nas cartas produzidas por massas populacionais menos abastadas em contextos de migração que comprovem maior ou menor desconhecimento das práticas da escrita convencionais?
2. O controle dos desvios ortográficos do tipo *escrita fonética* e *escriptualidade* (grafismos) poderia auxiliar na categorização dos redatores de um *corpus* de sincronias passadas para análises sociolinguísticas?

O artigo está organizado da seguinte forma. Primeiro, trazemos uma breve base teórica que fundamenta a discussão com informações adicionais sobre o contexto histórico-social do *corpus* e os aspectos metodológicos utilizados na análise. Após, apresentamos os resultados da quantificação geral e a proposta de gradiência para os redatores alfabetizados. Por fim, temos as considerações finais e as referências.

Aspectos teórico-metodológicos

A sociolinguística Histórica e o contexto histórico-social do corpus: breve enquadramento

A proposta está alinhada à (Socio)linguística Histórica que, nas últimas décadas, se remodelou para atender aos esforços da Linguística Histórica na elucidação dos processos de mudança linguística e da Sociolinguística contemporânea na análise de material escrito do passado. O maior desafio que a disciplina enfrenta é lidar metodologicamente com a *representatividade* da documentação disponível, uma vez que, em razão da escassez de dados e de seu caráter aleatório, não é fácil retratar todos os estratos da sociedade que se pretende estudar. Hernández-Campoy e Schilling (2012), por exemplo, questionam o fato de materiais documentais serem realmente representativos da comunidade estudada, já que as fontes não foram preservadas para este fim. Nessa perspectiva, tem-se procurado caracterizar socialmente o *corpus* acessível em acervos públicos e privados.

Elspass (2012) aponta que, por razões de valorização social, mesmo na época em que as classes de menor prestígio socioeconômico passam a ser alfabetizadas, pouca coisa está disponível nos arquivos públicos. Esse quadro se altera, entretanto, a partir dos séculos XIX e XX com as grandes guerras, por conta dos processos de migração generalizados na Europa e dos fortes deslocamentos populacionais. É nessa época que o volume de cartas produzidas por membros das classes mais baixas aumenta surpreendentemente. Como a distância era capaz de romper os vínculos, as cartas se tornam um fenômeno de massa que espelha o esforço da manutenção desses vínculos (Leite, 2017). Essas cartas privadas podem representar de maneira eficaz as primeiras fontes, na história das línguas do Ocidente, que seriam representativas de uma espécie de “oralidade histórica” da grande maioria da população.

Em termos do contexto histórico, as cartas analisadas no artigo, editadas por Cardoso (2020), se inserem no período da imigração portuguesa para o Brasil no século XX (1948-1960). Tal fase corresponde ao que Baganha (1994) chama de ciclo migratório transatlântico que vai até 1950¹ e tem o Brasil como destino principal do emigrante português: 75,7% entre 1880-1960, segundo Serrão (1982).

Os assuntos tratados nas cartas estão relacionados direta ou indiretamente à imigração de tipo familiar, típica dos anos 1950, e abordam os seguintes temas: preparativos burocráticos para a emigração; idas e vindas à Junta de Emigração; cartas de chamada; falecimentos; partilha de terras em Portugal e manutenção do vínculo de amizade e de laços familiares. No *corpus* analisado, tem-se uma rede de contato de três casais que se instalam no Brasil e que enviam cartas de chamada uns aos outros, além

¹ O ciclo intraeuropeu só ocorre a partir de 1960.

dos amigos dos dois lados do Atlântico². Estimuladas por esse contexto, pessoas de classes menos abastadas e pouco escolarizadas produziram textos escritos que figuram hoje como um rico material para a reconstrução da Sociolinguística Histórica da imigração portuguesa no Brasil a partir de uma *história da língua a partir de baixo*.

Proximidade e distância comunicativa: os aspectos do gradiente de habilidade com a escrita e o controle da escriptualidade e da escrita fonética

Um dos grandes problemas que se apresenta em um *corpus* remanescente de sincronias passadas é a falta de informações sobre os fatores externos envolvidos na história da documentação. Isso demanda uma atenção especial à caracterização do perfil social dos informantes, que é, na maioria das vezes, lacunar, sobretudo no caso de autores não ilustres. Para dar conta desse impasse, investe-se na busca de textos concebidos oralmente, mesmo que tenham sido registrados por meio gráfico, como é o caso das cartas pessoais, por serem produções escritas menos influenciadas pelas práticas da *escrita da distância*, no sentido de Koch e Oesterreicher (2007). Além disso, os redatores podem apresentar graus diferentes de habilidade com a escrita, e essa diferença se materializará em textos de concepção mais ou menos oral.

Segundo Koch e Oesterreicher (2007), a oposição básica entre *oralidade (imediatez comunicativa)* e *escripturalidade (distância comunicativa)* pode ser mais bem compreendida por meio de parâmetros de análise, que permitem situar as manifestações linguísticas no *continuum* conceitual (e de meio) oral-escrito. Analisando produções linguísticas em meio gráfico, por exemplo, há aquelas que, conceitualmente, se situam mais para o oral do que para o escrito, segundo seu caráter comunicativo, como é o caso das cartas particulares.

Tendo em vista os limites do artigo, o protocolo apresentado dará conta apenas de um aspecto utilizado como fator externo de análise: a habilidade com a escrita dos redatores dessas cartas de imigração. Em termos metodológicos, propõe-se o controle de desvios em relação a convenções da escrita, seja pela grafia irregular para sílabas complexas com /r/ e /l/ ou ainda pela transposição para a escrita de aspectos fonéticos como a ditongação *pessôua* por *pessoa*. Analisamos sobretudo o aspecto gráfico do material e não as características físicas atribuídas à exercitação da mão do redator, como o traçado inseguro e irregularidade na *empaginação* (Marquilhas, 2000). A adoção desse critério objetivo se justifica pelo fato de ser preferencialmente quantificável e permitir uma justa comparação entre os remetentes.

A análise dos domínios da *escriptualidade* e da *escrita fonética*, propostos em Barbosa (2017), servirá para caracterizar os escreventes a partir da tipologia dos

² Há referências à figura do “portador”, isto é, aquele que fazia a travessia levando artigos para os parentes e amigos.

caracteres gráficos por eles escritos, em contraste com o referencial modelar a 3ª edição atualizada do Dicionário contemporâneo da língua portuguesa Francisco Julio de Caldas Aulete, de 1948. Para as duas dimensões, os percentuais de frequência apresentados correspondem à quantidade dos dados analisados em relação ao total de palavras das cartas de cada remetente. A hipótese, para os remetentes alfabetizados, é a de que a frequência das dimensões analisadas pode ajudar a medir o grau de habilidade com a escrita que cada missivista possui, caracterizando-os em um *continuum* de habilidade.

Os dados encontrados foram separados por fenômeno e em seguida agrupados em duas categorias: *escriptualidade* e *escrita fonética*.

O domínio da *escriptualidade*

Barbosa (2017) considera a *escriptualidade* como um epifenômeno linguístico que, diferentemente dos índices grafofonéticos, é menos vinculado ao que se pronuncia e mais relacionado a “convencionalismos motivados por tradições culturais”. Ao escrever, o redator deve buscar em sua memória o convencionalismo da época para decidir, por exemplo, entre *paçando* ou *passando*. Barbosa inclui nessa dimensão a grafia de sílabas complexas com /r/ e /l/, a grafiação de dígrafos e de diacríticos, além dos latinismos³. Os grafismos e a representação de sílabas complexas foram agrupados por praticidade para facilitar a identificação da inabilidade, e por oposição aos símbolos gráficos pronunciados na escrita mais fonética, que é de outra dimensão para caracterizar as mãos inábeis.

Os fenômenos no âmbito da *escriptualidade* encontrados no *corpus* em análise são:

- a) **Notação de [ʒ]:** Dados com distintas representações (grafemas <g> e <j>) para [ʒ] em desacordo com o referencial modelar, como em *sugeito* por *sujeito*;
- b) **Abaixamento das vogais altas /u/ e /i/:** *molher* por *mulher* e *foturo* por *futuro* não seriam casos de escrita fonética, pois no português europeu o alteamento de [o] para [u] estava consumado já por volta de 1800 (Teyssier, 1982). O redator entende que o grafema <o> assume o valor de [u] em posição pretônica, porém, não sabe, quando pronuncia um [u] pretônico, em quais palavras deve escrever <u> ou <o>. O abaixamento da vogal /i/, como em *fenalizar* por *fnalizar* segue a mesma lógica: o redator sabe que, por vezes, o grafema <e> pode ser pronunciado como [i], mas não sabe em que palavras isso ocorre;
- c) **Ausência ou presença do grafema <h>:** *hontem* para *ontem*, em que é acrescentado um <h> numa possível analogia com a palavra *hoje*, ou *umanidade*

³ Tendo em vista a datação e origem de nossas cartas, os latinismos não foram observados, mas foram incluídas a notação de [ʒ] e das sibilantes.

por *humanidade*, com a falta do <h>. Trata-se de um verdadeiro grafismo (Barbosa, 2017), pois o <h>, em início de palavra, é um resquício etimológico sem valor fonético e o escriba precisa ter contato regular com textos escritos para memorizar as palavras com o <h>;

- d) Notação de sibilantes:** Ao longo de sua história, a grafia das sibilantes acabou por se descolar de sua realização oral e torná-la mais arbitrária, o que dificulta o seu aprendizado⁴. Não foi à toa, portanto, que a grafia da sibilante em desacordo com o referencial modelar foi o fenômeno mais abrangente do *corpus*: dos mais hábeis aos menos hábeis. A notação de sibilantes também foi, de longe, o fenômeno mais frequente no material e exigiu uma tipificação mais complexa: “<s> por <z>” (*afaseres* por *afazeres*) e “<c> por <ss>”, (*acinar* por *assinar*), entre outros.
- e) Notação de sílabas complexas:** A notação de sílabas com consoantes líquidas /r/ e /l/ foi apontada por Marquilhas (2000) como a característica mais recorrente das mãos inábeis renascentistas. Barbosa (2017) insere esse fenômeno ao lado dos grafismos (*escriptualidade*), por oposição à *escrita fonética*. Nas cartas analisadas, separamos em diferentes processos: epênteses (inserção de vogais para desfazer a sílaba complexa), como *compeletos* por *completos* (Crystal, 1988); metáteses (transposição de certos fonemas na cadeia falada) como em *emtertando* por *entretanto* (Dubois *et al.*, 1999). Foram incluídos nessa categoria dados como *ademira* por *admira*.
- f) Notação de vibrantes:** A representação grafemática irregular do /R/ forte em posição medial com o dígrafo <rr> e em posição inicial com <r> foi considerada como pertencente aos grafismos por razões óbvias: *aborecido* por *aborrecido* e *rregular* por *regular*.
- g) Representações do som [k]:** Foram quantificadas as grafias irregulares para o som [k], oriundas da confusão entre <c>, <q>, <cu> e o dígrafo <qu> na representação de [k]: *quando* por *quando*, *perqam* por *percam*, *cria* por *queria*.
- h) Representação da nasal:** A incompreensão de aspectos convencionais da representação da nasalidade na escrita apresentou variação tanto em posição medial como em *anbos* por *ambos* (contexto de uso de <m> antes de -p- e -b-), como em posição final na escrita dos ditongos nasais átonos escritos com <m> e não com til pelas normas gráficas previstas: *entregavão* por *entregavam*, *vom* por *vão*.
- i) Ausência de nasalidade:** A nasalidade simplesmente podia não ser representada, seja por ausência de diacrítico (*estau* por *estão*), seja pela ausência de <m> ou <n> (*asiosa* por *ansiosa*, *ibarcar* por *embarcar*) (Santiago, 2019).

⁴ Para detalhes sobre esse processo, consultar Teyssier (1982, p. 50-53) e Raposo *et al.* (2013, p. 96).

O domínio da *escrita fonética*: os índices grafofonéticos

Barbosa (2017) classifica como índices grafofonéticos as grafias que apontam para um redator estagnado na fase da escrita fonética quando da aquisição da escrita. Isso ocorre quando a “representação gráfica de sons vocálicos e consonantais busca formas de imitar a pronúncia e tende a se afastar das convenções gráficas” (Barbosa, 2017, p. 24-25).

No material analisado, alguns dos índices grafofonéticos repassam para a escrita o processo de elevação e de recuo do vocalismo átono, fenômeno comum a todas as variedades do português europeu. Foram incluídos nesta categoria alguns fenômenos dialetais da região Norte de Portugal. O principal deles é o betacismo.

- a) **Alteamentos:** Fenômeno mais relevante na categoria escrita fonética que foi abordado tendo em vista as suas consequências para as grafias em desacordo com o referencial modelar: o grafema <o> pode se referir às realizações [o], [ɔ] e [u] e o grafema <e> às realizações [e], [ɛ] e [i]. Trata-se de uma característica marcante do português europeu que, por volta de 1800, já se havia generalizado, tanto no contexto pretônico quanto postônico (Teyssier, 1982). Foram classificados em: “grafia <i>”, como em *ilétrica* por elétrica e *salpicôis* por salpicões; e “grafia <u>”, como em *duente* por doente.
- b) **Betacismo:** É um fenômeno dialetal decorrente da neutralização, na pronúncia nortenha, da fricativa bilabial /b/, realizada como oclusiva [b] ou fricativa (ou espirante) [β] (Cintra, 1983). Origina-se da confusão na escolha dos grafemas e <v>: *botos* por *votos*; *liverdade* por *liberdade*. O betacismo foi incluído entre os índices grafofonéticos por ser uma confusão gráfica que espelha uma característica da pronúncia dos remetentes.
- c) **“S beirão”:** Segundo Cintra (1983), os distritos de Viseu e do Porto, locais de origem da maior parte das missivas, fazem parte dos dialetos setentrionais e, mais especificamente, dos dialetos baixo-minhotos-durienses-beirões, cuja característica principal é a possível ocorrência das fricativas ápico-alveolares na realização dos fonemas /s/ e /z/, mais ou menos palatalizadas, sendo a mais palatalizada conhecida como “s beirão”. Nessas regiões, o sistema de quatro sibilantes do português foi simplificado em apenas duas sibilantes apicoalveolares, uma surda e outra sonora (Raposo *et al.*, 2013). Essa pronúncia é também chamada “pronúncia do s como x ou como j” (Cintra, 1983, p. 142-143) e “pronúncia assobiada” (Raposo *et al.*, 2013, p. 90). Alguns dados evidenciam uma aproximação gráfica desse falar nortenho: *curachão* por *coração* e *degejava* por *desejava*.

Outros processos mais gerais foram considerados índices grafofonéticos porque atuaram, principalmente, na caracterização particular do sistema vocálico do português europeu.

- a) **Aférese:** Queda de um fonema inicial ou supressão da parte inicial (uma ou mais sílabas) de uma palavra (Dubois *et al.*, 1999): *letria* por *aletria* (um tipo de massa alimentícia), possível pronúncia dialetal, e *ficuldade* por *dificuldade*.
- b) **Redução (síncope, apócope):** Em razão da atonicidade das vogais pretônicas no português europeu, alguns dados transferem, para a escrita, essa atonicidade, e outros chegam a evidenciar seu desaparecimento da pronúncia no interior (síncope), *aparceu* por *apareceu*, ou no fim da palavra (apócope), *val* por *vale* (apócope).
- c) **Prótese:** Inserção de um som no início da palavra, como em *arresceba* por *receba*, possivelmente por razões dialetais.
- d) **Vogal temática de verbo:** Vasconcellos (1970) comenta que para a 1ª pessoa do plural nos verbos de primeira conjugação, tanto no Norte quanto no Sul, diz-se, no presente do indicativo, *–emos*, como em *amemos*. Também no pretérito perfeito, o autor aponta a possibilidade da pronúncia da terminação *–emos* na primeira conjugação. Foram incluídos aqui os casos que possivelmente representam graficamente essa pronúncia, rotulados como vogal temática por falta de uma terminologia melhor: *fiqemos* por *ficamos* e *estemos* por *estamos*.
- e) **Centralização:** Transposição, para a escrita, da centralização das vogais anteriores ou da confusão na grafia de certas palavras: *amaricanas* por *americanas* e *sartidão* por *certidão*. Esse fenômeno foi encontrado nas cartas de comércio estudadas por Barbosa (1999), tanto em posição tônica (*tanha* por *tenha*) quanto átona (*parnambuco* por *Pernambuco*).
- f) **Nasalização:** Mesmo sem sabermos ao certo se houve hesitação na escrita ou se foi uma questão de pronúncia, a nasalização foi considerada uma característica fonética: *trantando* por *tratando* e *anrranjou* por *arranjou*.
- g) **Ditongação:** Inserção, na pronúncia, de semivogais em contextos que geram ditongos na grafia de algumas palavras: *bouas* por *boas*. Mais da metade desses dados referem-se à inserção da semivogal [y] antes da consoante [ʒ]: *veijo* por *veja*.
- h) **Monotongação:** Redução de ditongos: *troxesse* por *trouxesse*.
- i) **Representação da líquida [λ]:** Algumas ocorrências refletem uma possível “despalatalização” na pronúncia: *le* por *lhe*.

Resultados gerais por tipo de desvios: escriptualidade e escrita fonética

Descrição do método de análise e resultados brutos

Para caracterizar os autores das cartas do *corpus* tendo em vista a proposta de controle do grau de habilidade com a escrita, selecionamos 99 das 154 cartas editadas por Cardoso (2020), procurando configurar uma amostra mais homogênea e equilibrada. Assim, chegamos a 19 remetentes distintos que produziram cartas entre 1948 e 1960, obtendo 2.129 dados de desvios ortográficos.

A **Tabela 1**, na página seguinte, apresenta um panorama geral do levantamento feito nas cartas com os números brutos da quantificação de todos os dados obtidos. No topo da tabela, tem-se a lista dos 19 remetentes que foram anonimizados com as letras iniciais dos seus nomes por ordem alfabética; à esquerda, estão as categorias controladas (*escriptualidade* e *escrita fonética*) com a subdivisão dos fenômenos incluídos em cada uma delas. Na coluna da direita, há o número de dados referente a cada categoria e a cada fenômeno.

Em uma primeira análise bruta do material, observamos que os dados referentes aos fenômenos pertencentes à dimensão da *escriptualidade* (i.e., a *grafismos* e grafia para sílaba complexa) são numericamente mais significativos (1.422 ocorrências – 67%) do que aqueles incluídos na *escrita fonética* (i.e., índices *grafofonéticos*) com 707 dados (33%). Excetuando o remetente MDP, todos os remetentes apresentam dados nessa categoria, como se vê na **Tabela 1**.

O **Gráfico 1** apresenta a proporção de cada fenômeno no domínio da *escriptualidade* em relação ao total de dados. A notação da *sibilante* foi o fenômeno com o maior índice percentual nessa categoria. O remetente MDP é o único que não apresenta nenhum desvio desse fenômeno. A *troca de <q> por <g>* só ocorreu em MA.

O **Gráfico 2** mostra as frequências dos fenômenos de *escrita fonética* em relação ao total de dados. Nesse caso, foi o *alteamento* das vogais (289 dados - 13,5%), seguido pelo *betacismo* (213 dados - 9.9%), que tiveram os maiores índices dos desvios levantados nas cartas. Novamente, o remetente MDP é o único que não apresenta nenhuma ocorrência de alteamento. A *nasalização* só apareceu em CAR e MMJ.

Uma análise bruta dos desvios e da própria quantidade de fenômenos encontrados pode apontar tendências gerais do material. Não é estranho que tenham sido encontradas ocorrências categorizadas como *escriptualidade* em todos os remetentes, independentemente do grau de habilidade, já que nessa categoria foram incluídos os fenômenos relacionados a aspectos convencionais da escrita. Como afirma Lemle (1999, p. 31), no caso de algumas convenções da escrita, até mesmo pessoas cujo grau de habilidade é alto podem estar sujeitas a esse tipo de desvio, já que “quando duas ou mais letras rivalizam na simbolização de um mesmo som na mesma posição, a decisão por um ou outro é, do ponto de vista fonológico, arbitrária”. O *betacismo*,

Tabela 1 – Desvios levantados por categoria e autoria nas cartas de imigração portuguesa: todos os remetentes

Categoria e fenômeno	Remetentes																	Todos os remetentes			
	AB	AD	ADR	AM	ANT	CAM	CAR	DIA	DIA	FER	GEO	LC	MA	MAN	MDP	MGL	ML		MMJ	PAL	QU
<i>Escreptualidade</i>	19	57	8	11	4	17	331	22	22	193	82	7	95	39	1	266	9	212	32	17	1422
Sibilante	19	39	3	5	2	10	142	12	12	91	49	2	49	20	-	144	4	99	10	14	714
Nasalidade não padrão	-	3	4	1	1	1	57	1	1	27	3	3	5	5	1	31	-	64	5	-	212
Abaixamento	-	3	-	-	-	-	68	3	3	33	14	-	14	2	-	37	1	13	9	2	199
Letra H	-	4	-	3	-	-	29	1	1	5	10	-	4	4	-	16	-	3	2	1	82
Sílaba complexa	-	3	-	2	-	3	14	-	12	3	3	1	8	3	-	15	1	7	1	-	73
Ausência de nasal	-	2	1	-	-	-	8	1	16	16	-	-	4	3	-	9	1	13	2	-	60
Som [k]	-	-	-	-	-	-	6	3	6	6	2	-	2	1	-	11	-	13	-	-	44
Repr. de [ʒ]	-	-	-	-	-	2	5	-	2	2	-	-	1	1	-	3	-	-	3	-	17
Vibrante	-	2	-	-	1	1	2	1	1	1	1	-	5	-	-	-	-	-	-	-	14
Posteriorização	-	1	-	-	-	-	-	-	-	-	-	1	1	-	-	-	-	-	-	-	3
Anteriorização	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	2
Troca de q por g	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	2
<i>Escrita fonética</i>	2	8	2	2	5	6	168	21	21	84	18	3	67	22	-	138	3	126	28	4	707
Alteamento	2	3	1	2	3	3	46	5	5	34	7	1	8	12	-	60	1	97	2	2	289
Betacismo	-	1	1	-	1	1	72	2	2	38	3	-	20	3	-	40	-	8	23	-	213
Redução	-	1	-	-	-	1	11	-	1	1	1	-	8	1	-	13	1	6	1	-	45
Vogal temática	-	-	-	-	-	-	20	-	-	-	-	2	1	3	-	15	1	1	1	-	44
Anteriorização	-	1	-	-	-	1	4	14	3	3	1	-	13	-	-	2	-	1	-	-	40
Ditongação	-	2	-	-	1	-	3	-	1	6	6	-	3	2	-	3	-	2	1	1	25
Monotongação	-	-	-	-	-	-	3	-	4	4	-	-	-	1	-	3	-	1	-	-	12
Centralização	-	-	-	-	-	-	2	-	2	2	-	-	6	-	-	-	-	-	-	-	10
Líquida	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	8	-	-	9
Aférese	-	-	-	-	-	-	2	-	-	-	-	-	4	-	-	-	-	-	-	-	6
Prótese	-	-	-	-	-	-	3	-	1	1	-	-	1	-	-	-	-	-	-	1	6
S beirão	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	2	-	1	-	-	5
Nasalização	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	1	-	-	2
Total Geral	21	65	10	13	9	23	499	43	277	100	10	162	61	1	404	12	338	60	21	2129	

Fonte: Cardoso (2020, p. 118).

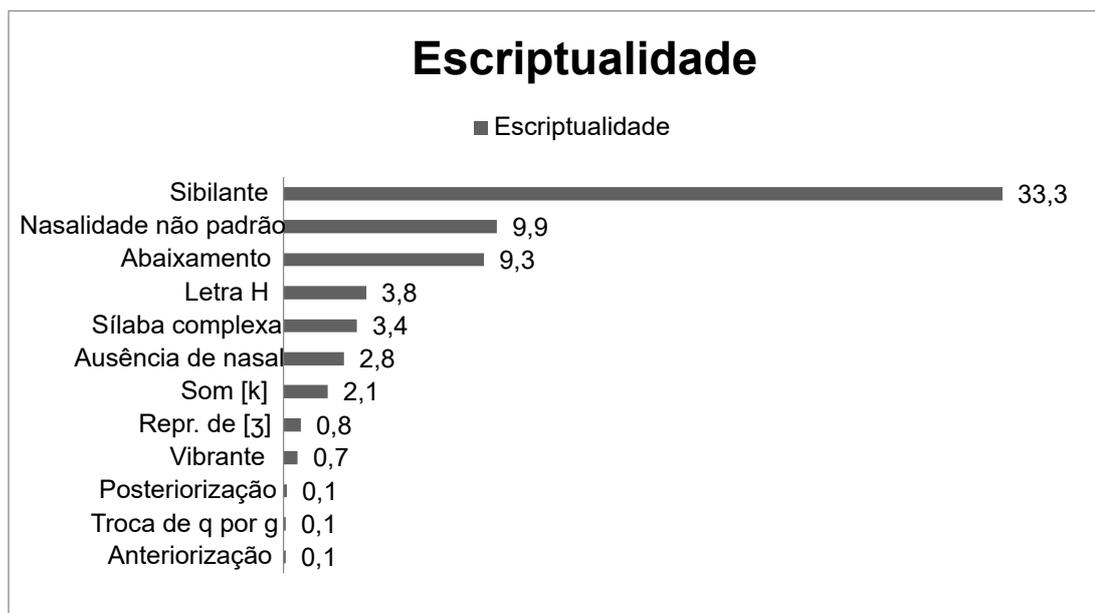


Gráfico 1 – Percentual de fenômenos da categoria esriptualidade.
Fonte: Cardoso (2020, p. 120).

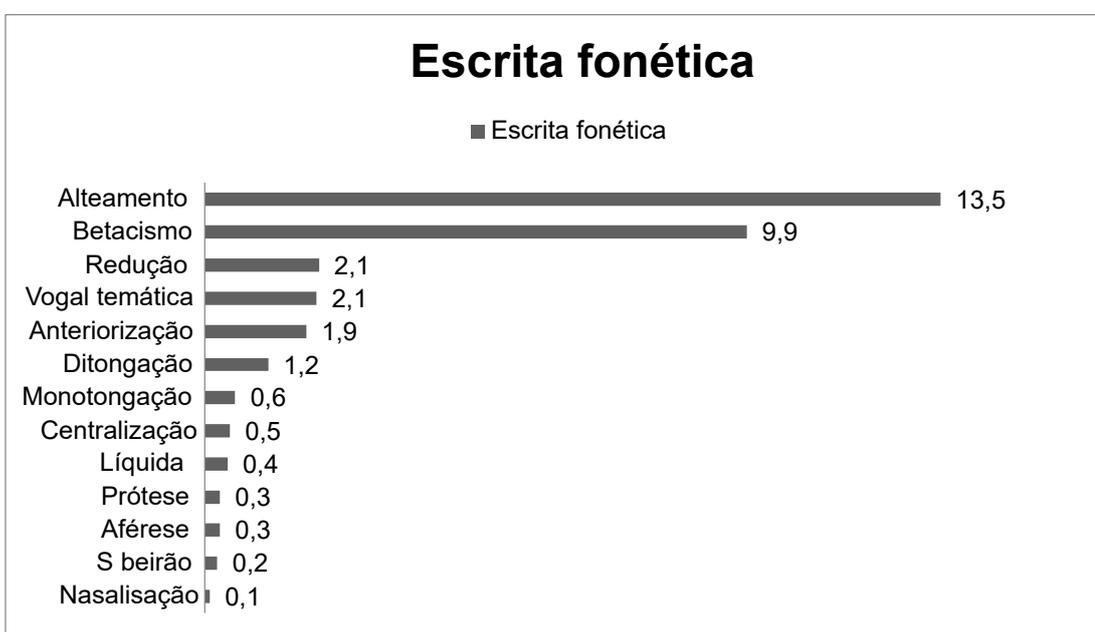


Gráfico 2 – Percentual de fenômenos da categoria escrita fonética.
Fonte: Cardoso (2020, p. 121).

conhecida característica fonética dialetal do Norte, também foi percentualmente significativo possivelmente por termos cartas escritas em grande parte por portugueses dessa região.

Além de observações gerais, é fundamental realizar uma análise mais qualitativa, sobretudo porque a proposta prevê estabelecer com precisão o grau de habilidade dos remetentes com a escrita. Ademais, adotamos um método de nivelamento dos números brutos, já que alguns escreveram mais palavras do que outros. A título de exemplo, enquanto há sete cartas de AB, de MDP há apenas uma. Inspirados em trabalhos anteriores (Barbosa, 2017; Santiago, 2019), propomos, na próxima seção, uma análise dos resultados nas duas categorias propostas (*escriptualidade e escrita fonética*) para caracterizar os remetentes alfabetizados⁵ em um *continuum* de habilidade com a escrita. O objetivo é postular um fator externo para diferenciar os remetentes, caracterizando o material como uma fonte pertinente por suas marcas de oralidade em termos conceituais (Koch; Oesterreicher, 2007).

Proposta de gradiente de habilidade dos remetentes alfabetizados

Para formulação de um *continuum* de habilidade dos 15 remetentes alfabetizados, apresentado na [Tabela 2](#), levamos em conta a frequência de todos os fenômenos quantificados em relação ao total de palavras de cada remetente. Em “total”, tem-se a soma das frequências dos dois aspectos quantificados (*escriptualidade e escrita fonética*) para hierarquizar os remetentes por grupo. A hipótese era a de que uma menor frequência de desvios caracterizaria a maior habilidade de escrita de um informante/grupo, ao passo que frequências altas sinalizariam que o remetente teria menor habilidade com modelos de escrita. Os remetentes foram comparados especificamente no que se refere à frequência de desvios, para constituir uma gradiência interna ao *corpus*.

Os remetentes foram separados em três grupos de habilidade dentro do universo do *corpus*. O grupo 1 é constituído pelos cinco remetentes de maior habilidade, que se encontram no topo da gradiência por apresentar os menores percentuais de desvios. O grupo 2 reúne os cinco remetentes de habilidade média. O grupo 3 contém os cinco remetentes que apresentaram maior percentual de desvios. Essa separação objetiva apenas uma organização dos remetentes para facilitar a análise dos dados. A [Tabela 2](#), a seguir, apresenta o resultado da proposta de gradiência.

⁵ Não serão considerados neste artigo os quatro missivistas que, se, segundo os depoimentos de familiares, eram sabidamente analfabetos e ditaram suas cartas (FER, MA, MAN e MG). Para mais informações consultar Cardoso (2020).

Tabela 2 – Os remetentes por grupos de habilidade com a escrita.

Grupo	Remetente	Esriptualidade	Escrita fonética	Total	Grau
Grupo 1 Remetentes mais hábeis	MDP	0.16%	–	0.16%	
	CAM	0.32%	0.11%	0.43%	
	LC	0.48%	0.21%	0.69%	
	AB	1.15%	0.12%	1.27%	
	AM	1.17%	0.21%	1.38%	
Grupo 2 Remetentes de habilidade média	ADR	1.84%	0.46%	2.30%	
	QU	2.88%	0.72%	3.60%	
	ML	3.32%	1.11%	4.43%	
	AD	4.50%	0.63%	5.13%	
	GEO	4.33%	0.95%	5.28%	
Grupo 3 Remetentes menos hábeis	PAL	3.05%	2.67%	5.73%	
	CAR	4.19%	2.11%	6.30%	
	DIA	4.09%	3.58%	7.67%	
	ANT	5.48%	6.85%	12.33%	
	MMJ	12.45%	7.40%	19.85%	

Fonte: Cardoso (2020, p. 125).

Na sequência apresentam-se os traços de cada grupo de remetentes postulado tendo em vista as diferenças no tocante aos fenômenos, tanto em termos de quantificação quanto de variedade nos tipos de desvios.

Grupo 1: remetentes mais hábeis

O grupo 1 agrega redatores com o menor percentual nos fenômenos elencados. Na esteira de Barbosa (2005), toma-se como referência os remetentes mais hábeis, que servirão como redatores balizares/modelo. Os remetentes desse grupo apresentaram frequências baixas de desvios (menos de 2%). Vejamos:

Tabela 3 – Desvios encontrados no Grupo 1.

Categoria	Fenômeno	Remetentes / nº de palavras					Grupo 1 (9993)
		AB (1653)	AM (940)	CAM (5311)	LC (1457)	MDP (632)	
<i>Esriptualidade</i>	Sibilante	19	5	10	2	–	36
	Nasalidade não padrão	–	1	1	3	1	6
	Sílaba complexa	–	2	3	1	–	6
	Letra H	–	3	–	–	–	3
	Repr. de [3]	–	–	2	–	–	2
	Vibrante	–	–	1	–	–	1
	Posteriorização	–	–	–	1	–	1
	Esriptualidade (todos)	19 (1.15%)	11 (1.17%)	17 (0.32%)	7 (0.48%)	1 (0.16%)	55 (0.55%)
<i>Escrita fonética</i>	Alteamento	2	2	3	1	–	8
	Vogal temática do verbo	–	–	–	2	–	2
	Betacismo	–	–	1	–	–	1
	Redução	–	–	1	–	–	1
	Anteriorização	–	–	1	–	–	1
	Escrita fonética (todos)	2 (0.12%)	2 (0.21%)	6 (0.11%)	3 (0.21%)	–	13 (0.13%)
Total Geral		21 (1.27%)	13 (1.38%)	23 (0.43%)	10 (0.69%)	1 (0.16%)	68 (0.68%)

Fonte: Cardoso (2020, p. 127).

A **Tabela 3** com os dados dos mais hábeis do *corpus* mostra que a presença dos desvios é bastante baixa, tanto em termos do número de ocorrências (de 1 a 5 dados, com exceção das *sibilantes*) como em relação aos percentuais de frequência (menos de 0.50% em sua maioria). As ocorrências são bastante esporádicas, com os supostos desvios ocorrendo em um ou outro remetente. Se levarmos em conta, por exemplo, o número de ocorrências identificadas nas cartas desses redatores (68 dados) em relação ao número total de palavras escritas (9993), constata-se que esse conjunto apresenta um baixíssimo índice de desvios (0,68% - abaixo de 1%), o que qualifica esses redatores do grupo como bastante hábeis.

O fenômeno *escriptualidade* que sobressaiu no grupo foi a representação das sibilantes. Foram identificados 36 dados: AB teve 19 dados; CAM (um procurador), 10 ocorrências; AM, 5 dados; e LC apenas 2.

A representação gráfica oscilante das consoantes sibilantes não costuma ser usada como parâmetro de habilidade dos redatores, porque aparece com frequência na escrita de escritores hábeis e inábeis. Na escrita dos remetentes menos hábeis, a flutuação se revelará bem mais expressiva, do ponto de vista percentual, do que percebemos nas cartas do grupo 1. Os dados referentes às sibilantes foram separados nos seguintes subtipos, como se vê na **Tabela 4**:

Tabela 4 – Tipos de desvios no fenômeno sibilante (Grupo 1).

Tipo de sibilante	Remetente				Grupo 1
	AB	AM	CAM	LC	
<s> por <z>	18	1	5	1	25
<z> por <s>	–	2	2	1	5
<s> por <c>	–	1	1	–	2
<ç> por <s>	–	–	1	–	1
<ç> por <ss>	–	1	–	–	1
<ch> por <x>	–	–	1	–	1
<c> por <s>	1	–	–	–	1
Total Geral	19	5	10	2	36

Fonte: Cardoso (2020, p. 129).

A maioria dos dados (25/36) se limita à representação do som /z/, como em: *atrazo* por *atraso*, *autorisando* por *autorizando*, *certesa* por *certeza*, *despresarem* por *desprezarem*, *diser* por *dizer*, *estranhesa* por *estranheza*, *faser* por *fazer*, *gosando* por *gozando*, *presadíssimo* por *prezadíssimo*, *praso* por *prazo*, *prejuiso* por *prejuízo*, *satisfaser* por *satisfazer*, *veses* por *vezes*. A maioria desses tipos ocorreu apenas uma vez, com exceção de *presado* por *prezado* que apareceu quatro vezes nas cartas de AB. Há outros casos menos frequentes para representação de /s/ com os grafemas <c>, <ç>, <s>: *sisma* por *cisma*, *preça* por *pressa*, *pretenções* por *pretensões*. Nesse grupo, é o remetente AB que apresenta uma maior frequência de desvio na grafia sibilantes: 18 em 19 dados se referem à troca de <s> por <z>.

Com relação aos índices grafonômicos, a frequência maior está no alteamento das vogais, que ocorrem predominantemente nas cartas de CAM (0.11) para representar a vogal /e/: *inferna* por *enferma*, *mâi* por *mãe*. Nas cartas de MDP não foram encontrados dados desse tipo, como consta da [Tabela 5](#).

Tabela 5 – Tipos de desvios no fenômeno alteamento (Grupo 1).

Tipos de Alteamento	Remetente				Grupo 1
	AB	AM	CAM	LC	
vogal /o/	1	–	1	1	3
vogal /e/	1	2	2	–	5
Total Geral	2	2	3	1	8

Fonte: Cardoso (2020, p. 130)

Analisando contrastivamente esses cinco remetentes, nota-se uma diferença de comportamento no que se refere aos fenômenos que aparecem em cada redator. Enquanto nas cartas do nosso redator-modelo CAM, há uma diversidade maior de quatro aspectos grafonômicos e cinco tipos diferentes de grafismos; nas cartas de AB, presumidamente menos hábil, os dados se concentram em dois aspectos: representação das sibilantes e alteamento vocálico. Isso se deve possivelmente à diferença entre o número de palavras escritas por AB (1.653) e CAM (5.311).

A escrita de CAM contém ainda (e esta é uma característica rara no material), a prática recorrente de abreviaturas (*m/* por *meu*; *n/* por *nossa*, *p.p* por próximo passado; *p.f.* por *próximo futuro*) e duas ocorrências de grafia etimologizante “ótima” (CAM01 e CAM04)⁶. Destaca-se a presença de uma grafia abolida antes do acordo ortográfico de 1911: o ditongo “ae” em *comerciaes* por *comerciais*, *constitue* por *constitui* e *possue* por *possui*. Mais uma vez, tem-se uma grafia que reitera um perfil com fortes traços de erudição.

Em síntese, podemos afirmar que esses remetentes, com percentual de desvios baixo entre os remetentes alfabetizados, são os que conhecem melhor o referencial modelar escolhido, pois quanto menor o percentual de desvios em relação ao total de palavras maior a habilidade com a escrita. Pelo menos quanto a CAM, outras características apresentadas também nos levam a concluir que se trata de um remetente bastante hábil. Passaremos, em seguida, à análise do grupo de remetentes de habilidade média.

⁶ Pode-se opor a esta grafia, forte indício de habilidade, a grafia “ótimamente”, presente na carta AM02.

Grupo 2: remetentes de habilidade média

Como remetentes de habilidade média foram considerados aqueles que apresentaram percentuais entre 2.30% e 5.28%. A **Tabela 6** apresenta os resultados de cada um. Sublinhamos os fenômenos que ocorrem apenas nesse grupo e não no anterior:

Tabela 6 – Desvios encontrados no Grupo 2.

Categoria	Fenômeno	Remetentes					Grupo 2 (4422)
		AD (1266)	ADR (435)	GEO (1894)	ML (271)	QU (556)	
Esriptualidade	Sibilante	39	3	49	4	13	108
	Abaixamento	3	–	14	1	2	20
	Letra H	4	–	10	–	1	15
	Nasalidade não padrão	3	4	3	–	–	10
	Sílaba complexa	3	–	3	1	–	7
	Ausência de nasal	2	1	–	1	–	4
	Vibrante	2	–	1	–	–	3
	Anteriorização	–	–	–	2	–	2
	Som [k]	–	–	2	–	–	2
	Posteriorização	1	–	–	–	–	1
Esriptualidade (todos)	57 (4.50%)	8 (1.84%)	82 (4.32%)	9 (3.32%)	16 (2.87%)	172 (3.89%)	
Escrita fonética	Alteamento	3	1	7	1	2	14
	Ditongação	2	–	6	–	1	9
	Betacismo	1	1	3	–	–	5
	Redução	1	–	1	1	–	3
	Anteriorização	1	–	1	–	–	2
	Prótese	–	–	–	–	1	1
	Vogal temática do verbo	–	–	–	1	–	1
Escrita fonética (todos)	8 (0.63%)	2 (0.46%)	18 (0.95%)	3 (1.10%)	4 (0.72%)	35 (0.79%)	
Total Geral	65 (5.13%)	10 (2.30%)	100 (5.27%)	12 (4.42%)	20 (3.59%)	207 (4.68%)	

Fonte: Cardoso (2020, p. 133).

Nos resultados da tabela do grupo médio, as categorias *escriptualidade* e *escrita fonética* tiveram não só um percentual mais alto do que o grupo 1 (3.89% e 0.79%, respectivamente, contra 0.55% e 0.13%), mas também maior número de dados, embora os remetentes do grupo 2 tenham escrito menos que os do grupo 1 (9.993 palavras para o grupo 1 e 4.422 palavras para o grupo 2).

As diferenças nos dois grupos ocorreram também na diversidade dos fenômenos observados, o que ratifica a menor habilidade por parte dos remetentes do grupo 2 em relação ao 1. Na categoria *escriptualidade*, além dos fenômenos observados no grupo 1, identificamos mais três aspectos: *abaixamento*, *ausência de nasal* e o *som [k]*. As sibilantes foram novamente os desvios mais produtivos. O *abaixamento* foi o segundo mais frequente na categoria *escriptualidade* (20 ocorrências). O fenômeno foi frequente nas cartas da remetente GEO (14 /1894 palavras - 0.73%).

Na categoria *escrita fonética*, aparecem fenômenos como a *ditongação* e a *prótese* que não ocorreram entre os mais hábeis.

A **Tabela 7** apresenta os resultados da representação das *sibilantes*. Houve um aumento da frequência (108/4.422 palavras - 2.44%, em 2, contra 83/9993 - 0,83%, em 1) e maior diversidade de tipos: 13 tipos de desvios diferentes no grupo 2 contra 7 tipos no grupo 1 dos mais hábeis.

Tabela 7 – Tipos de desvio no fenômeno sibilante (Grupo 2).

Tipo de Sibilante	Remetentes					Total
	AD	ADR	GEO	ML	QU	
<s> por <z>	22	2	37	1	7	69
<ch> por <x>	2	–	4	–	–	6
<c> por <s>	6	–	–	–	–	6
<s> por <ss>	3	–	–	2	1	6
<s> por <c>	1	–	4	–	–	5
<c> por <ss>	1	–	–	–	2	3
<ss> por <ç>	–	1	–	1	1	3
<ç> por <s>	1	–	–	–	2	3
<ss> por <c>	1	–	–	–	1	2
<ç> por <ss>	2	–	1	–	–	3
<s> por <sc>	–	–	1	–	–	1
<z> por <s>	–	–	1	–	–	1
<ss> por <s>	–	–	1	–	–	1
Total Geral	39	3	49	4	14	109

Fonte: Cardoso (2020, p. 134).

O tipo mais produtivo de desvio na representação das sibilantes foi novamente em <s> por <z>, mas aparecem outros tipos menos frequentes como: <ch> por <x>, <c> por <s>, <s> por <ss> e <s> por <c>.

Nos dados de <s> por <ss>, tipo de *sibilante* inexistente para o grupo 1, os remetentes AD e QU apresentaram desvios na representação do sufixo de pretérito imperfeito do subjuntivo: *conta-se* por *contasse* e *escreve-se* por *escrevesse* (remetente AD); e *fica se* por *ficasse* (remetente QU). A remetente ML, porém, apresentou, em sua escrita, outros indícios: *empregase* por *empregasse* e *passagem* por *passagem*, o que evidencia uma incompreensão do valor do grafema <s> em contexto intervocálico, que não é [s] e sim [z].

A **Tabela 8** mostra que a *nasalidade não padrão* foi identificada em 3 dos 5 remetentes do grupo 2. Os dados foram: <ão> por <am>, em *acharão* por *acharam*, *aumentarão* por *aumentaram*, *ficarão* por *ficaram* e *tenhão* por *tenham*. Apesar de terem escrito menos, os remetentes do grupo 2 apresentaram mais dados de *nasalidade não padrão* do que o grupo 1, confirmando o seu menor grau de habilidade com a escrita.

Tabela 8 – Tipos de desvio no fenômeno nasalidade não padrão (Grupo 2).

Tipos de nasalidade não padrão	Remetentes			
	AD	ADR	GEO	Grupo 2
<ão> por <am>	2	–	2	4
<m> por <n>	–	3	–	3
<n> por <m>	1	1	1	3
Total Geral	3	4	3	10

Fonte: Cardoso (2020, p. 135).

Na categoria *escrita fonética*, foram relevantes os fenômenos do *alteamento* e da *ditongação*. O primeiro, já presente no grupo 1, quadriplicou sua frequência em relação ao grupo anterior: 14/4422 palavras - 0.32% contra 8/9.993 - 0.08%. Todos os remetentes do grupo 2 apresentam ocorrências de *alteamento*, como consta da **Tabela 9**.

Tabela 9 – Tipos de desvio encontrados no fenômeno alteamento (Grupo 2).

Tipo de alteamento	Remetentes					Total
	AD	ADR	GEO	ML	QU	
vogal /e/	2	1	2	1	2	8
vogal /o/	1	–	5	–	–	6
Total Geral	3	1	7	1	2	14

Fonte: Cardoso (2020, p. 136).

A *ditongação*, que estava ausente do grupo 1, aparece aqui em três remetentes: AD, GEO e QU. A remetente GEO tem como fenômenos principais o *abaixamento* e a *letra h*. Contudo, ao analisar seus 14 dados de abaixamento, vemos que 3 deles se referem à grafia *emenso* por *imenso* e outros 3 se referem à grafia *comprimento* por *cumprimento*. Ao analisar os dados de *letra h*, foram 3 ocorrências de *hainda* por *ainda* e 2 da grafia *há* por *à*. Talvez uma análise que levasse em conta esse fator permitiria a GEO uma posição diferente da qual lhe atribuímos aqui.

Grupo 3: remetentes menos hábeis

O grupo dos menos hábeis teve os percentuais mais altos tanto para desvios no domínio da *escriptualidade* (5.73%) quanto da *escrita fonética* (3.06%). A Tabela 10 apresenta a diversidade de fenômenos. Em relação ao grupo 1, foram identificados dois novos fenômenos de *escriptualidade* (*som [k]* e *abaixamento*) e seis para *escrita fonética* (*aférese*, *anteriorização*, *centralização*, *ditongação*, *líquida*, *monotongação*, *nasalização*, *prótese* e *s beirão*). Em relação ao grupo 2, houve um fenômeno a mais: *representação de [ʒ]*.

Os fenômenos mais frequentes do grupo 3 na categoria *escriptualidade* foram, como ocorreu no grupo 1, a grafia das *sibilantes* e a *nasalidade não padrão*. O terceiro fenômeno mais frequente (*abaixamento*) estava, entretanto, ausente no grupo 1, embora tenha aparecido entre os de habilidade média.

Na análise do fenômeno *sibilantes* nesse grupo 3 (menos hábeis), vide Tabela 11, percebe-se uma maior variedade de tipos, além das frequências altas. Foram 265 ocorrências do que estamos rotulando de *sibilantes*, o que representa 2.33% de todas as palavras escritas pelos remetentes do grupo, o que contrasta, por exemplo, com as 36 ocorrências (0.03%) identificadas entre os remetentes do grupo 1, e 108 no grupo 2 (2.44%). Para a *representação da sibilante*, embora não se note aumento percentual em relação ao grupo 2, percebe-se uma clara progressão na variedade de tipos: 7 para o grupo 1, 14 tipos para o grupo 2 e agora são 21 desvios diferentes no grupo 3:

Com exceção de ANT, todos os missivistas incorrem em, no mínimo, 7 tipos diferentes de desvios referentes a *sibilantes*. O remetente MMJ, por exemplo, apresenta um alto percentual de troca de <s> por <ss> em diferentes itens lexicais: *noso* por *nosso* (4 vezes), *nosa* por *nossa* (4 vezes), *nosos* por *nosso*s (3 vezes), *pasando* por *passando* (4 vezes), *masos* por *maços*, *iso* por *isso* (4 vezes) e *esquesendo* por *esquecendo*. Tal diversidade e regularidade reitera a incompreensão da variação contextual do grafema <s>, que, em contexto intervocálico, não tem o som de [s], mas é assim usado pelo remetente. Quanto à remetente CAR, houve sistematicamente a troca do <c> por <s> em contexto inicial ou intervocálico em itens que se repetem nas cartas: *siente* por *ciente* (4 vezes) e nas variações do verbo “receber” (*reseve* por *recebe*, *resever* por *receber*, *reseveu* por *recebeu*, *resevi* por *recebi*)⁷.

⁷ Não se pode afirmar com propriedade que haja comportamentos sistemáticos dos outros remetentes, já que ANT e DIA só têm uma carta e PAL tem apenas duas.

Tabela 10 – Desvios encontrados no Grupo 3.

Categoria	Fenômeno	Remetentes/ nº de palavras					Grupo 3 (11365)
		ANT (73)	CAR (7954)	DIA (587)	MMJ (1703)	PAL (1048)	
<i>Esriptualidade</i>	Sibilante	2	142	12	99	10	265
	Nasalidade não padrão	1	57	1	64	5	128
	Abaixamento	–	68	3	13	9	93
	Letra H	–	29	1	3	2	35
	Ausência de nasal	–	8	1	13	2	24
	Sílaba complexa	–	14	–	7	1	22
	Som [k]	–	6	3	13	–	22
	Representação de [ʒ]	–	5	–	–	3	8
	Vibrante	1	2	1	–	–	4
	Esriptualidade (todos)	4 (5.48%)	331 (4.19%)	22 (4.09%)	212 (12.45%)	32 (3.05%)	601 (5.28%)
<i>Escrita fonética</i>	Alteamento	3	46	5	97	2	153
	Betacismo	1	72	2	8	23	106
	Vogal temática do verbo	–	20	–	1	1	22
	Anteriorização	–	4	14	1	–	19
	Redução	–	11	–	6	1	18
	Líquida	–	1	–	8	–	9
	Ditongação	1	3	–	2	1	7
	Monotongação	–	3	–	1	–	4
	Prótese	–	3	–	–	–	3
	Centralização	–	2	–	–	–	2
	Aférese	–	2	–	–	–	2
	Nasalisação	–	1	–	1	–	2
	S beirão	–	–	–	1	–	1
	Escrita fonética (todos)	5 (6.85%)	168 (2.11%)	21 (3.58%)	126 (7.40%)	28 (2.67%)	348 (3.06%)
Total geral	9 (12.33%)	499 (6.30%)	43 (7.67%)	338 (19.85%)	60 (5.73%)	949 (8.35%)	

Fonte: Cardoso (2020, p. 137)

Tabela 11 – Tipos de desvio no fenômeno sibilante (Grupo 3).

Tipo de sibilante	Remetente					Grupo 3
	ANT	CAR	DIA	MMJ	PAL	
<s por z>	–	46	–	35	1	82
<s> por <ss>	1	–	2	32	–	35
<s> por <c>	–	26	1	6	1	34
<ç> por <c>	–	20	1	1	1	23
<ss> por <ç>	–	11	2	–	1	14
<z> por <s>	–	8	–	1	4	13
<ç> por <ss>	–	4	–	6	–	10
<c> por <ç>	1	4	–	3	–	8
<s> por <ç>	–	1	–	7	–	8
<ss> por <c>	–	5	2	–	–	7
<c> por <ss>	–	3	–	3	–	6
<ch> por <x>	–	4	–	1	1	6
<x> por <ch>	–	4	–	1	–	5
<c> por <s>	–	2	–	2	–	4
<c> por <x>	–	–	1	1	1	3
<ss> por <x>	–	–	2	–	–	2
<c> por <sc>	–	1	–	–	–	1
<s> por <sc>	–	1	–	–	–	1
<s> por <x>	–	–	1	–	–	1
<sc> por <c>	–	1	–	–	–	1
<zs> por <ss>	–	1	–	–	–	1
Total Geral	2	142	12	99	10	265

Fonte: Cardoso (2020, p. 139)

Como se vê na **Tabela 12**, a *nasalidade não padrão*, que correspondia a apenas 6 casos no grupo 1 (0.06%) e a 10 no grupo 2 (0.22%), entre os menos hábeis é mais frequente: são 129 dados (1.07%). Eram 4 tipos de *nasalidade não padrão* no grupo 1 e 3 tipos no grupo 2, ao passo que, para o grupo 3, são 7 tipos distintos. Assim, a *nasalidade não padrão* se revelou numérica e percentualmente relevante:

Tabela 12 – Tipos de desvio no fenômeno nasalidade não padrão (Grupo 3).

Tipo de nasalidade não padrão	Remetentes					Grupo 3
	ANT	CAR	DIA	MMJ	PAL	
<n> por <m>	1	35	–	20	1	57
<m> por <n>	–	8	1	31	2	42
<ão> por <am>	–	9	–	11	–	20
<am> por <ão>	–	4	–	–	–	4
<om> por <ão>	–	–	–	1	2	3
til por <n>	–	–	–	1	–	1
<óns> por <óes>	–	1	–	–	–	1
Total Geral	1	57	1	64	5	128

Fonte: Cardoso (2020, p. 140).

Destacam-se aqui as trocas de: <n> por <m>, como em *anbos* por *ambos* e *compadre* por *compadre*; e <m> por <n> como em *disemdo* por *dizendo* e *emtende* por *entende*. Tais dados evidenciam uma incompreensão dos contextos de uso dos grafemas <m> e <n>.

Na análise da categoria *escrita fonética*, observa-se que o *alteamento* aparece novamente como o fenômeno mais frequente, assim como visto nos grupos 1 e 2, em especial o alteamento da vogal /e/, como se vê na **Tabela 13**:

Tabela 13 – Tipos de desvio no fenômeno alteamento (Grupo 3).

Tipo de alteamento	Remetentes					Grupo 3
	ANT	CAR	DIA	MMJ	PAL	
vogal /e/	2	22	–	76	1	101
vogal /o/	1	24	5	21	1	52
Total Geral	3	46	5	97	2	153

Fonte: Cardoso (2020, p. 140)

Quanto ao alteamento, o remetente MMJ demonstra uma regularidade de comportamento, pois a conjunção aditiva “e” aparece sistematicamente grafada com <i>, e dados como *ista* por *está*, *istamos* por *estamos*, *istao* por *estão*, *istavamos* por *estávamos* evidenciam que a escrita fonética desse remetente não é um mero lapso ou descuido. A remetente CAR, do mesmo modo, apresenta regularidade no alteamento de algumas palavras como *duença* por *doença* (3 ocorrências) e *duente* por *doente* (5 ocorrências). Em outras, fica patente sua insegurança na grafia da palavra *fotografia* que apresenta polimorfismo gráfico: *fotugrafia*, *futografia* (2 ocorrências) e *futugrafia*.

A partir do segundo fenômeno mais frequente da categoria *escrita fonética*, o quadro é diferente do que se apresenta no grupo 1, pois aparecem entre os fenômenos mais frequentes o *betacismo* (0.88%), a *vogal temática do verbo* (0.18%) e a *redução* (0.15%), seguidos de categorias ausentes do grupo 1, como *anteriorização* e *líquida*. Novamente, é possível notar tendências em certos remetentes. Nas cartas de CAR, localizamos: o *betacismo* (72 dados), como em *cabaquinho* por *cavaquinho*, *bolta* por *volta* e *saves* por *sabes*; e a *vogal temática do verbo*, como em *estemos* por *estamos* e *precisemos* por *precisamos*.

Dos 72 dados de *betacismo* de CAR, a maior parte é de troca de por <v> (*saves* por *sabes*) e casos de <v> por (*cabaquinho* por *cavaquinho*), como se vê na **Tabela 14**:

Tabela 14 – Tipos de desvio no fenômeno betacismo (remetente CAR).

Tipo de betacismo	CAR	Total Geral
Betacismos com 	9	9
Betacismos com <v>	63	63
Total Geral	72	72

Fonte: Cardoso (2020, p. 141).

É possível identificar uma característica marcante no remetente DIA: a *anteriorização* da vogal (14 dados), em dados como: *graças é Deus* por *Graças a Deus*, *é carta* por *a carta*, e *é meu mano* por *a meu mano*.

O remetente MMJ, possivelmente o mais velho remetente, apresenta uma característica peculiar e única no material: o uso de formas anacrônicas para algumas palavras, se for levado em conta como modelo os anos de 1950. As informações sobre esse remetente situam seu nascimento no ano de 1898 e, assim, MMJ, tinha por volta de 50 anos quando escreveu suas cartas. Essa característica de sua escrita fica evidenciada nas grafias *dahi* por *daí* e *dacta* por *data*. Nas cartas de 1948, constam ainda consoantes geminadas em *nella* por *nela* (alternando com *nêla*), *annos* por *anos*, *bocca* por *boca*.

Esse remetente parece ter adquirido um modelo de escrita etimologizante quando jovem e continua a reproduzi-lo, pelo menos em algumas palavras, anos depois. No grupo 3, MMJ é o remetente que apresenta alto percentual de *escrita fonética*, em grande parte por casos de alteamentos. Se suas grafias etimologizantes poderiam ser elementos de erudição no século XIX, já não é o caso nos anos em que essas cartas foram escritas.

Sumarizando e discutindo os resultados dos alfabetizados

A **Tabela 15** traz uma média geral dos resultados por grupo, diferenciando, segundo os cálculos realizados, o perfil dos mais hábeis, de habilidade média e menos hábeis.

Tabela 15 – Percentuais de desvios por categoria e por grupo.

Grupo	Número de palavras	Número bruto de dados / porcentagem
Grupo 1 (mais hábeis)	9.993	Esriptualidade: 55 (0.55%) Escrita fonética: 13 (0.13%) Total: 68 (0.68%)
Grupo 2 (habilidade média)	4.422	Esriptualidade: 172 (3.89%) Escrita fonética: 35 (0.79%) Total: 207 (4.68%)
Grupo 3 (menos hábeis)	11.365	Esriptualidade: 601 (5.28%) Escrita fonética: 348 (3.06%) Total: 949 (8.34%)

Fonte: Cardoso (2020, p. 143).

A quantificação da **Tabela 15** não representa a média dos grupos. É um perfil “abstrato” das médias dos 3 grupos por gradiência. Obviamente indivíduos bastante diferentes foram incluídos em um mesmo grupo, mas foi possível, a partir dessas médias, tipificar o material que permitirá análises futuras de outros fatos linguísticos.

O grupo 1 é o mais homogêneo de todos e nele as porcentagens totais não atingem 1%. Esses redatores seriam, segundo os critérios adotados, os mais proficientes *na escrita da distância* e, portanto, produziram textos menos transparentes em relação à língua oral. O fato de termos apenas remetentes homens reitera que o acesso à educação pelas mulheres foi, por muito tempo, restrito. O grupo 2 apresentou porcentagens totais de desvios de 2.30% (ADR) até 5.27% (GEO). O grupo 3, o mais heterogêneo, varia de 5.73% (PAL) até 18.85% (MMJ). Os remetentes que se destacaram como os menos hábeis foram ANT e MMJ.

A quantidade de dados na categoria *esriptualidade* foi maior do que os de *escrita fonética* para quase todos os remetentes (exceto ANT). Dados da categoria *esriptualidade* aparecem mesmo nos remetentes mais hábeis, mas os fenômenos incluídos nessa categoria sempre aparecem em conjunto com os da *escrita fonética* que nunca influencia isoladamente. Nesse sentido, esses resultados confirmam a hipótese de Santiago (2019), quando afirma que a *esriptualidade* é o fator que melhor caracterizaria a escrita inábil, embora, no século XX, os índices grafofonéticos se tornaram

significativos para ajudar na definição do grau de habilidade. A *escriptualidade* parece ser um bom parâmetro para diferenciar graus de habilidade em remetentes mais hábeis.

Por um lado, uma tendência postulada não foi aparentemente confirmada: embora os percentuais de *escriptualidade* do grupo 3 sejam maiores, não houve maior variedade de fenômenos dessa categoria em relação aos demais grupos (foram 9 no grupo 3 e 10 no grupo 2). A hipótese era a de que quanto menos hábeis fossem os remetentes e quanto mais escrevessem, mais tipos diferentes de fenômenos seriam encontrados. Para a *escriptualidade*, constatou-se que a frequência obtida nos fenômenos levantados foi mais relevante do que a variedade de tipos propriamente dita.

Por outro lado, outra postulação foi confirmada: quanto mais palavras temos de um remetente, mais chances há de ocorrerem tipos e fenômenos diferentes em seu material. A remetente CAR, por exemplo, escreveu mais (7.954 palavras) e apresentou mais fenômenos diferentes (21, entre *escriptualidade* e *escrita fonética*). Mesmo assim, houve remetentes que escreveram menos palavras, mas apresentaram percentual mais alto do que outros que escreveram muito. Embora tenham escrito pouco, os remetentes ANT e DIA possuem percentual bastante diferente dos de CAM e AB, que escreveram mais. No mesmo raciocínio, ao se comparar remetentes com números de palavras equivalentes, como AB (1.653 palavras), GEO (1.894 palavras) e MMJ (1.703 palavras), observamos percentuais diferentes, estando cada um em uma categoria de habilidade. Assim, a frequência de uso se mostrou, de modo geral, mais relevante do que a diversidade de fenômenos e tipos encontrados.

Aparentemente alguns fenômenos foram balizares e, em algum grau, seriam sinalizadores de menor habilidade com a escrita na comparação com o grupo 1 (modelar):

- i. da categoria *escriptualidade*: “ausência de nasal”, “abaixamento” e “som [k]” só ocorreram nos grupos 2 e 3;
- ii. da categoria *escrita fonética*: “ditongação” e “prótese” só aparecem nos grupos 2 e 3;
- iii. apenas no grupo 3 estão presentes para *escriptualidade* “representação de [ʒ]” e para *escrita fonética* “representação da líquida”, “Monotongação”, “centralização”, “nasalização” e “S beirão”.

É possível atribuir um peso diferente às categorias, destacando aquelas que aparecem apenas no grupo 3 e não aparecem em todos os grupos. Um desvio de monotongação teria um peso maior como prova de inabilidade do que uma troca de sibilante de tipo <s> por <z> por aparecer apenas no grupo 3. Podemos dizê-lo com certeza, porque a diferença entre palavras escritas do grupo 1 para o grupo 3 não é tão grande: o grupo 1 escreveu 9.993 palavras e no 3 foram 11.365. A monotongação, quando ocorre, virá acompanhada de muitos outros fenômenos, como é o caso dos desvios encontrados nos remetentes MMJ e CAR.

A esse respeito, merece destaque a categoria “sílabas complexas”, que apareceu em todos os grupos de remetentes, inclusive nos redatores hábeis como CAM. Como Barbosa (1999), verificamos que mesmo textos de mais hábeis apresentam inversão de /r/ e a diferença entre hábeis e não hábeis em relação a esse fenômeno é apenas a proporção de incidência.

Considerações finais

O artigo procurou aplicar, com base em Barbosa (2017) e Santiago (2012; 2019), alguns protocolos metodológicos que ajudam na categorização social de redatores com graus diferentes de habilidade com modelos de escrita, quando não há informações tão precisas sobre o perfil social dos redatores.

Mesmo cientes que outras dimensões de análise são fundamentais para caracterizar a produção escrita mais ou menos influenciada por propriedades de *imediatez comunicativa* (cf. Koch; Oesterreicher, 2007), a análise dos desvios de *escriptualidade* e de *escrita fonética* mostraram relativa eficácia na tipificação de um *corpus* escrito do passado, como Barbosa já defendia.

Para responder às duas questões iniciais, confirmamos que as cartas analisadas podem fornecer, em graus diferentes para cada missivista, um texto menos influenciado pelas práticas correntes da escrita. Nas cartas de imigrantes portugueses do início do século XX, as características grafofonéticas já evidenciam a transposição, para a escrita, de características da pronúncia do português europeu, tais como *altea-mentos* e *ditongações* e, mais especificamente, do dialeto do norte de Portugal, como os *betacismos*. Na gradiente proposta, mostramos que os missivistas menos hábeis (grupo 3), por exemplo, apresentaram evidências do menor grau de habilidade com a escrita no conjunto analisado, o que permite inferir inclusive, para novas análises, que seus textos seriam os mais “transparentes” em relação à língua oral.

Em nossa análise do grau de habilidade, na forma do contato com textos escritos, concluiu-se que a variedade de fenômenos observados não acompanha necessariamente o aumento no percentual de desvios. Alguns desvios ocorreram apenas nos remetentes de menor grau de habilidade. Esses fenômenos são possivelmente “mais graves” ou indicativos de uma baixa habilidade com a escrita do que aqueles observados nos remetentes mais hábeis. Para lidar com um número menor de cartas, em um sistema que atribuísse pesos aos fenômenos, um desvio como a grafia das sibilantes poderia pontuar menos, por ser frequente até mesmo entre os mais hábeis. No fim das contas, esse desvio não caracteriza, apenas por sua ocorrência, a inabilidade: é na quantidade de desvios dessa natureza que reside a diferença entre um remetente mais hábil e um menos hábil.

Referências

BAGANHA, Maria Ioannis. As correntes emigratórias portuguesas no século XX e o seu impacto na economia nacional. *Análise Social: Revista do Instituto de Ciências da Universidade de Lisboa*, XXIX, n. 128, p. 959-980, 1994. Disponível em: http://analise-social.ics.ul.pt/?page_id=14. Acesso em: 10 maio 2019.

BARBOSA, Afrânio Gonçalves. *Para uma história do português colonial: aspectos linguísticos em cartas do comércio*. 1999. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999.

BARBOSA, Afrânio Gonçalves. Tratamento dos corpora de sincronias passadas da língua portuguesa no Brasil: recortes grafológicos e linguísticos. In: LOPES, Célia Regina dos Santos (org.). *A norma brasileira em construção: fatos linguísticos em cartas pessoais do século XIX*. Rio de Janeiro: UFRJ, FAPERJ, 2005. p. 25-43.

BARBOSA, Afrânio Gonçalves. O controle de marcas de inabilidade na escrita alfabética e a identificação das mãos inábeis em corpora histórico-diacrônicos. *Revista da ABRALIN*, v. 16, n. 2, p. 19-43, 2017.

BARBOSA, Afrânio Gonçalves; LIMA, Alexandre Xavier. O controle indireto de perfis sócio-históricos em corpora histórico-diacrônicos: a identificação de graus de letramento pela grafia etimológica no século XIX. In: CASTILHO, Ataliba T. de. (coord.). *História do Português Brasileiro v. 2: Corpus diacrônico do português brasileiro*. São Paulo: Contexto. 2019. p. 168-205.

CALDAS AULETE. *Dicionário contemporâneo da língua portuguesa*. 3. ed. Lisboa: Parceria Antônio Maria Pereira, 1948.

CARDOSO, Rafael Rodrigues da Silva. *De Portugal ao Brasil: Edição semidiplomática e descrição de cartas privadas da imigração portuguesa para o Brasil no século XX*. 2020. Tese (Doutorado em Letras Vernáculas (Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2020.

CINTRA, Luís Filipe Lindley. *Estudos de dialectologia portuguesa*. Lisboa: Sá da Costa Editora, 1983.

CONDE SILVESTRE, Juan Camilo. *Sociolingüística Histórica*. Madrid: Gredos, 2007

CRYSTAL, David. *Dicionário de Linguística e Fonética*. Rio de Janeiro: Jorge Zahar Editor, 1988.

DUBOIS, Jean, et al. *Dicionário de Linguística*. São Paulo: Editora Cultrix, 1999.

ELSPASS, Stephan. The use of private letters and diaries in sociolinguistic investigation. In: HERNÁNDEZ-CAMPOY, Juan M.; CONDE-SILVESTRE, Juan C. (ed.). *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, 2012. p. 156-169.

HERNÁNDEZ-CAMPOY, Juan M.; SCHILLING, Natalie. The Application of the Quantitative Paradigm to Historical Sociolinguistics: Problems with the Generalizability Principle. In: HERNÁNDEZ-CAMPOY, Juan M.; CONDE-SILVESTRE, Juan C. (ed.). *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, 2012. p. 63-79.

KOCH, Peter; OESTERREICHER, Wulf. Oralidad y escrituralidad a la luz de la Teoría del Lenguaje. In: KOCH, Peter; OESTERREICHER, Wulf. *Lengua Hablada en La Romania: español, francés, italiano*. Madrid: Editorial Gredos, 2007. p. 20-42.

LEITE, Ana Carolina da Silveira. Cartas de mulheres: história social da cultura escrita de e/imigrantes portuguesas no Brasil (1896-1929). *Cem Cultura Espaço e Memória*, v. 8, p. 357-372, 2017.

LEMLE, Miriam. *Guia Teórico do Alfabetizador*. São Paulo: Editora Ática, 1999

MARQUILHAS, Rita. Mãos Inábeis nos Arquivos da Inquisição. In: *A Faculdade das letras: Leitura e escrita em Portugal no séc. XVII*. Lisboa: Imprensa Nacional-Casa da Moeda, 2000. p. 228-266.

RAPOSO, Eduardo Paiva B.; NASCIMENTO, Maria Fernanda Bacelar do; MOTA, Maria Antónia Coelho da; SEGURA, Luísa; MENDES, Amália; VICENTE, Graça; VELOSO, Rita. (org.). *Gramática do Português*. 2 v. Lisboa: Fundação Calouste Gulbenkian, 2013.

ROMAINE, Suzanne. *Socio-Historical Linguistics: Its Status and Methodology*. Cambridge: Cambridge University Press, 1982.

SANTIAGO, Huda da Silva. *Um estudo do português popular brasileiro em cartas pessoais de “mãos cândidas” do sertão baiano*. 2012. Dissertação (Mestrado em Estudos Linguísticos) - Programa de Pós-graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana, 2012.

SANTIAGO, Huda da Silva. *A escrita por “mãos inábeis”: uma proposta de caracterização*. 2019. Tese (Doutorado) - Programa de Pós-Graduação em Língua e Cultura, Universidade Federal da Bahia, Salvador, 2019.

SERRÃO, Joel. *A emigração portuguesa*. 4. ed. Lisboa: Livros Horizonte, 1982.

SILVA, Érica Nascimento. *Cartas amorosas de 1930: o tratamento e o perfil sociolinguístico de um casal não-ilustre*. 2012. Dissertação (Mestrado) - Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

TEYSSIER, Paul. *História da Língua Portuguesa*. São Paulo: Sá da Costa Editora, 1982.

VASCONCELLOS, José Leite de. *Esquisse d'une Dialectologie Portugaise*. 2. ed. Lisboa: Centro de Estudos Filológicos, 1970. Disponível em: <http://purl.pt/160>. Acesso em: 23 ago. 2019.