

# LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS  
DA UNIVERSIDADE DO PORTO

VOL. 16, 2021

## FICHA TÉCNICA

*Linguística*  
*Revista de Estudos Linguísticos da Universidade do Porto*  
Vol. 16, 2021

ISSN: 1646-6195

Periodicidade: Anual

Diretor:  
António Leal

Editores:  
Faculdade de Letras da Universidade do Porto  
e Centro de Linguística da Universidade do Porto

Capa:  
José Osswald

Impressão e acabamentos:  
Invulgar - Artes Gráficas

Depósito Legal:  
248653/06

Tiragem:  
130 Exemplares

Os artigos publicados estão sujeitos a “peer review”.

Esta edição é integralmente financiada pelo Centro de Linguística da Universidade do Porto, ao abrigo do Programa de Financiamento FCT-UIDB/00022/2020 (Fundação para a Ciência e a Tecnologia).

A Revista está registada no DOAJ e Latindex e está indexada na base de dados Fonte Académica.

<http://ler.letras.up.pt/site/default.aspx?qry=id04id191&sum=sim>

## Espaço da Direção

O volume 16 da *Linguística* marca uma mudança na direção da revista.

Tendo sido fundada em 2006, foi dirigida inicialmente pela Professora Doutora Maria da Graça Pinto (2006-2009), que a lançou e editou os primeiros quatro volumes. Seguiu-se a Professora Doutora Ana Maria Brito, que, após onze anos como diretora (de 2010 a 2020), deixou as tarefas de edição da revista, cargo que desempenhou de forma exemplar, o que está patente nos dez volumes que foram por si editados durante este período.

Neste momento de início de funções de uma nova direção da revista, é tempo de agradecer e prestar tributo ao trabalho que ambas desenvolveram, que muito prestigiou o CLUP e a FLUP, assim como assumir um compromisso de continuar o legado que deixaram, criando as condições necessárias para que a revista continue a crescer e possa acompanhar as novas tendências e formatos de disseminação da ciência.

O primeiro artigo deste volume é da autoria de Cíntia Alcântara e Carmen Matzenauer, da Universidade Federal de Pelotas. Trata-se de um estudo diacrónico sobre o Francês, que adota os pressupostos da Fonologia Autossegmental na abordagem da vocalização da lateral em coda interna. As autoras defendem que o surgimento, em Francês, de vogais frontais arredondadas (inexistentes em Latim) tem subjacentes fatores fonético-fonológicos e está relacionado com a evolução da lateral em posição de coda.

O segundo artigo é da autoria de Jéssica Gomes e Celeste Rodrigues, da Universidade de Lisboa. Neste trabalho, as autoras apresentam os resultados de um estudo efetuado, a partir de dados do *corpus* EFFE-On, sobre as produções escritas de crianças de Chaves, Porto e Lisboa, no sentido de averiguar do impacto, na escrita, da (não) manutenção da distinção fonológica entre a africada pós-alveolar surda /tʃ/ e a fricativa pós-alveolar surda /ʃ/. As autoras concluem, entre outros aspetos, que a existência dessa distinção fonológica em certas zonas dialetais tem reflexo direto na maior facilidade das crianças dessas zonas em distinguirem os contextos de utilização do grafema <x> e do dígrafo <ch>.

Liliana Correia e Cristina Flores, da Universidade do Minho, são as autoras do último artigo deste volume, no qual apresentam o *Questionário*

*Sociolinguístico Parental para Famílias Emigrantes Bilingues* (QuesFEB). Este questionário foi concebido para ser usado no âmbito do bilinguismo de herança e permite não só a recolha de informação de tipo biográfico e sociolinguístico sobre cada criança e respetivo contexto, mas também a avaliação do *input* linguístico a que a criança está exposta. As autoras descrevem, no artigo, as diversas partes do QuesFEB e o método de cálculo das variáveis mais importantes.

Este número da revista termina com duas recensões. A primeira delas, da autoria de Inês Cantante (Universidade do Porto), tem como objeto o volume *The Semantics of Gradability, Vagueness, and Scale Structure. Experimental Perspectives*, editado por Elena Castroviejo, Louise McNally e Galit Sassoon.

A segunda recensão, da autoria de Purificação Silvano (Universidade do Porto), aborda o livro *Linguistic Linked Data. Representation, Generation and Applications*, de Philipp Cimiano, Christian Chiarcos, John P. McCrae e Jorge Gracia.

Para terminar, um agradecimento especial aos dezanove avaliadores que contribuíram, de forma anónima, com o seu excelente trabalho para a produção deste número da revista *Linguística* (FCT-UIDB/00022/2020).

Porto, julho de 2021

António Leal

# SUMÁRIO

Espaço da direção .....	3
Artigos	
A vocalização da lateral posvocálica em Francês .....	9
<i>Cíntia Alcântara e Carmen Matzenauer</i>	
O grafema <x> e o dígrafo <ch>: um estudo longitudinal do desempenho ortográfico de crianças de três dialetos portugueses .....	39
<i>Jéssica Gomes e Celeste Rodrigues</i>	
Questionário sociolinguístico parental para famílias emigrantes bilingues (QuesFEB): uma ferramenta de recolha de dados sociolinguísticos de crianças falantes de herança .....	75
<i>Liliana Correia e Cristina Flores</i>	
Recensões	
Elena Castroviejo, Louise McNally e Galit Sassoon (Eds.). The Semantics of Gradability, Vagueness, and Scale Structure. Experimental Perspectives .....	105
<i>Inês Cantante</i>	
Philipp Cimiano; Christian Chiarcos; John P. McCrae; Jorge Gracia (2020). <i>Linguistic Linked Data. Representation, Generation and Applications</i> . Springer International Publishing.....	111
<i>Purificação Silvano</i>	



Artigos



# A vocalização da lateral posvocálica em francês

Cíntia da Costa Alcântara

cintiaca09@gmail.com

*Universidade Federal de Pelotas (Brasil)*

Carmen Lúcia Barreto Matzenauer

carmen.matzenauer@gmail.com

*Universidade Federal de Pelotas/CNPq (Brasil)*

## ABSTRACT

Vocalization of the lateral in syllabic coda provided French with forms that have rounded posterior vowels, such as /o, u/ (cf. *salvus* > fr. *sauvs* /o/ 'salvo'; *pulmonem* > *poumon* /u/ 'pulmão'), and rounded front ones, such as /ø, y/ (cf. *filtrum* > *feutre* /ø/ 'feltro'; *pul(i)cem* > *puce* /y/ 'pulga'). The latter did not exist in the Latin vowel system; thus, they were unknown, unlike the former. In the diachronic context, this study focuses on the evolution of Latin forms bearing 'l pinguis', which resulted in French words with rounded vowels, either [- post] /ø/ or [+ post] /o/. This paper argues that one of the sources of rounded front vowels in French is the Latin sequence of vowel + lateral liquid in the medial position of the word. The explanation is provided with the support of Autosegmental Phonology, by the features that comprise the internal structure of the segments of the sequence. The corpus under study was collected in historical grammars, etymological dictionaries and historical compendia of French phonetics and phonology.

## KEY-WORDS

diachrony of French; vowel system; postvocalic lateral; feature theory

## RESUMO

A vocalização da lateral em coda silábica rendeu ao francês formas com vogais posteriores arredondadas, como /o, u/ (cf. *salvus* > fr. *sauvs* /o/ 'salvo'; *pulmonem* > *poumon* /u/ 'pulmão'), mas também formas com vogais frontais arredondadas, conforme /ø, y/ (cf. *filtrum* > *feutre* /ø/ 'feltro'; *pul(i)cem* > *puce* /y/ 'pulga'), estas inexistentes no latim. Em panorama diacrônico, este estudo focaliza a evolução de formas latinas portadoras do 'l pinguis' que resultaram em vocábulos do francês portadores de vogal arredondada, seja [- post] /ø/, seja [+ post] /o/. Defende-se estar, na sequência latina de vogal + líquida lateral, em posição medial de palavra, uma das fontes da presença de vogais frontais arredondadas do francês, o que se explica, com o suporte da Fonologia Autossegmental, pelos traços que compõem a estrutura interna dos segmentos da sequência. O corpus estudado partiu de gramáticas históricas, dicionários etimológicos e compêndios históricos de fonética e fonologia francesas.

## PALAVRAS-CHAVE

diacronia do francês; sistema vocálico; lateral posvocálica; teoria de traços

## 1. Introdução

A história linguística do francês registra quatro momentos – francês antigo (sécs. IX-XIII); francês médio (sécs. XIV-XVI); francês clássico (sécs. XVII-XVIII); e francês moderno (sécs. XIX-XXI) (Veiga, 1965). O francês antigo é o palco em que se estabelece a lateral velarizada posvocálica e é também nele que tal elemento consonantal sofre o processo de vocalização, conforme será apresentado neste artigo.

O presente estudo tem por foco a vocalização da lateral, foneticamente velarizada, em coda silábica na diacronia do francês<sup>1</sup>, especificamente a coda silábica medial de palavra. O objetivo é descrever e analisar, sob os pressupostos da Fonologia Autossegmental, a evolução de formas latinas portadoras do ‘l *pinguis*’ que resultaram em vocábulos do francês portadores de vogal arredondada, seja [-post], a exemplo de /ø/ e /y/, seja [+ post], a exemplo de /o/ (e.g. *filtrum* > *feutre* /*føtre*/ ‘futre’; *pul(i)cem* > *puce* /*pyse*/ ‘pulga’; *alter* > *autre* /*otre*/ ‘outro’)<sup>3</sup>.

É uma herança latina a líquida lateral /l/ acolhida não somente pela língua francesa, mas também por outras línguas neolatinas, como o português e o italiano, por exemplo. No francês *standard* contemporâneo<sup>4</sup>, a lateral coronal anterior é a única representante na classe das líquidas laterais, podendo ocupar tanto a posição de ataque quanto a de coda silábica; porém, nem sempre foi assim: o francês também já contou com a lateral palatal /ʎ/, em

<sup>1</sup> Este estudo trata do francês *standard*, nascido do dialeto falado em Île de France, onde fixaram residência os reis de França e sua corte.

<sup>2</sup> Segundo Camara Jr. (1976: 49), a variante *pinguis* da líquida lateral latina /l/ apresentava “uma articulação posterior, provavelmente velar”. O termo *pinguis* é uma denominação atribuída ao gramático latino Consentius, no séc. V d. C. para descrever a variante lateral [+post] do /l/ em coda. Outro gramático latino, Plínio O Velho, no séc. I d. C., já o denominara *plenus* (Sen 2015:15-6). É possível encontrar-se na literatura atual as expressões *dark l*, em inglês, e *l* *sombre*, em francês.

<sup>3</sup> O objeto deste estudo restringe-se ao processo diacrônico da lateral em coda silábica de que resultaram as vogais posteriores arredondadas /o/ e /u/ e as vogais frontais arredondadas /ø/ e /y/, deixando de abordar a vogal frontal arredondada média-baixa /œ/, como elemento resultante da vocalização da lateral posvocálica, em virtude de, com essa motivação específica, este elemento vocálico ter ocorrido somente enquanto forma transitória/intermediária, ou seja, no processo evolutivo de que resultou a vogal média-alta frontal arredondada /ø/ (cf. chev[ɛ]s > chev[ew]s > chev[œw]s > chev[œs] > chev/o/ *cheveux*).

<sup>4</sup> Já no século XVI, após haver cessado o período em que se deu o processo de vocalização da líquida lateral em coda, os vocábulos com /l/ posvocálico – empréstimos ao italiano (cf. *balcon* ‘sacada’; *altesse* ‘alteza’; *galbe* ‘garbo’) ou reconstruções vocabulares (cf. *malgré* ‘apesar de’; *malfaire* ‘fazer mal’; *malcontent* ‘descontente’) – manifestavam, de acordo com Bourciez (1921), a lateral coronal anterior [l].

seu inventário fonológico. Esta aproximante, de cujo tratamento o presente estudo não se ocupa, foi vocalizada para [j], tendo vigorado no sistema consonantal da língua francesa até meados do séc. XIX<sup>5</sup>.

No Quadro 1, a seguir, têm-se, à esquerda, os sistemas vocálicos do latim clássico e do latim vulgar; à direita, tem-se o sistema vocálico do francês<sup>6</sup>.

Quadro 1 – Sistemas vocálicos do latim (clássico e vulgar) e do francês.

Vogais Latinas		Vogais Francesas	
Lat. Clássico > Lat. Vulgar			
<u>ī</u> - <u>i</u>	<u>ū</u> - <u>ū</u>	i	u
ĩ → e	ũ → o	e	o
ē → e	ō → o	ε	ɔ
ě - <u>ε</u>	ǔ - <u>ő</u>		
	ā → a		a
	ǎ → a		
		y	
		ø	
		œ	

Observa-se, no Quadro 1, que o sistema vocálico do latim vulgar, no quadro sublinhado, não mais contava com a ‘duração’ como parâmetro distintivo, como o fizera o latim clássico; trabalhava somente com a informação de tonicidade. Na sílaba tônica, o desenvolvimento do sistema vocálico latino ocorre pela aparição no sistema de “duas vogais médias-baixas /ε/ (< ě lat. clás.) e /ɔ/ (< ő lat. clás.), as quais se opõem a /e/ (< lat. clás. {ē, ĭ}) e /o/ (< {ō, ũ}); com relação à sílaba átona, a série anterior {ĩ, ē, ě} passa para /e/ e a posterior {ő, ō, ũ}, para /o/” (cf. Ségéral & Scheer 2015: 117; Camara Jr. 1976). E essa nova configuração vocálica, presente no latim vulgar, foi herdada por línguas românicas, como o francês e o português. O francês, contudo, acresceu novos segmentos ao inventário segmental – a série arredondada de vogais frontais, /y, ø, œ/ – desconhecidas no sistema latino, conforme é apresentado no Quadro 1.

E da sequência latina de vogal seguida de líquida lateral velarizada

<sup>5</sup> A literatura diagnóstica esse fato (cf. Tranel 1987; Grevisse 1993).

<sup>6</sup> No quadro, à esquerda, as flechas apontam para as vogais do latim vulgar, derivadas do sistema vocálico do latim clássico; à direita, estão em destaque as vogais frontais arredondadas que integram a fonologia do francês contemporâneo.

resultou em francês uma vogal arredondada, de caráter posterior, /u, o/, ou frontal, /y, ø/. No francês antigo (sécs. IX-XIII), a lateral velarizada, em coda, seguida ou não de outra consoante, apresentava características de elemento velarizado, [ʃ] – uma característica herdada do ‘l’ *pinguis* (velar) latino. A vigência desta variante se deu até cerca do séc. XI<sup>7</sup>, período após o qual a língua passou a contar com uma forma vocalizada [w].

Ressalta-se, não obstante, que em vocábulos derivados o francês conta com a presença da sequência constituída de vogal e líquida lateral (coronal), em (1) mostrado.

(1) Formas derivadas em V + /l/<sup>8</sup>

latim	francês antigo	francês atual	português	francês - forma deriv.	português
pulmonem	> p[ot̪]mon	> p/ <u>u</u> /mon	‘pulmão’	pulmonaire	‘pulmonar’
balsamum	> b[at̪]me	> b/ <u>o</u> /me	‘bálsamo’	balsamique	‘balsâmico’
capillos	> chev[eʃ]s	> chev/ <u>ø</u> /	‘cabelos’	chevelure	‘cabeleira’

Os vocábulos derivados, em (1), funcionam como evidências morfofonológicas que atestam uma etapa pregressa da língua, com a presença da lateral em coda silábica. Por outro lado, as formas sincrônicas com a vogal arredondada (cf. *pu/mon*; *b/o/me*; *chev/ø/l*) guardam informações sob o rótulo de traços distintivos, na linha de Clements & Hume (1995)<sup>9</sup>.

O presente estudo se atém aos fatos linguísticos ocorridos em coda interna à palavra não derivada, conforme já referido, e utilizou, como

<sup>7</sup> De acordo com Nyrop (1899: 278), as primeiras manifestações de vocalização da lateral *pinguis* teriam sido observadas ao final do período Galo-Romano – por volta da metade do séc. IX. Os exemplos mais antigos remontariam às cartas latinas do séc. X, em que se encontram nomes próprios de origem germânica, grafados com a vogal *u* na posição destinada à líquida lateral (ao lado de *a*) (e.g. *Rainaudus*, *Giraut*, *Gauterius*). No séc. XII, ainda era possível encontrar-se grafada a líquida lateral, a exemplo de *albe* (> *aube* ‘amanhecer’); *altre* (> *autre* ‘outro’); *chevals* (> *chevaux* ‘cavalos’); *palme* (> *paume* ‘palma’); *talpe* (> *taupe* ‘toupeira’); *falcon* (> *faucon* ‘falcão’) etc. O autor relata que deveria ocorrer, à época, “a pronúncia de um tipo de *l* muito reduzido, cuja articulação consonântica mal devia ser percebida”

<sup>8</sup> Neste artigo optou-se por apresentar as formas fonéticas e as formas fonológicas apenas nos segmentos foco da análise.

<sup>9</sup> Adota-se a proposta teórica de Clements & Hume (1995) para a análise dos dados do presente artigo em razão da visão não-linear do funcionamento do componente fonológico das línguas, com o tratamento autosssegmental dos traços que compõem a estrutura interna dos segmentos e a formalização dos processos que envolvem segmentos e traços por meio de uma representação geométrica.

*corpus*, gramáticas históricas, dicionários etimológicos e compêndios históricos de fonética e fonologia francesas.

## 2. A vocalização da lateral em coda interna

Segundo Nunes (1989 [1951]: 132), o francês é a língua românica em que se verifica “mais extensamente e com maior regularidade” o processo de vocalização do ‘l’ velarizado. Na literatura encontram-se afirmações semelhantes à de Nunes com respeito ao francês, seja antes ou depois do referido autor português (cf. Kolovrat 1923; Müller 2011; Recasens 2012). Segundo Sen (2015), a vocalização da lateral em coda consistiria em uma das evidências da influência indireta protagonizada pela estrutura silábica na mudança sonora.

A lateral velarizada, nos primórdios do francês antigo, compunha uma sequência com a vogal-núcleo<sup>10</sup>, /a/, /e/, /ɛ/, /i/, /o/, /ɔ/ ou /u/, ou seja, quaisquer que fossem as propriedades fonológicas da vogal, a lateral posvocálica manifestava-se como segmento velarizado, à borda da sílaba, em meio ou fim de palavra, em (2) ilustrado. Da mesma forma, pouco importavam as propriedades segmentais da consoante vizinha à líquida lateral em coda interna, posto que o fenômeno de vocalização se repetia.

### (2) Sequência V + /l/

latim	francês antigo	francês atual	português
pulmonem	> p[ot̪]mon	> p/ <u>u</u> /mon (poumon)	‘pulmão’
balsamum	> b[at̪]me	> b/ <u>o</u> /me (baume)	‘bálsamo’
capillos	> chev[et̪]s	> chev/ <u>o</u> / (cheveux)	‘cabelos’
pul(i)cem	> p[ut̪]ce	> p/ <u>y</u> /ce (puce)	‘pulga’

<sup>10</sup> Segundo Brachet (1872: 31-2), a líquida lateral em coda, precedida de uma vogal (a, e, i, o, u), persistiu no francês primevo (cf. *mollis* = mol; *malva* = malve; *porcellus* = pourcel); porém, foi abrandada em u (cf. *mou* ‘mole’; *mauve* ‘malva’; *pourceau* ‘porco’) ao redor da metade do séc. XII.

Da sequência latina de V + /l/, em (2) ilustrada, desenvolveu-se no francês uma vogal [+arred] – seja /u/ (cf. *poumon* p/u/mon), seja /o/ (cf. *baume* b/o/me), ou, ainda, /ø/ [-post, +arred] (cf. *cheveux* chev/ø/)<sup>11</sup> e /y/ (cf. *puce* p/y/ce). Contudo, a passagem de um momento a outro na história da lateral posvocálica (V + /l/) não se deu automaticamente; apresenta etapas, o que faz com que se formulem duas questões a que se pretende responder neste estudo:

- (i) Quais seriam as causas fonético-fonológicas a desencadear a mudança de um sistema sem vogais frontais arredondadas (latim vulgar) para outro sistema com essas vogais (francês)?
- (ii) A inserção de vogais arredondadas, frontais ou posteriores, na fonologia do francês mostraria relação com a história da lateral posvocálica na língua?

Nas seções que seguem, apresenta-se o processo evolutivo de cada uma das sequências de vogal seguida de líquida lateral (V + /l/) – /a/, /ɔ/, /o/, /u/ e /i/ –, das quais resultaram, na língua francesa, vogais arredondadas: /o/, /u/, /ø/. É sob os pressupostos da Fonologia Autossegmental e da Geometria de Traços (Clements & Hume 1995) e da proposta de Martinet (1955), relativamente à evolução dos sistemas fonológicos, que serão apresentadas possíveis respostas para as indagações formuladas. Basilar para a análise foi o reconhecimento de que os segmentos que compõem as sequências de vogal + líquida lateral compreendem, em sua estrutura interna, traços organizados em uma hierarquia, funcionando como unidades autossegmentais, o que permite explicar por que derivaram, nos movimentos diacrônicos do francês, vogais arredondadas, com a ocorrência especialmente de processos de vocalização da lateral, de assimilação e de coalescência; os traços estão no cerne dos processos que dão origem às vogais /o/, /u/, /ø/ que integram a fonologia do francês.

---

<sup>11</sup> A vogal média-baixa frontal arredondada /œ/ existiu como elemento resultante da vocalização da lateral posvocálica somente enquanto forma transitória/intermediária, ou seja, no processo evolutivo de que resultou a vogal média-alta frontal arredondada /ø/ (cf. chev[et]s > chev[ew]s > chev[œw]s > chev[œ]s > chev/ø/ *cheveux*), conforme foi preliminarmente descrito na nota 3.

## 2.1 Sobre a sequência /a/

Da sequência *al* derivou-se, no sistema linguístico do francês, a vogal arredondada /o/. A seguir, no Quadro 2, apresenta-se o processo evolutivo por que passou a lateral em coda, resultando na vogal oral /o/, passando pela velarização da consoante lateral e pela sua subsequente vocalização (cf. [a] > [aw] > /o/). Adianta-se que tal processo de vocalização da lateral velarizada, em coda interna, atingiu todas as formas que carregavam, nos primórdios do francês antigo, a referida sequência de vogal e aproximante. À esquerda, estão as formas oriundas do latim vulgar, na coluna central, estão dispostas as formas evolutivas francesas e, à direita dessas, as formas correspondentes em português. As formas que carregam um asterisco são alvos hipotéticos, que acompanhariam o percurso evolucionar das demais formas em vogal /a/ seguida da líquida lateral velarizada. Seguem-se comentários.

Quadro 2 - Vocalização da líquida lateral em coda medial – resultado: /o/.

	/ <u>l</u> / [aʎ] > [aw] > /o/	
<u>a</u> ltare	<b>aʎ</b> tel > <b>aw</b> tel > <b>au</b> tel [aʎ] > [aw] > /o/	‘altar’
b <u>a</u> lsamum	<b>baʎ</b> me > <b>ba</b> wme > <b>ba</b> ume [aʎ] > [aw] > /o/	‘bálsamo’
<u>c</u> alvus	<b>caʎ</b> ve > <b>ca</b> wve > <b>cha</b> uve [aʎ] > [aw] > /o/	‘calvo’
<u>f</u> alconem	<b>faʎ</b> cun* > <b>fa</b> wcun > <b>fa</b> ucon [aʎ] > [aw] > /o/	‘falcão’
<u>g</u> albinus	<b>gaʎ</b> ne > <b>ga</b> wne > <b>ga</b> une [aʎ] > [aw] > /o/	‘amarelo’
<u>m</u> alva	<b>maʎ</b> ve > <b>ma</b> wve > <b>ma</b> uve [aʎ] > [aw] > /o/	‘malva’
<u>p</u> alma	<b>paʎ</b> me > <b>pa</b> wme > <b>pa</b> ume [aʎ] > [aw] > /o/	‘palma’

palpebra	paɫpebra* > pawpebra > paupièrè [aɫ] > [aw] > /o/	‘pálpebra’
salmonem	saɫmun > sawmun > saumon [aɫ] > [aw] > /o/	‘salmão’

Interessam ao presente trabalho as formas evolutivas francesas, as quais são, na primeira etapa, constituídas de uma sequência de vogal baixa posterior, /a/, seguida de líquida lateral, /l/, ou seja, V + /l/. Essa sequência teria passado por três etapas evolutivas: (i) a sequência de vogal /a/ seguida da líquida lateral /l/, a qual se manifestava velarizada foneticamente; (ii) o grupo de vogal /a/ seguida de glide dorsal [w]; e (iii) a vogal oral /o/<sup>12</sup>, resultante de um processo de coalescência da sequência V + glide, ou seja, a/l/ [ɫ] > a[w] > /o/ [o]. Ademais, na sequência tautossilábica /al/, segundo Kolovrat (1923:79), a vocalização da líquida lateral ter-se-ia processado independentemente tanto da tonicidade da sílaba, quanto do ponto de articulação da consoante seguinte – dental, palatal ou labial (cf. *altare* > *utel* ‘altar’; *caltiare* > *chusser* ‘calçar’; *salvare* > *sauver* ‘salvar’; *talpa* > *taupe* ‘toupeira’).

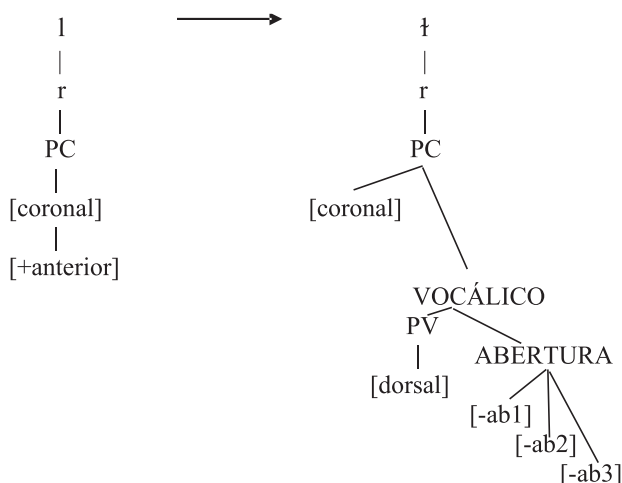
Deve-se observar que a líquida lateral vocalizada, passando a comportar-se como glide dorsal, necessariamente permaneceu em coda silábica, ao lado da vogal-núcleo, /a/, do que resultou, nesse momento, um ditongo decrescente, [aw], o qual, por sua vez, passou à vogal /o/, em etapa subsequente<sup>13</sup>.

Observe-se, na Figura 1, a representação autossegmental do processo de velarização da líquida lateral, /l/ > [ɫ], ilustrada na forma francesa arcaica [’aɫbe], sincronicamente vocalizada – [’ob] *aube* ‘alvorecer’. Os fundamentos da Fonologia Autossegmental permitem a formalização da lateral velarizada como um segmento complexo, por conter, associada à articulação primária consonantal coronal da lateral /l/, uma articulação secundária vocálica. Veja-se a Figura 1.

<sup>12</sup> A vogal média-alta posterior /o/ no francês é representada ortograficamente por ‘au’, ‘eau’ e ‘o’, esta última grafia pode vir ou não acompanhada de acento circunflexo.

<sup>13</sup> Segundo Bourciez (1921: 239), as palavras *malgré* ‘malgrado’; *malfaire* ‘fazer mal’; *malcontente* ‘descontente’ foram refeitas/reconstruídas, pois, no francês antigo, eram *maugré*, *maufaire*, *maucontent* (cf. francês moderno *maudire* ‘maldizer’; *maussade* ‘rabugento’, etc.). O autor ressalta ainda que palavras como *balcon* ‘sacada’; *altesse* ‘altura’; *galbe* ‘garbo’; *soldat* ‘soldado’, etc. são empréstimos do italiano que entraram para o francês no séc. XVI.

Figura 1 - Velarização da líquida lateral em coda.



Fonte: Quédnau (1994)<sup>14</sup>

Como se pode observar, na Figura 1, para que se forme o segmento complexo reconhecido como a líquida velarizada, é necessário que haja dois pontos de articulação – um primário, ou de consoante, e o outro secundário, ou de vogal. A constrictão primária representa a líquida como um segmento [coronal]; a constrictão secundária assinala um elemento [dorsal], com traços de abertura que caracterizam uma vogal alta posterior ou dorsal.

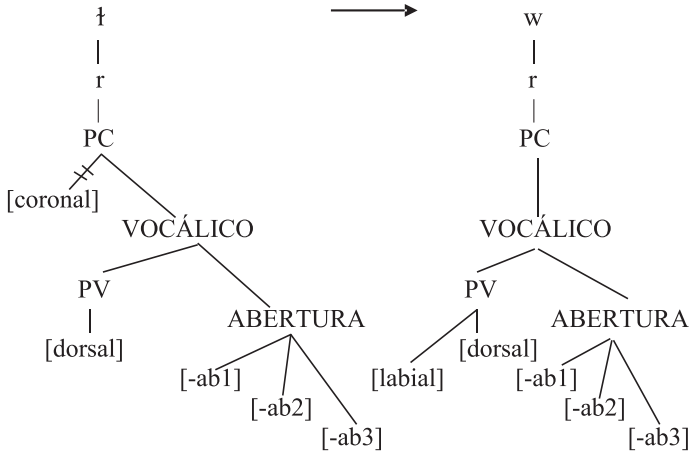
Na Figura 2<sup>15</sup>, a seguir, representa-se o processo de vocalização da soante lateral<sup>16</sup>, sob a Geometria de Traços (Clements & Hume 1995).

<sup>14</sup> A interpretação da líquida lateral velarizada, sob o modelo da Geometria de Traços (Clements 1985, 1989, 1991), é assumido por Quédnau (1994: 145) para o tratamento da lateral posvocálica no português gaúcho. À luz da Geometria de Traços, a lateral velarizada é interpretada como consoante complexa, com duas articulações orais: uma consonantal, representada, na Figura 1, pelo traço [coronal] sob o nó Ponto de Consoante (PC), outra vocálica, representada pela presença do nó Vocálico, também dependente do nó PC. Vale acrescentar ainda que, no momento do acréscimo do nó Vocálico com o ponto [dorsal], o traço [+anterior], dependente do [coronal], sob o ponto de C, é desligado por relação implicacional; prevalece então o ponto [dorsal]. O resultado é uma lateral dorsalizada.

<sup>15</sup> O glide [w] assim como a vogal [u] por ser um segmento labiodorsal, em sua estrutura interna, também se manifesta com o traço [labial].

<sup>16</sup> Embora resulte um ditongo a partir da vocalização da líquida lateral, considera-se que o glide ocupa, na estrutura silábica, a posição de coda, seguindo a proposta de Mascaró (1989) e Bisol (1994).

Figura 2 - Vocalização da líquida lateral em coda.



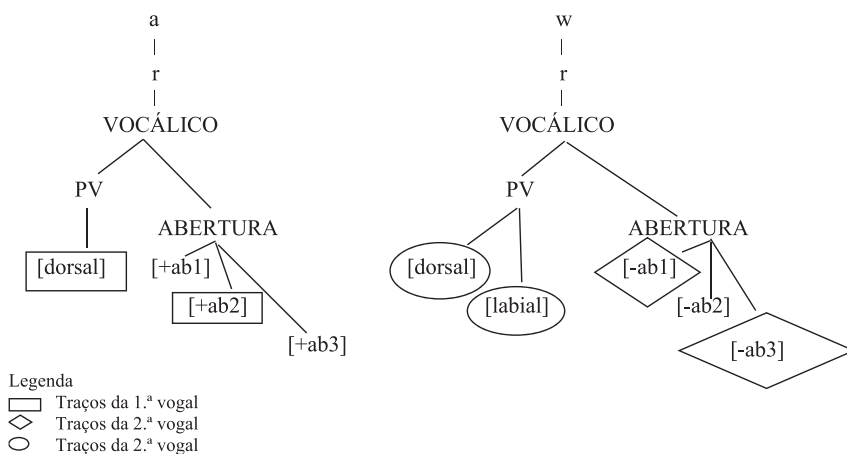
Fr. ant. /albe/ [ʼawbe] (> fr. mod. *aube* ‘alvorecer’)

Fonte: as autoras

O fenômeno de vocalização da lateral velarizada em coda, na Figura 2, ocorrido no francês, assim como no português do Brasil, é formalizado como um mecanismo de desassociação do traço [coronal], imediatamente sob o nó Ponto de C, que caracteriza segmentos consonantais. A informação que resta na hierarquia concerne a um elemento vocálico, sob o constituinte Nó Vocálico – os traços [dorsal] e [labial], sob o nó Ponto de V – e os traços de Abertura. Está concluída, assim, a vocalização da líquida lateral.

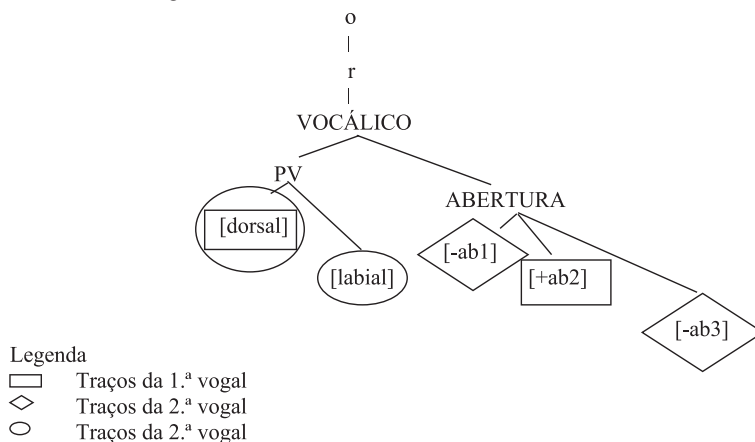
Contudo, no francês, o processo evolutivo não cessou a marcha: o resultado da vocalização de /l/, ou seja, o glide dorsal, passa a integrar a sequência de V + glide dorsal. Essa sequência é, então, submetida a novo processo, do qual redundará a vogal arredondada /o/. O processo responsável pela emergência desse segmento vocálico é a coalescência de traços, na Figura 3 apresentada, e cujo resultado final está na Figura 4.

Figura 3 - Unidades envolvidas no processo de coalescência de traços.



Fonte: as autoras

Figura 4 - Resultado da coalescência de [a] e [w].



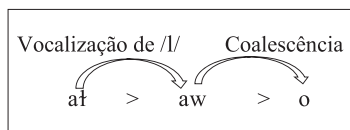
Fonte: as autoras

Na Figura 3, tem-se a formalização das unidades envolvidas no processo de coalescência da vogal /a/ e do glide dorsal [w], fazendo resultar a vogal /o/, representado na Figura 4. Note-se que são emprestados ao novo

segmento, pela vogal /a/, os traços [dorsal], de Ponto de Vogal, e [+ab2], de Abertura (destacados pelo retângulo). Com respeito aos traços do glide [w], são utilizados [dorsal] e [labial], que caracterizam o ponto de articulação do segmento vocálico/glide dorso-labial (destacados pelo círculo); quanto aos traços de Abertura, [-ab1] e [-ab3] são chamados a integrar o novo segmento (destacados pelo losango). Assim, o segmento que emerge da coalescência de /a/ e [w], na Figura 4 mostrado, carrega traços de ambos os segmentos, reunindo-os para a formação de outro segmento que deles difere: trata-se da vogal arredondada /o/, já integrante do sistema fonológico do francês, e conhecida do latim vulgar. Uma análise dos traços que integram a estrutura interna da sequência vogal + lateral é indispensável para o entendimento da emergência da vogal arredondada /o/. O entendimento de que os traços, como autosegmentos que, de forma coocorrente, constituem os segmentos, está na base da ocorrência dos processos de que derivam também as outras vogais arredondadas do francês que estão no foco do presente estudo.

Na Figura 5, apresenta-se a evolução do grupo /al/ até chegar à vogal /o/, assinalando-se-lhe as diferentes fases.

Figura 5 - Evolução do grupo /al/ a partir da manifestação fonética velarizada de /l/:  
vocalização de /l/ e coalescência.



Fonte: as autoras

Observe-se que, ao longo da evolução da sequência /al/, até chegar à forma resultante na vogal oral /o/, a partir da manifestação fonética velarizada de /l/, ocorreram dois processos fonológicos: a vocalização da líquida lateral, na Figura 2, que redundou no ditongo decrescente [aw], e a coalescência, na Figura 4, com a reunião de dois segmentos distintos de cuja junção nasceu um terceiro, diferente daqueles que lhe deram origem.

## 2.2 Sobre as sequências /ɔl/ - /ol/ - /ul/

Das sequências /ɔl/, /ol/, /ul/ derivou-se, no sistema linguístico do francês, a vogal alta dorsal arredondada /u/, o que é ilustrado, no Quadro 3, a seguir, com a presença da vocalização da líquida lateral. A vogal /u/, portanto, foi derivada a partir de três ordens de vocábulos: (i) oriundos de formas latinas com a sequência /ɔl/<sup>17</sup>; (ii) decorrentes de vocábulos latinos com a sequência /ol/; e (iii) provenientes de formas latinas com a sequência /ul/. Reitera-se que o francês tratou o glide dorsal [w] como parte da sequência V+glide. Dessa sequência resultaram ditongos decrescentes<sup>18</sup>, que redundaram, em etapa subsequente, na vogal /u/, no presente caso. No quadro apresentado, encontram-se, à esquerda, as formas oriundas do latim vulgar; na coluna central, estão dispostas as formas evolutivas francesas, seguidas, à direita, das formas correspondentes em português. As formas marcadas por asteriscos são construções abstratas. Seguem-se comentários.

Quadro 3 - Vocalização da líquida lateral em coda medial: resultado: /u/.

	/ _ l	
	<b>ɔł &gt; ow &gt; u</b>	
poll(i)cem	pɔłce > powce > <i>pouce</i> [ɔł] > [ow] > /u/	'polegar'
mol(e)re	mɔłdre > mowdre > <i>moudre</i> [ɔł] > [ow] > /u/	'moer'
	<b>ol &gt; ow &gt; u</b>	
coll(o)care	colcher > cowcher > <i>coucher</i> [ol] > [ow] > /u/	'deitar'
soll(i)citare	solcier > sowcier > <i>soucier</i> [ol] > [ow] > /u/	'inquiatar'

<sup>17</sup> Considera-se aqui a forma dos vocábulos no latim vulgar.

<sup>18</sup> Kolovrat (1923: 100-1) defende que esses ditongos não teriam coexistido, tampouco apresentado o mesmo timbre. O autor assevera que a passagem de /ɔl/ e /ol/ às formas correspondentes com a vogal alta dorsal labial, nas Figuras 7 e 8, explicam-se da mesma sorte que o grupo *al*, ou seja, o fenômeno ocorre, no francês antigo, em virtude do timbre duro do *l* diante de consoante.

	<b>uʔ &gt; oʔ &gt; ow &gt; u</b>	
ascul <u>te</u> re	escuʔter* > escoʔter > écowter > éc <u>ou</u> ter [uʔ] > [oʔ] > [ow] > /u/	‘escutar’
cul <u>p</u> abilis	cuʔpable* > coʔpable > cowpable > <u>cou</u> pable [uʔ] > [oʔ] > [ow] > /u/	‘culpado’
cul <u>t</u> ellus	cuʔtel* > coʔtel > cowteau > <u>cou</u> teau [uʔ] > [oʔ] > [ow] > /u/	‘faca’
ful <u>g</u> ur	fuʔdre* > foʔdre > fowdre > <u>fo</u> udre [uʔ] > [oʔ] > [ow] > /u/	‘relâmpago’
sulfur	suʔfre > soʔfre > sowfre > <u>sou</u> fre [uʔ] > [oʔ] > [ow] > /u/	‘enxofre’

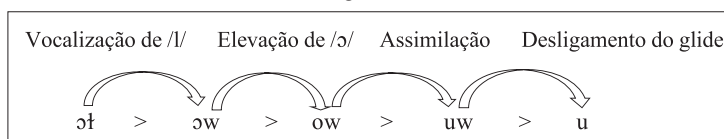
Interessam ao presente trabalho as formas evolutivas francesas constituídas de (i) sequência de vogal média-baixa posterior, /ɔ/, seguida da líquida lateral, /l/, [ʔ] (cf. p[ɔʔ]ce > p/u/ce ‘pouce’; m[ɔʔ]dre > m/u/dre ‘moudre’); (ii) sequência de vogal média-alta posterior, /o/, seguida da líquida lateral, /l/, [ʔ] (cf. c[ɔʔ]cher > c/u/cher ‘coucher’; s[ɔʔ]cier > s/u/cier ‘soucier’); (iii) sequência de vogal alta posterior, /u/, acrescida da líquida lateral, /l/, [ʔ] (cf. p[uʔ]mon > p[oʔ]mon > p/u/mon ‘poumon’).

Com respeito aos grupos de vocábulos que carregam a sequência /ɔl/ (i) e /ol/ (ii), em posição interna, os casos são restritos, mas neles o fenômeno de vocalização da líquida lateral velarizada operou da mesma forma que no grupo de casos mais extensos, em /ul/. No grupo de vocábulos com a sequência originalmente formada de vogal alta dorsal, /u/, acrescida da líquida lateral velarizada, [ʔ], em posição medial, daí /ul/ (iii), a vogal teria sofrido, inicialmente, um processo de abaixamento, passando, assim, de vogal alta dorsal a vogal média-alta dorsal, cujo resultado foi a sequência /ol/. Tal sequência sofreu, então, os mesmos processos fonológicos que acometeram a sequência originalmente formada de vogal média-alta dorsal mais a líquida lateral – /ol/.

Para Kolovrat (1923), nas sequências /ɔl/ e /ol/, a líquida lateral foi vocalizada em [w] independentemente da tonicidade da vogal dorsal labial – média-baixa, /ɔ/ (cf. pill(i)ce(m) > pɔʔce > pouce ‘polegar’) ou média-alta, /o/ (cf. sollicitare > soʔcier > soucier ‘inquietar’; collocare > coʔcare

> *coucher* 'deitar') ou da consoante que se lhes seguisse. A sequência V+glide foi base para três novos processos: coalescência, nas Figuras 3-4 apresentado, e assimilação e desligamento do glide, nas Figuras 6-8 a seguir. Do processo pelo qual a vogal /o/ assimilou a altura do glide resultou a sequência [uw], em que os dois elementos vocálicos compartilham os mesmos traços vocálicos. Por essa razão, entra em operação o OCP<sup>19</sup>, que desliga o glide [w], fazendo resultar a vogal /u/. Seguem-se comentários.

Figura 6 - Evolução do grupo /ɔl/ a partir da manifestação fonética velarizada de /l/: vocalização de /l/, elevação da vogal média-baixa labial, assimilação e desligamento do glide.



Fonte: as autoras

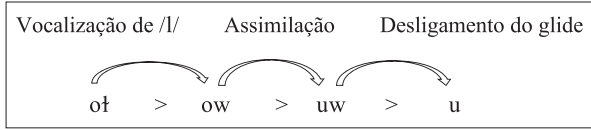
Na Figura 6, têm-se os três processos ocorridos na passagem de /ɔl/ a /u/. Primeiramente, acredita-se ter ocorrido a vocalização da líquida lateral velarizada, que passou a apresentar, então, os traços de um glide dorsal [w]<sup>20</sup>, em virtude de tal elemento compartilhar a rima com uma vogal. Esse processo teria desencadeado a elevação da vogal média-baixa dorsal ao patamar de vogal média-alta, por assimilação do traço [-ab3] que integra os traços de altura do glide [w] – [-ab1, -ab2, -ab3]. Por fim, houve a assimilação de /o/ aos traços de [w], formando a sequência [uw], o que faz com que, por OCP, seja desligado o glide [w] – portador dos mesmos traços segmentais da vogal /u/ – e resulte um só segmento vocálico, a vogal alta dorsal /u/, como núcleo da sílaba.

Na Figura 7, a seguir, têm-se os processos fonológicos que desencadearam a passagem de /ol/ a /u/, no francês.

<sup>19</sup> OCP (*Obligatory Contour Principle* – Princípio do Contorno Obrigatório) opera nas línguas regulando a tendência geral à evitação de elementos adjacentes iguais (McCarthy 1988).

<sup>20</sup> A formalização de tal processo na geometria está apresentada na Figura 2. Considerando-se idêntica a estrutura interna de [u] e [w], esta também mostrada na Figura 2, em obediência ao OCP, o resultado da sequência [uw] teria de conduzir à forma /u/.

Figura 7 - Evolução do grupo /ol/ a partir da manifestação fonética velarizada de /l/:  
vocalização de /l/, assimilação e desligamento do glide.

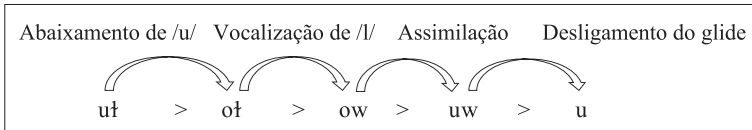


Fonte: as autoras

Com respeito à sequência *ol*, portanto, ocorreu a vocalização da líquida lateral, formando-se, com isso, o ditongo decrescente [ow]. Na etapa seguinte, surgiu a vogal /u/, pelo fato de a vogal /o/ ter sido assimilada aos traços de [w], daí [uw], sendo que, com a operação de OCP, há o apagamento do glide, do que resultou a vogal alta dorsal /u/.

Na Figura 8, apresentam-se os processos fonológicos responsáveis pela mudança de /ul/ a /u/, no francês.

Figura 8 - Evolução do grupo /ul/ a partir da manifestação fonética velarizada de /l/:  
abaixamento de /u/, vocalização de /l/, assimilação e desligamento do glide.



Fonte: as autoras

Relativamente à sequência *ul*, dois caminhos foram observados, de um resultando a vogal /u/, do outro, a vogal frontal arredondada /y/.

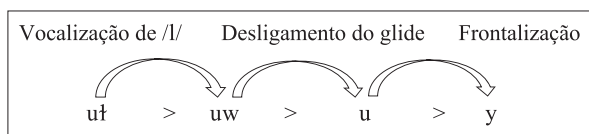
Na Figura 8, está representado o processo de que resultou a vogal /u/ a partir da sequência *ul*: ocorreu, num primeiro momento, o processo de abaixamento da vogal alta posterior, dando origem, assim, à sequência /ol/, cujas etapas subsequentes já foram apresentadas e ilustradas neste texto, na Figura 7. Repete-se, assim, a partir da etapa /ol/ [ot], o mesmo que se processou com a sequência original *ol*, ou seja, deu-se a vocalização da líquida lateral, de que resultou a formação do ditongo decrescente [ow]. Na etapa seguinte, ocorreu o processo de assimilação da vogal /o/ aos traços do glide dorsal, fazendo surgir a sequência [uw] e, por desligamento do glide,

resultou a vogal alta posterior/dorsal /u/.

Houve casos, entretanto, em que a sequência *ul* resultou na vogal alta frontal arredondada /y/ – extremamente raros são os exemplos em que esse fenômeno ocorreu. De acordo com Bourciez (1921: 39), “a combinação do ‘l’ vocalizado com a vogal ‘u’ não deixou rastro em francês”<sup>21</sup> (cf. *pul(i)cem* > *p[uɫ]ce* > \**p[uw]ce* > \**p[u]ce* > *p/y/ce* *puce* ‘pulga’; *pul(i)cella* > *p[uɫ]cella* > \**p[uw]celle* > \**p[u]celle* > *p/y/celle* *pucelle* ‘donzela’). Nesse segundo caminho evolutivo, a sequência *ul* teria apresentado a vocalização da lateral em coda, de que se seguiu [uw] e, em sendo a sequência [uw] formada de dois elementos iguais adjacentes, desde logo operou o OCP, em consonância com o que já foi representado nas Figuras 6, 7 e 8, implicando o desligamento do glide. A vogal /u/ resultante, nesse caso particular, no entanto, não se manteve: por ser /u/ o núcleo original da sílaba, essa vogal (mesmo resultante da sequência [uw]) recebeu, em alguns itens lexicais, o tratamento evolutivo dado, no francês, para a vogal /u/ do latim – a vogal longa do latim clássico –, ou seja, foi convertida na vogal alta frontal arredondada /y/, pelo processo de frontalização, com a manutenção do traço [+arredondado].

Apresentam-se, na Figura 9, os processos fonológicos responsáveis pela mudança de /ul/ a /y/, no francês.

Figura 9 - Evolução do grupo /ul/ a partir da manifestação fonética velarizada de /l/:  
vocalização de /l/, desligamento do glide e frontalização.



Fonte: as autoras

Observe-se que, de acordo com o processo evolutivo representado na Figura 9, a vogal alta frontal arredondada /y/ apenas seria resultante da vogal alta posterior /u/ quando derivou da sequência vogal + lateral em coda silábica como um segundo caminho evolutivo, já que também essa

<sup>21</sup> O mesmo ocorreu, segundo Bourciez (*op.cit.*), com o ‘l’ vocalizado ao compor sequência com a vogal ‘i’ (cf. *filius* > francês antigo *fi[i]s* > *fi[w]s* > *fi[j]s* *fil*s ‘filho’).

sequência *ul*, seguindo outro percurso, poderia resultar na vogal /u/, como representado na Figura 8.

No caso mostrado na Figura 9, estaria sendo seguida a tendência geral da evolução da vogal /u/ latina, pois, segundo relata a literatura (cf. Bourciez 1921; Nyrop 1899), na passagem do latim ao francês, a vogal /u/ evoluiu para a vogal /y/ francesa. Esse movimento se deveria a uma mudança de ponto de articulação, sem a influência de vizinhança linguística, ou seja, tratar-se-ia de uma mudança fonética de caráter “independente” (e.g. lat. *dura* > fr. *dure* d/y/re ‘duro’ *ap.* Bourciez<sup>22</sup> 1921:15; lat. *muru* > fr. *mur* m/y/r ‘muro’; lat. *crucifigere* > fr. *crucifier* cr/y/cifier ‘crucificar’ (*ap.* Nyrop 1899:113).

Nos restritos casos em que da sequência *ul* houve o resultado da vogal alta frontal arredondada, ainda se poderia cogitar que a presença de /y/ teria sido desencadeada por assimilação do traço [coronal] de consoante(s) adjacente(s) ou de vogal coronal da sílaba seguinte (cf. pul(i)cem > fr. ant. p[uʔ]ce > \*p[uw]ce > \*p[u]ce > p/y/ce *puce* ‘pulga’) – seria, então, um processo de coronalização da vogal /u/, motivada por assimilação, com a manutenção do traço [+arredondado].

Para Kolovrat (1923: 120), no entanto, teria havido outro percurso evolutivo; para o autor, a única possibilidade é que a líquida lateral, na sequência de /u/ + /l/, haja sido perdida em etapa precedente à redução de /u/ à vogal frontal arredondada /y/ (cf. /ul/ > /u/ > /y/). Segundo Kolovrat, nos empréstimos que carregam /l/ em sílabas cujo núcleo é /y/ ou /i/, o elemento consonantal /l/ se mantém (cf. *filtre* ‘filtro’; *occulte* ‘oculto’; *sultan* ‘sultão’). Em outras palavras, a líquida lateral em coda não desencadeia, nesses casos, qualquer processo fonológico que possa envolver as vogais tautossilábicas /y/ ou /i/. Defende também que, mesmo em palavras vernáculas, também a vogal frontal arredondada /y/ não derivou de sequência de vogal + lateral e que, portanto, a natureza fonológica da vogal alta frontal contrastiva /y/ não estaria relacionada à presença da líquida em coda.

<sup>22</sup> Da mesma forma, relata o autor (*op. cit.*: 13), o ‘a’ final do latim “se conservou no francês sob a forma enfraquecida de ‘e’ surdo [ɛ] (que se tornou mudo mais tarde)”. Em outras palavras, a vogal ‘a’ do latim, em posição final de palavra, sofreu um processo de enfraquecimento na passagem do latim ao francês, de que resultou um ‘e’ surdo [ɛ], o qual mais tardiamente na história da língua francesa tendeu a deixar de ser pronunciado – daí ‘e mudo’ (e.g. lat. *dura* > fr. *dure* ‘duro’).

Retomando-se o processo evolutivo mostrado na Figura 9 e defendendo-se a possibilidade de mudança  $u > y$ , bem como se reconhecendo os numerosos casos, na sincronia do francês, com a sequência /y/, interpreta-se que a presença de /y/, em exemplos como *p/y/ce* (*puce* ‘pulga’) e *v/y//gaire* (*vulgaire* ‘vulgar’), teria motivações ainda diferentes daquelas acima listadas: em palavras como *p/y/ce*, a vogal frontal arredondada /y/ derivou de /u/, resultante da vocalização da lateral em coda, conforme já foi dito acima e mostrado na formalização na Figura 9. Por outro lado, a vogal /y/ presente em palavras que mantêm a lateral, como *v/y//gaire*, deriva da assimilação, pela vogal /u/, da propriedade [-posterior] da consoante lateral em coda; nestes casos, deixou de haver a velarização da lateral em coda<sup>23</sup> e a lateral, não se vocalizando, manteve-se nas palavras. Este fato aconteceu em momento subsequente na evolução – esses diferentes tipos de processo evolutivo de que decorreu a vogal frontal arredondada /y/ são evidências de movimentos da língua francesa em diferentes estágios de sua história, conforme pode ser atestado no Quadro 4, em vocábulos franceses provenientes de formas latinas.

Observa-se, no Quadro 4, que a maioria das palavras elencadas entrou no francês em etapa subsequente ao séc. XIII, o que poderia, então, ser tomado como evidência de um movimento evolutivo diferente, relativamente à sequência de /u/ + líquida lateral, àquele encetado em período precedente da história da língua.

Destaca-se que os processos aqui relatados como motivadores da emergência, no francês, tanto da vogal /u/ como da vogal /y/, têm a sua naturalidade respaldada nos traços que integram a estrutura interna dos segmentos neles envolvidos.

---

<sup>23</sup> O processo de velarização da lateral na coda silábica do francês seguiu o movimento do latim vulgar rumo à vocalização, mas cessou em determinado momento da evolução da língua.

Quadro 4 – A frontalização da vogal arredondada na sequência /yl/.

latim > francês	Datação/período	
adultus > <i>adulte</i>	1380	‘adulto’
auscultare > <i>ausculter</i>	1500	‘auscultar’
bulbus > <i>bulbe</i>	1400	‘bulbo’
consultare > <i>consulter</i>	1410	‘consultar’
crepusculum > <i>crépuscule</i>	1596	‘crepúsculo’
culpabilis > <i>culpabilité</i>	1791	‘culpa’
cultura > <i>culture</i>	1509	‘cultura’
difficultas > <i>difficulté</i>	séc. XIII/1200	‘dificuldade’
fulgurare > <i>fulgurer</i>	1862	‘fulgurar’
fulminare > <i>fulminer</i>	1300	‘fulminar’
indulgentia > <i>indulgence</i>	1564	‘indulgência’
insultus > <i>insulte</i>	1535	‘insulto’
pulmonarius > <i>pulmonaire</i>	1572	‘pulmonar’
pulpa > <i>polpe</i> > <i>pulpe</i>	1503/1105	‘polpa’
resultatum > <i>résultat</i>	1589	‘resultado’ (< lat. escolas.)
sculptura > <i>sculpture</i>	1380	‘escultura’
sepulcrum > <i>sépulcre</i>	séc. X/1100	‘sepulcro’
sepultura > <i>sépulture</i>	1100	‘sepultura’
tumultus > <i>tumulte</i>	1100	‘tumulto’
vulgaris > <i>vulgaire</i>	1552	‘vulgar’

### 2.3 Sobre a sequência /il/

A sequência *il* foi uma das formas de cuja evolução se originou a vogal frontal média-alta arredondada /ø/, no inventário fonológico do francês (cf. *it* > *et* > *ew* > *ø*), a qual era desconhecida do latim vulgar, conforme já referido neste texto. Apresenta-se, no Quadro 5, a seguir, a vocalização da líquida lateral, a partir da sequência /il/. Seguem-se comentários.

Quadro 5 - Vocalização da líquida lateral em coda – resultado: /ø/.

	/ __ l	
	<b>it &gt; et &gt; ew &gt; ø</b>	
fil trum	<b>fe tter &gt; fewtre &gt; fe utre</b> [eʔ] > [ew] > /ø/	'feltro'
fil (i)caria	<b>fi tgeria &gt; fet gère &gt; fewgère &gt; fe ugère</b> [iʔ] > [eʔ] > [ew] > /ø/	'samambaia'
capil l(o)s	<b>cheve ts &gt; cheve ws &gt; cheve ux</b> [eʔ] > [ew] > /ø/	'cabelo'

No quadro evolutivo apresentado, a vogal frontal arredondada /ø/ – ortograficamente representada *eu* –, oriunda da sequência francesa /eʎ/, somente passou a vigorar no sistema vocálico do francês ao redor do fim do séc. XII, de acordo com Nyrop (1899: 201). Observe-se, no Quadro 5, que a primeira forma evolutiva francesa, coletada a partir de dados disponíveis em obras especializadas, constitui-se de vogal média-alta frontal, /e/, seguida de líquida lateral velarizada, [ʔ], decorrente de um processo de abaixamento (cf. *fi|tgeria* > *fet|gère*; etc.), assim como já ocorria em latim vulgar<sup>24</sup>. Na etapa seguinte, tem-se a aplicação do processo de vocalização da líquida lateral, de que resulta um glide dorsal, em posição subsequente à vogal /e/, formando-se, assim, um ditongo decrescente, [ew]. Na etapa subsequente, os integrantes do ditongo são submetidos a um processo de coalescência, cujo resultado é a vogal frontal arredondada, /ø/. Por fim, registra-se que os casos de vocalização da lateral velarizada na sequência /eʎ/ são restritos, mas neles o mencionado processo de vocalização operou da mesma sorte que nos demais, mostrados nos Quadros 2 e 3.

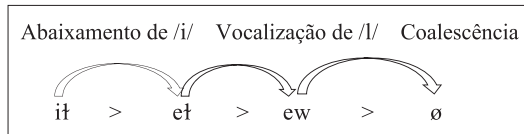
Com respeito à origem das vogais frontais arredondadas do francês, a literatura (Nyrop 1899; Bourciez 1921) observa que da presença, no latim vulgar, da vogal média-baixa posterior /ɔ/, em sílaba aberta e tônica, resultam (i) a vogal média-baixa frontal arredondada /œ/, em sílaba fechada (cf. *hora* > *oure* > *heure* h/œ/re 'hora'; *buove* > *buef* > *bœuf* b/œ/f

<sup>24</sup> Na obra de Quintilhano, encontra-se esse mesmo processo de abaixamento da vogal (resultado de dissimilação de altura) em *magester* e *leber* por *magister* e *liber* (cf. Brachet 1872: 157-8).

'boi'), e (ii) a vogal média-alta frontal arredondada /ø/, em sílaba aberta (cf. *focus* > *feu* f/ø/ 'fogo'; *bœuf*s b/ø/fs 'bois'). Segundo Bourciez (*op. cit.*), a primeira mudança por que passou /ɔ/ foi ter-se ditongado em [wɔ] e, após, no início do séc. XI, ter passado ao ditongo crescente [wɛ], ou antes [yɛ], escrito *ue*. A passagem do ditongo [yɛ] a um som simples – a vogal média-baixa frontal arredondada /œ/<sup>25</sup> –, desconhecida do latim e da maioria das línguas românicas, teria ocorrido, provavelmente, ao final do séc. XII. Não há, portanto, nos autores, referência ao fato de a procedência de vogais frontais arredondadas no francês estar vinculada a sequências originalmente compostas por vogal + líquida lateral.

Na Figura 10, apresentam-se os processos fonológicos envolvidos na evolução do grupo *eI* (<il) até chegar à vogal /ø/. Seguem-se comentários.

Figura 10 - Evolução do grupo /iI/ a partir da manifestação fonética velarizada de /I/: abaixamento de /i/, vocalização de /I/ e coalescência.



Fonte: as autoras

Incidem três processos fonológicos sobre a sequência /iI/. O primeiro é o abaixamento da vogal frontal alta /i/ para a média-alta correspondente, /e/, de que resulta o grupo [eʔ]<sup>26</sup>. O processo seguinte consiste na vocalização de /I/, daí resultando o glide dorsal. Por compartilhar a rima com a vogal-núcleo, forma-se, por conseguinte, o ditongo decrescente [ew], a vogal-núcleo /e/ seguida do glide dorsal [w]. O ditongo decrescente [ew] passa, então, por um processo de coalescência, de cuja aplicação assomou novo segmento vocálico no sistema do francês – a vogal média-alta frontal arredondada /ø/ – par opositivo da vogal frontal média-alta /e/. Os traços que integram as estruturas internas dos segmentos é que oferecem as condições para a aplicação dos processos que resultaram, de acordo com o exposto, na vogal /ø/ na fonologia do francês.

<sup>25</sup> A manifestação escrita da vogal média-baixa frontal arredondada /œ/ pode ser *eu, oeu, oe, oue*

<sup>26</sup> Observe-se que o processo de abaixamento de vogal alta precedendo /I/ já havia sido registrado na sequência *ul* – veja-se a Figura 8.

### 3. Sumarização

É chegado o momento de retornar-se às questões anteriormente formuladas – (i) a possibilidade de haver causas fonético-fonológicas responsáveis pelo desencadeamento da mudança de um sistema sem vogais frontais arredondadas (latim) para outro com essas vogais (francês) e (ii) a indagação a propósito de a inserção de vogais arredondadas, frontais ou posteriores, ser capaz de mostrar alguma relação com a história da lateral posvocálica na língua francesa.

É com base em Martinet (1955) que se responde afirmativamente ao primeiro questionamento, em (i). Segundo o autor, dentre as causas principais da mudança fonética na história das línguas estão:

- (a) a assimetria dos articuladores, o que impede o estabelecimento de um sistema fonológico simétrico;
- (b) as forças internas dentro do sistema – a atração de um fonema por um espaço vazio no sistema, o que o faz mudar a sua articulação e seus traços distintivos a fim de preencher aquele espaço como membro de uma oposição sistemática.

Entende-se que essas duas causas poderiam estar subjacentes à emergência de vogais frontais arredondadas no francês. A causa em (a) responderia pela frontalização de /u/, com a manutenção do seu traço de arredondamento, como resultado da “assimetria dos articuladores” na cavidade bucal – o grande espaço fonético (articulatório e acústico) entre as vogais altas [u] e [i] oferecia as melhores condições para a sua ocupação por outro segmento, o que fez emergir a vogal frontal arredondada [y], que foi fonologizada. O maior espaço articulatório na parte frontal da boca pode justificar o número maior de oposições vocálicas anteriores/coronais em relação à região posterior. Por sua vez, a causa em (b) concerniria à efetiva ação que houve das forças internas dentro do sistema, o que ocorreu em dois eixos: primeiramente, emergiram segmentos pelo fato de que o espaço das vogais frontais arredondadas estava vazio no sistema; em segundo lugar, tendo em vista a presença da vogal frontal arredondada alta /y/, a analogia e a busca de simetria do sistema levaram à emergência das contrapartes

arredondadas das outras duas alturas de vogais frontais: da vogal frontal arredondada média alta /ø/ e da vogal frontal arredondada média baixa /œ/. Esta causa se vê fortalecida ao observar-se a simetria das vogais frontais arredondadas com as vogais frontais não arredondadas: os dois tipos de vogais frontais ocupam simetricamente três alturas, preenchendo de forma paralela os espaços fonético-fonológicos, conforme rerepresentado em (3), originalmente Quadro 1.

(3) Sistema vocálico do francês

i	y	u
e	ø	o
ɛ	œ	ɔ
a		

Com respeito ao segundo questionamento em (ii), ou seja, se a inserção de vogais arredondadas, frontais ou posteriores, no francês mostraria alguma relação com a história da lateral posvocálica na língua, tem-se também uma resposta afirmativa. Essa seria, aliás, uma terceira motivação, que também está vinculada às forças internas dentro do sistema (causa em (b) acima registrada), desta vez relacionada à lateral posvocálica: diz respeito à efetiva ação das propriedades (representadas por traços) do(s) segmento(s) originário(s) da mudança. Dito de outra forma, a velarização da lateral (/l/ → [ʎ]), sob a visão autosegmental dos traços, atribuiu, à consoante lateral, uma articulação secundária vocálica, portadora do traço [+arred] ([labial]), na Figura 1 apresentado. O traço [+arred], por sua vez, foi mantido no processo subsequente de vocalização ([ʎ] → [w]), dado que, à luz da Geometria de Traços, a vocalização da lateral posvocálica é responsável por desligar a articulação primária consonantal, mantendo a articulação vocálica (com o traço [labial]/[+arredondado]), ilustrado na Figura 2 – a manutenção desse traço responde também por vogais arredondadas, frontais ou posteriores.

Propõe-se que essas três causas, em ação interativa, desencadearam a mudança fonético-fonológica que, do latim, resultou nas vogais

arredondadas, frontais ou posteriores, presentes na fonologia do francês contemporâneo. No Quadro 6, sintetizam-se os movimentos diacrônicos e as suas motivações com foco exclusivo nas vogais frontais arredondadas da língua.

Quadro 6 – Surgimento de vogais frontais arredondadas no francês.

1º) /y/ - **vogal alta frontal arredondada**

Motivações:

a) o espaço fonético (frontal) maior

→ Frontalização de /u/

b) as forças internas dentro do sistema

→ Derivado de sequência V+lateral ( $V+I > V+t > V+w > V$ )

2º) /ø/ - **vogal média-alta frontal arredondada**

Motivações:

a) o espaço fonético (frontal) maior

b) as forças internas dentro do sistema

→ Analogia com /y/ e busca de simetria

→ Derivado de sequência V+lateral ( $V+I > V+t > V+w > V$ )

3º) /œ/ - **vogal média-baixa frontal arredondada**

Motivações:

a) o espaço fonético (frontal) maior

b) as forças internas dentro do sistema

→ analogia com /y/ e /ø/ e busca de simetria

→ Derivado de sequência V+lateral ( $V+I > V+t > V+w > V$ )

Tratando-se agora exclusivamente das três vogais frontais arredondadas da língua francesa, pelo Quadro 6 observa-se que apresentariam mais de uma fonte de surgimento no sistema, sendo prevalentes as motivações determinadas pela configuração fonológica do inventário vocálico da língua.

Quanto às informações relativas à estrutura silábica, na passagem do latim ao francês, tem-se que sílabas fechadas por líquida lateral, foneticamente uma lateral velarizada, tautossilábica a uma vogal, resultaram, ao longo do processo evolutivo do francês, em uma só vogal – /u/, /o/, /y/ ou /ø/. Com respeito à sílaba aberta no latim, a literatura (Nyrop 1899; Bourciez 1921) relata que da presença de vogal média-baixa posterior /ɔ/, em sílaba aberta e tônica, resultaram no francês (i) a vogal média-baixa frontal arredondada /œ/, em contexto de sílaba fechada (e.g. latim *hora* > francês antigo *oure* > francês moderno *heure* h/œ/re [œR] ‘hora’; latim *buove* > francês antigo

*buef* > francês moderno *bœuf* b/œ/f [bœf] ‘boi’); e (ii) a vogal média-alta frontal arredondada /ø/, em contexto de sílaba aberta (e.g. latim *focus* > francês *feu* f/ø/ [fø] ‘fogo’; *bœufs* b/ø/fs [bø] ‘bois’). Outrossim, a literatura também aponta que a vogal alta posterior /u/ do latim foi interpretada no francês como a vogal alta frontal arredondada /y/ (e.g. latim *durum* > francês *dur* d/y/r [dyʀ] ‘duro’; latim *maturum* > francês *mûr* m/y/r [myʀ] ‘maduro’; latim *pluma* > francês *plume* pl/y/me [plym] ‘pluma’). Note-se, por fim, que nos três últimos casos citados (/ɔ/ > /œ/; /ɔ/ > /ø/ e /u/ > /y/), ou seja, naqueles em que o surgimento de vogais frontais arredondadas apresenta outra origem que não a da sílaba fechada por líquida lateral, a mudança de ponto de articulação operou num movimento de frontalização: o que era posterior no latim, passou a frontal no francês. Segundo Nyrop (*op. cit.*: 112), em se tratando de mudanças que envolvem ponto de articulação, essas ocorrem ordinariamente de trás para frente. Os movimentos são todos expressos com base em propriedades dos segmentos que são expressas, na fonologia da língua, por meio de traços distintivos.

#### 4. Considerações finais

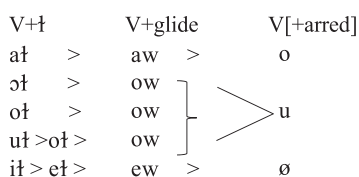
No francês primevo, havia um grupo de vogais frontais não arredondadas, herdadas do latim vulgar, /i, e, ε/, a que se somava uma série posterior/dorsal de vogais arredondadas/labiais – também disponíveis no latim vulgar – /u, o, ɔ/, além da vogal baixa /a/. As sequências de vogal e líquida lateral em coda, vigentes no francês antigo, ao menos inicialmente, tinham como núcleo vogais já integradas ao sistema. Essas sequências segmentais, com a evolução, puderam formar novos segmentos na língua – as vogais frontais arredondadas –, cuja emergência e fonologização se deu, pois, em etapa posterior da diacronia da língua francesa, no francês antigo. O presente estudo traz, como uma das fontes da presença de vogais arredondadas do francês, tanto das posteriores /u/, /o/, como das frontais arredondadas, a sequência, em latim, de vogal + líquida lateral em posição medial de palavra.

As formas resultantes da vocalização da líquida lateral velarizada em coda de sílaba (/l/ [ɫ] → [w]), ilustradas no Quadro 2 (formas et /o/), no

Quadro 3 (formas em /u/) e no Quadro 5 (formas em /ø/), preservam o traço [+arred] do elemento vocalizado – o glide dorsal, [w] –, em coda silábica interna à palavra. Observe-se que a manifestação fonética da lateral /l/ como velarizada está na origem de processos de que resultaram vogais fonológicas no francês; os movimentos evolutivos aqui destacados evidenciam a articulação entre um fenômeno de natureza fonética e outro de natureza fonológica. Além da articulação entre fonética e fonologia, os fatos diacrônicos aqui destacados também mostram articulação entre diferentes tipos de unidades fonológicas – traços, segmentos e sílabas –, confirmando o que defende Sen (2015), com respeito ao fato de a vocalização da lateral em coda consistir em uma das evidências da influência indireta protagonizada pela estrutura silábica na mudança sonora.

Em (4), a seguir, apresenta-se a sequência de V + /l/ em seus diferentes estágios, até resultar na vogal arredondada oriunda do processo de vocalização da lateral posvocálica no francês.

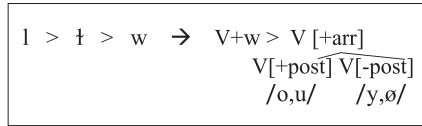
#### (4) Evolução da sequência V + /l/ no francês



O traço [+arred], em (4), emerge nas formas resultantes da vocalização de /l/ em coda (/l/ → [w]) como decorrente dos processos de **posteriorização ou velarização + vocalização**, sofridos pela lateral posvocálica. Porém, a evolução não cessa nessa etapa: o glide dorsal [w] foi interpretado, na fonologia do francês, como parte da sequência **V + glide**, ou seja, como unidade constituinte de um ditongo, ensejando o desencadeamento de dois outros processos – coalescência e assimilação – que deram origem a vogais arredondadas, já pertencentes ao sistema e conhecidas do latim vulgar (/u/ e /o/), além de frontais arredondadas – inexistentes no latim vulgar.

Sintetiza-se, na Figura 11, o percurso evolucionar por que passou a lateral posvocálica em francês.

Figura 11 - O francês e a evolução da lateral em coda.



Fonte: as autoras

No francês, observa-se, portanto, que a lateral em coda silábica, em seu processo evolutivo, passou por um processo de vocalização de /l/ ( $l > \text{ɫ} > w$ )<sup>27</sup>, de que resultou a sequência de  $V+w$  ( $l > \text{ɫ} > w \rightarrow V+w$ ). Ao prosseguir a marcha, a lateral posvocálica /l/ foi tratada, no sistema linguístico do francês, como parte da sequência **V+glide** ( $V+w$ ), a qual sofreu processo e resultou numa só vogal ( $V+w > V [+arred]$ ), seja ela [+post] ([+post, +arred] – /o/, /u/) ou [-post] ([-post, +arred] – /y/, /ø/).

Os segmentos vocálicos resultantes da vocalização da lateral posvocálica foram vogais arredondadas e, mais do que isso, os efeitos da vocalização da lateral em coda redundaram, no francês, em novas incorporações vocálicas: as vogais frontais arredondadas.

Por fim, defende-se que a adoção do enfoque teórico assumido por Clements & Hume (1995), que trabalha com a visão não-linear do funcionamento do componente fonológico das línguas – a partir do tratamento autosssegmental dos traços que compõem a estrutura interna dos segmentos e a formalização dos processos que envolvem segmentos e traços por meio de uma representação geométrica –, foi de fundamental importância não só por permitir formalizar a contento a trajetória por que passou a lateral posvocálica até a sua vocalização e os subsequentes desdobramentos de vogal-núcleo + glide dorsal, no processo evolutivo do francês, como também por permitir explicar uma das fontes de que se originaram as vogais arredondadas posteriores /o/, /u/ e também as vogais frontais arredondadas no sistema vocálico do francês.

<sup>27</sup> No português do Brasil o processo evolutivo da lateral em coda foi até essa etapa, ou seja, parou no glide dorsal [w]. Isso ocorreu porque tratou o /l/ posvocálico como segmento independente da vogal, assim, o espaço fonológico de /l/ foi preservado (cf. a[ɫ]ma > a[ɫ]ma > a[w]ma ‘alma’).

## REFERÊNCIAS

- Bisol, Leda. 1994. Ditongos derivados. *D.E.L.T.A.* 10 (n.º especial): 123-140.
- Bourciez, Edouard. 1921. *Précis historique de phonétique française*. Paris: Librairie Klincksieck. Retirado da Internet em 16/02/2017: <https://archive.org/details/prcishistoriqu00bouruoft>
- Brachet, Auguste. 1872. *Dictionnaire étymologique de la langue française*. Paris: J.Hetzel. Retirado da Internet em 15/02/2017: <https://ia600201.us.archive.org/19/items/DictionnaireetymologiqueDeLaLangueFrancaise>
- Camara Júnior, Joaquim Mattoso. 1976. *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão.
- Clements, George Nick. 1985. The geometry of phonological features. *Phonology Yearbook*. 2: 225-252.
- Clements, George Nick. 1989. A unified set of features for consonants and vowels. Ms, Cornell University.
- Clements, George Nick. 1991. Place of articulation in consonants and vowels: a unified theory. *Working Papers of the Cornell Phonetics Laboratory*. 5: 77-123. Ithaca: Cornell University.
- Clements, George Nick & Hume, Elizabeth. 1995. The internal organization of speech sounds. In: Goldsmith, John A. (Ed.). *The handbook of phonological theory*. Cambridge: Blackwell, 245-306.
- Grevisse, Maurice. 1993. *Le bon usage*. Paris: Duculot.
- Kolovrat, George de. 1923. *Étude sur la vocalisation de la consonne l dans les langues romanes*. Paris: Jouve & Cie. Éditeurs. Retirado da Internet em 14/02/2017: [https://ia802606.us.archive.org/4/items/tudesurlavocal00kolouoft/tudesurlavocal00kolouoft\\_bw.pdf](https://ia802606.us.archive.org/4/items/tudesurlavocal00kolouoft/tudesurlavocal00kolouoft_bw.pdf)
- Martinet, André. 1955. *Économie des changements phonétiques. Traité de phonologie diachronique*. Berne: Éditions A. Francke S.A.
- Mascaró, Joan. 1989. On the Form of Segment Deletion and Insertion Rules. *Probus*. 19(1): 31-61.
- McCarthy, John. 1988. Feature geometry and dependency: A review. *Phonetica*. 45. Disponível em: [http://works.bepress.com/john\\_j\\_mccarthy/35/](http://works.bepress.com/john_j_mccarthy/35/)
- Müller, Danielle. 2011. Developments of the lateral in occitan dialects and their romance and crosslinguistic context. *Linguistics*. Université Toulouse le Mirail – Toulouse II. English. <NNT : 2011TOU20122>. <tel-00674530>

- Nunes, José Joaquim. 1989. *Compêndio de gramática histórica portuguesa*. 9.<sup>a</sup> ed. Lisboa: Clássica Editora [1951].
- Nyrop, Kr. 1899. *Grammaire historique de la langue française*. Paris: A. Picard & Fils. Retirado da Internet em 16/11/2016: <http://gallica.bnf.fr/ark:/12148/bpt6k240021>
- Quednau, Laura R. 1994. A vocalização variável da lateral. *Letras de Hoje*. 29(4): 143-151.
- Recasens, Daniel. 2012. A phonetic interpretation of sound changes affecting dark // in Romance. In: M. J. Solé. & D. Recasens (Eds.). *The Initiation of Sound Change. Production, Perception and Social Factors*. Amsterdam: John Benjamins, 57-76.
- Ségéral, Philippe & Scheer, Tobias. 2015. De la communauté de destin des voyelles en syllabe fermée dans l'évolution du latin vulgaire à l'ancien français. *Diachroniques*. 5: 105-152.
- Sen, Ranjan. 2015. *Syllable and Segment in Latin*. Nova Iorque: Oxford University Press.
- Tranel, Bernard. 1987. *The sounds of French: an introduction*. Cambridge: Cambridge University Press.
- Veiga, Cláudio. 1965. *Gramática nova do francês*. 4.<sup>a</sup> ed. São Paulo: Editora do Brasil S/A.

# O grafema <x> e o dígrafo <ch>: um estudo longitudinal do desempenho ortográfico de crianças de três dialetos portugueses<sup>1</sup>

Jéssica Gomes

jgomes4@campus.ul.pt

*Faculdade de Letras da Universidade de Lisboa, Centro de Linguística da  
Universidade de Lisboa (Portugal)*

Celeste Rodrigues<sup>2</sup>

celesterodrigues@campus.ul.pt

*Faculdade de Letras da Universidade de Lisboa, Centro de linguística da  
Universidade de Lisboa (Portugal)*

## ABSTRACT

The possible preservation of the phonological distinction between the voiceless post-alveolar affricate /tʃ/ and the voiceless post-alveolar fricative /ʃ/ - corresponding to the graphic forms <ch> and <x>, respectively - in only a few dialects in the north of the country could be motivating different behaviors from children speaking different regional varieties of Portuguese in their first years of schooling. Since this study has not yet been done, this paper intends to analyze the written productions of Portuguese children from three linguistic areas of Portugal, represented here by children from the cities of Lisbon, Porto, and Chaves, about the spellings <x> and <ch>, to understand whether there is any relationship between the preservation or non-preservation of this phonological distinction and the children's orthographic performance. Thus, the written productions of students belonging to the three dialectal regions mentioned above will be analyzed based on data from the EFFE-On corpus (Rodrigues, Lourenço-Gomes, Alves, Janssen & Gomes 2015) with the phonological distinction between /tʃ/ and /ʃ/ having been preserved in some and neutralized in others. Therefore, it will be possible to measure the impact that the students' dialect can have on their orthographic performance, as well as perceive any differences in performance related to the grapheme <x> vs. the digraph <ch>. Two groups of children were created for the analysis: a group with typical and another with atypical phonological development children. The differences observed point to a need to deepen the discussion concerning the sound aspect of language in the classroom, beginning in the first years of schooling - especially during the literacy phase.

---

<sup>1</sup> A redação final deste artigo beneficiou da contribuição de dois revisores anónimos que solícitamente nos providenciaram algumas referências bibliográficas, o que muito agradecemos.

<sup>2</sup> Este trabalho beneficiou de apoio por parte do Centro de Linguística da Universidade de Lisboa - CLUL 2020: UIDB/00214/.

## KEY-WORDS

Orthographic knowledge development; Primary School; Dialectology; European Portuguese

## RESUMO

A possível conservação da distinção fonológica entre o segmento fonológico africado pós-alveolar surdo /tʃ/ e o fricativo pós-alveolar surdo /ʃ/ - correspondente às formas gráficas <ch> e <x>, respetivamente - em apenas alguns dialetos do norte do país poderá motivar comportamentos distintos das crianças das diferentes regiões dialetais do português nos primeiros anos de escolaridade. Uma vez que esse estudo ainda não foi feito, este trabalho pretende analisar a escrita de crianças portuguesas de três áreas linguísticas em Portugal, aqui representadas por crianças das cidades de Lisboa, Porto e Chaves no que respeita às grafias <x> e <ch>, com propósito de compreender se existe alguma relação entre a permanência (ou não) desta distinção fonológica e o desempenho ortográfico das crianças. São assim analisadas as produções escritas de estudantes pertencentes às três regiões dialetais acima indicadas, umas em que a distinção fonológica persiste e outras em que houve neutralização da oposição entre /tʃ/ e /ʃ/ com base em dados do corpus EFFE-On (Rodrigues, Lourenço-Gomes, Alves, Janssen & Gomes 2015). Assim, será possível perceber o impacto que o dialeto utilizado pelos estudantes pode ter no seu desempenho ortográfico, além de eventuais diferenças entre os desempenhos relativos ao grafema <x> vs. dígrafo <ch>. Dois grupos de crianças foram criados para a análise: um grupo de crianças com desenvolvimento fonológico típico e outro com desenvolvimento fonológico atípico. As diferenças observadas apontam para a necessidade de aprofundamento do tratamento da vertente sonora da língua em sala de aula, desde os primeiros anos de escolaridade - sobretudo, na fase de alfabetização.

## PALAVRAS-CHAVE

Desenvolvimento ortográfico; Ensino Básico; Dialectologia; Português Europeu

## 1. Introdução

### 1.1 Relações entre a oralidade e a escrita

A aquisição da oralidade e a aprendizagem do sistema ortográfico são processos cognitivos independentes, mas inter-relacionados. A escrita é aprendida mais tardiamente do que a oralidade, geralmente, em contextos de ensino explícito e formal, enquanto esta última é adquirida espontaneamente, mediante continuada exposição ao *input* linguístico envolvente (Lourenço-Gomes, Rodrigues & Alves 2016). Esta diferença nos dois processos de desenvolvimento das componentes, oral e escrita, da língua faz com que a criança tenha uma representação sonora da cadeia

falada, antes da aprendizagem da escrita, implicando que, posteriormente, estabeleça uma associação entre a representação sonora interiorizada e a imagem gráfica correspondente (Delgado-Martins 1996). A apropriação da escrita alfabética pela criança é um processo complexo e moroso, que envolve competências linguísticas como a complexidade da estrutura segmental (segmentos fonológicos e representações fonéticas) e a frequência com que esta ocorre no léxico. Envolve também competências metalinguísticas e condicionantes extralinguísticas como o conhecimento da norma ortográfica convencionada (única ou múltipla), o desenho dos caracteres utilizados, o grau de insistência do professor na estrutura gráfica, o grau de exposição da criança a material gráfico, tanto através da leitura como da escrita, etc. Em suma, a aprendizagem da escrita exige um conjunto de conhecimentos complexos e específicos (Carragher 1985; Cassar & Treiman 1997; Baptista, Viana & Barbeiro 2011; Rodrigues & Lourenço-Gomes 2016; Vale & Sousa 2017).

Os sistemas de escrita apresentam, geralmente, diferentes conjuntos de símbolos e uma maior ou menos regularidade das regras com as quais representam as unidades sonoras das línguas. De entre os vários sistemas de escrita possíveis, a língua portuguesa adota um sistema de escrita alfabético em que, na maior parte dos casos, cada grafema representa um segmento fonológico somente (Velooso 2005). Contudo, não é isto que se verifica sempre para o português. As línguas dotadas de escrita alfabética têm sido agrupadas em dois grandes grupos: línguas com escrita fonemicamente transparente e línguas com escrita fonemicamente opaca, embora o nível de transparência/opacidade seja gradual. O primeiro grupo apresenta uma correspondência maximamente regular, sistemática e biunívoca entre os segmentos fonológicos e os símbolos gráficos discretos. O segundo grupo, por sua vez, é caracterizado pela ausência desta representação biunívoca entre fonema-grafema (*ibidem*). Segundo Seymour (1997), do grupo de línguas transparentes, fazem parte línguas como o finlandês, o italiano e o espanhol. No extremo oposto, isto é, nas línguas opacas, encontra-se o inglês. O português situa-se numa posição intermédia, ou seja, é considerado uma língua com grau intermédio de opacidade (Castro & Gomes 2000). Isto acontece porque o sistema ortográfico do português apresenta alguns exemplos de opacidade fonémica. Um desses exemplos é o caso do grafema

<x>, que pode assumir diferentes formas fonéticas, não sendo todas elas contextualmente motivadas. Um outro exemplo de opacidade é o facto de a grafia apresentar marcas etimológicas, como é o caso do dígrafo <ch> (Velooso 2005). Assim, pode-se concluir que as representações gráficas, em análise no presente trabalho, são exemplos de opacidade fonémica (pelo menos, para algumas crianças, posto que <ch> e <x> representam um único segmento fonológico) e, conseqüentemente, são estruturas potencialmente geradoras de algumas dificuldades na aprendizagem da escrita.

A representação ortográfica das sibilantes, em geral, é complexa e potencialmente problemática, no Português Europeu (PE), uma vez que estas apresentam diferentes formas de registo ortográfico, quer através de grafemas, quer através de dígrafos (Freitas, Rodrigues, Costa & Castelo 2012: 132). No caso do segmento /ʃ/ em posição de ataque (único segmento existente, para as estruturas em análise neste trabalho, no sistema fonológico do português, que é baseado no funcionamento dos dialetos centro-meridionais), este pode assumir as seguintes representações gráficas:

Tabela 1 - Representação gráfica de /ʃ/

Fonema	Fone	Grafia	Exemplo
/ʃ/	[ʃ]	<x>	<b>x</b> aile; pe <b>x</b> e
		<ch>	<b>ch</b> ave; b <b>ch</b> o

Apesar de a representação gráfica de /ʃ/ ser considerada arbitrária por alguns falantes do PE, nomeadamente os das variedades centro-meridionais, é possível identificar contextos em que esta é previsível. Scliar-Cabral (2003) identifica para o português do Brasil (PB), mas igualmente válidos para o PE, os seguintes contextos previsíveis para a ocorrência do grafema <x>: quando estiver dentro de palavra depois dos ditongos [ej] (pe**x**e), [ow] (rou**x**inol) e [aj] (ca**x**ia) e quando estiver no interior da palavra antecedido da vogal nasal [ẽ] (en**x**ame). Nos restantes contextos, não existe uma regra que defina qual a representação ortográfica a ser adotada para o /ʃ/, em ataque silábico.

A literatura acerca da representação ortográfica de /ʃ/, para o PE, é relativamente escassa, sendo que a maior parte da investigação é relativa ao PB. Miranda (2010), num estudo sobre o erro ortográfico no PB, conclui que, nos casos em que há troca entre <x> e <ch>, 80% dos resultados mostram que a opção escolhida foi <ch>. A autora justifica este valor através da influência da frequência de uso de <x> e <ch>, pois, no léxico, há prevalência de palavras com <ch>. Por sua vez, num estudo sobre o desempenho ortográfico das crianças em relação às consoantes fricativas (Schier, Bert & Chacon 2013), verificou-se que, na escrita, se registavam erros ortográficos que envolvem a presença/ausência de vozeamento, tendo sido encontrados dados em que as crianças representaram as formas gráficas <x> e <ch>, que correspondem a /ʃ/, com as formas gráficas representantes de /ʒ/, nomeadamente, <j> e <g>.

O facto de poderem ser estruturas consideradas problemáticas, na aprendizagem da ortografia, já torna estas estruturas com <ch> e <x> interessantes para serem estudadas. Sem embargo, a motivação é ainda maior quando elas estão, também, relacionadas com a existência de variação dialetal, ou seja, quando existem diferentes formas de pronunciar uma palavra, consoante a área geográfica do falante, mas apenas uma forma de representação a nível ortográfico. Este ponto será discutido com mais detalhe no decorrer do trabalho.

## 1.2 A história do traço fonético: «pronúncia do ch como tx ou tch»

Em 1971, Cintra publica a “Nova proposta de classificação dos dialetos galego-portugueses”, identificando os dialetos conforme as áreas geográficas que adotam um conjunto comum de variantes linguísticas de certas estruturas. Das cinco características que o autor julgava serem facilmente perceptíveis pelos falantes e utilizadas pela própria comunidade linguística para determinar a origem geográfica dos indivíduos, destaca-se a descrição do traço fonético conhecido como «pronúncia do ch como tx ou tch», ou seja, quando a africada /tʃ/ em ataque (chá) contrasta fonologicamente com o segmento alvéolo-palatal /ʃ/ em ataque (xarope) (Rodrigues 2016). Este traço fonético, uma isófono binária, permite a identificação da área geográfica em que a distinção entre os fonemas /tʃ/ e /ʃ/ ainda se verificava daquelas em

que esta distinção fonológica já não existia, prevalecendo exclusivamente o segmento fonológico /ʃ/, tanto para o dígrafo <ch> como para o grafema <x>.

De acordo com Castro (2006), a oposição entre os dois fonemas já era registada desde a época do português medieval, com a africada palatal surda /tʃ/ a corresponder ao dígrafo <ch> e a fricativa palatal surda /ʃ/ a corresponder ao grafema <x>. Contudo, este último som ocorria em poucos contextos, sendo, como tal, pouco frequente. Assim, a africada /tʃ/ era mais produtiva, estando presente no galego e no português como evolução dos grupos consonânticos iniciais com L: PL- (PLUVIA) CL- (CLAVE) FL- (FLAMMA). Estes três grupos consonânticos tiveram em galego-português uma evolução peculiar, que terminou na africada /tʃ/ (*ibidem*: 25) e se mantém ainda hoje em dialetos do Norte de Portugal. Considera-se que, até ao século XVIII, a distinção fonológica permaneceu. Contudo, já na segunda metade do século XVII, os gramáticos começaram a referir confusões frequentes entre esses dois tipos, originadas por um processo de perda da africção do /tʃ/ que levou à sua convergência com a fricativa pós-alveolar, em particular, nas regiões centro e sul do país (Pérez 2015). De acordo com Teyssier (1982: 45), é nos textos manuscritos desse século que se começam a observar confusões entre as formas gráficas <ch> e <x>, dando como exemplo *xão* por *chão*. Segundo Pérez (2015), Lisboa teve um papel fundamental para explicar a difusão desta inovação, uma vez que a pronúncia da capital favoreceu a sua dispersão para outras regiões do país. Assim, a variação do Sul causou uma mudança linguística em diversas variedades da língua portuguesa. Porém, nem todas as regiões foram atingidas por esta perda da africção. Em qualquer caso, verifica-se que a ortografia – mais conservadora do que a língua – contempla uma diferença entre <ch> e <x> que é neutralizada fonética e fonologicamente em, pelo menos, toda a área dos dialetos centro-meridionais, levando Teyssier (1982: 46) a afirmar que este será um facto que “ocasionará um número infinito de erros em todos os alunos de Portugal”.

No trabalho sobre os dialetos portugueses, Pérez (2015) observa a existência de uma grande heterogeneidade em relação à distribuição geográfica da conservação da africada. Desta forma, em PE subsistem dois sistemas com segmentos distintos, ou seja, o sistema fonológico dos dialetos

setentrionais, com uma oposição entre as duas unidades, e o sistema dos dialetos centro-meridionais, onde esta distinção fonológica não existe, prevalecendo exclusivamente o fonema /ʃ/, tanto para o dígrafo <ch> como para o grafema <x> (Segura 2013: 95-100). Em relação às regiões das cidades analisadas no nosso artigo e considerando a classificação clássica de Cintra (1971), a região de Chaves e a região do Porto pertencem aos dialetos setentrionais. Por sua vez, a região de Lisboa pertence aos dialetos centro-meridionais. Quanto à realização da africada, de acordo com Castro (2006), esta africada permanece apenas nas montanhas do Alto Minho, de Trás-os-Montes e das duas Beiras interiores. Como tal, e considerando as regiões observadas no presente artigo, só a zona de Chaves é que preserva a africada /tʃ/, visto que em Lisboa e no Porto este segmento já não faz parte do inventário fonológico dos falantes devido à sua fusão com /ʃ/ (representado ortograficamente como <x>).

### 1.3 Descrição linguística das estruturas alvo

Do ponto de vista fonológico, em Mateus & Andrade (2000: 13) apresenta-se o sistema das sibilantes como sendo mais complexo em determinadas regiões dialetais setentrionais, se comparado com o do PE padrão e o do PB. Neste ponto do trabalho, pretende-se fazer a descrição autosegmental das estruturas em estudo, utilizando a geometria de traços (Clements & Hume 1995) e a teoria de subespecificação de traços radical (Archangeli 1988), adaptadas para o português pelos autores acima mencionados.

A consoante /ʃ/ é uma fricativa que pode ser descrita como Coronal não anterior (Figura 1, abaixo) e, nas regiões centro-meridionais do português, quando a consoante se encontra em Ataque de sílaba, corresponde ao grafema <x> e ao dígrafo <ch>.

Em algumas das variedades setentrionais, preserva-se, como já foi referido, uma oposição fonológica entre as duas unidades, a fricativa e africada. Isso pode promover a mais fácil aprendizagem dos dois grafemas, <x> e <ch>, por parte das crianças dessas regiões. A consoante /tʃ/ não se encontra disponível no inventário fonológico do PE padrão. É uma consoante africada que podemos representar como se encontra na Figura 2. Nos dialetos setentrionais em que surge, esta consoante corresponde

ao dígrafo <ch> ao passo que o som [ʃ] corresponde a <x>. Como tal, é possível verificar que, nestas variedades, estes segmentos fonológicos distinguem, por exemplo, um par mínimo como *taxar* [tə'ʃɐɾ] e *tachar* [tə'tʃɐɾ].

Figura 1- Representação das características da consoante /ʃ/.

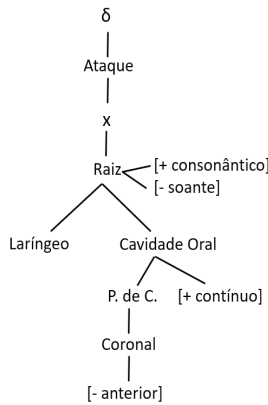
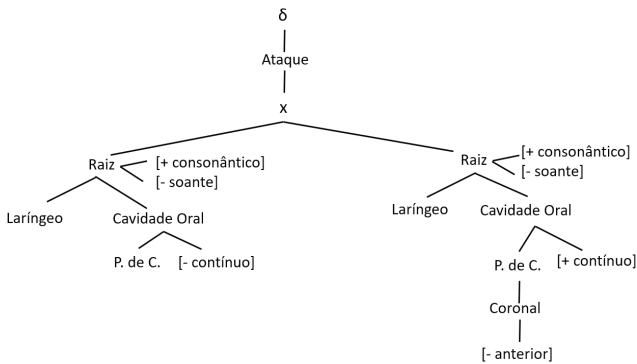


Figura 2 - Representação das características da consoante /tʃ/.

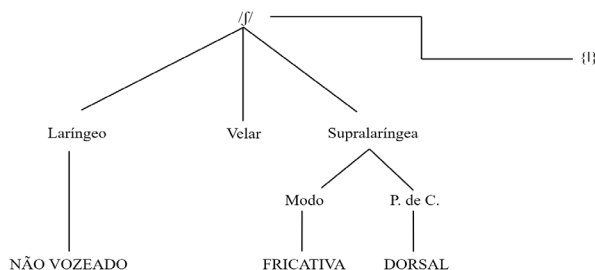


Como Mateus & Andrade (2000: 13) referem, estas distinções são recuperáveis através da ortografia.

No que respeita à estrutura fonológica das consoantes palatais em geral,

no entanto, a proposta da Figura 1 não é consensual. Veloso (2019) propõe que todas as consoantes palatais sejam classificadas com a introdução do Elemento //, enquanto propriedade autónoma, na sua estrutura – como representamos na Figura 3. Além disso, implica que o articulador interveniente na articulação destas consoantes seja o Dorsal, em vez do Coronal. A sua análise permite explicar de um modo natural o fenómeno de surgimento de uma semivogal no núcleo que antecede uma destas consoantes, típico de alguns dialetos setentrionais, como em: *cereja, fecha, telha, tenha* (reportado, por exemplo, por Silva 2017 e por Lao 2020).

Figura 3 - adaptada de Veloso (2019: 10).



#### 1.4 O erro ortográfico

Adotando uma perspetiva segundo a qual os desvios ortográficos constituem uma janela para o conhecimento que os estudantes já alcançaram da sua língua e da ortografia, evitaremos referir-nos a “erro ortográfico”, na medida em que acarreta uma conotação negativa, sublinhando conhecimentos não adquiridos. Com efeito, os desvios são características naturais das primeiras produções escritas das crianças, uma vez que a relação das crianças com a ortografia é relativamente recente. Desse modo, tentaremos perceber quais são as motivações desses desvios – aqui referidos como formas não convencionais (FNCs).

As FNCs são consideradas evidência empírica para aceder ao desenvolvimento do conhecimento ortográfico, sendo com frequência reflexo do conhecimento que os estudantes têm da fonologia da sua língua (Treiman 1998; Miranda 2009; Miranda & Matzenauer 2010; Veloso 2010).

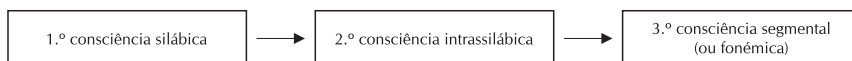
Para Meireles & Correa (2005), as FNCs refletem o nível de formação alcançado pelas crianças sobre o conhecimento ortográfico, ou seja, aquilo que a criança já adquiriu até àquele momento. Por sua vez, Carraher (1985) apresenta as FNCs não como estruturas aleatórias, mas como estruturas particulares de conceber a escrita, indicando uma compreensão que se vai modificando ao longo do percurso escolar das crianças. No trabalho de Veloso (2010), as primeiras produções escritas - em específico, as produções gráficas que revelam um conhecimento não consolidado de todas as convenções ortográficas da língua que ocorrem no 1.º e 2.º anos de escolaridade - são descritas como manifestações externas conducentes à caracterização da língua-l, permitindo uma descrição indireta do conhecimento fonológico.

Existem várias tipologias de erros ortográficos, mas é apenas relevante tratar, neste contexto, os erros considerados *erros de uso* ou *erros grafemáticos* (Pinto 1998; Rio-Torto 2000). De acordo com as mesmas autoras, estes erros afetam somente a representação gráfica, e não a forma como a palavra deve ser pronunciada.

## 2. Consciência fonológica

Neste ponto do trabalho, pretende-se apresentar de forma sucinta informação sobre a consciência fonológica, uma das dimensões da consciência linguística mais importante e benéfica para a aprendizagem da escrita. Segundo Freitas, Alves & Costa (2007: 13), a consciência fonológica é “a capacidade de explicitamente identificar e manipular as unidades do oral e exprime-se através de comportamentos metafonológicos (consciência explícita)”. A consciência fonológica pode ser de vários tipos. Freitas et al. (2012: 182) organizam-na em quatro níveis: consciência de palavra, consciência silábica, consciência intrassilábica e consciência segmental ou fonémica, sendo a ordem de desenvolvimento dos tipos de consciência fonológica a seguinte:

Figura 4 - Ordem de desenvolvimento dos tipos de consciência fonológica (fonte: Freitas et al. 2012: 182).



Como é visível pela figura 4, estes tipos de consciência fonológica não se desenvolvem simultaneamente, sendo possível observar que o primeiro tipo de consciência fonológica a ser desenvolvida é a consciência silábica, seguida da consciência intrassilábica e, por fim, a consciência segmental. As autoras consideram que este último tipo de consciência fonológica se desenvolve com a aprendizagem do código alfabético. Desta forma, defendem a existência de uma interação entre a ortografia e este tipo de consciência linguística, segundo a qual o conhecimento da ortografia pode remodelar em alguns aspetos o conhecimento fonológico.

Segundo Veloso (2010), existe relação de influência recíproca entre consciência fonológica e o conhecimento ortográfico nas fases iniciais da aprendizagem da norma ortográfica nas línguas com escrita alfabética. Assim, a consciência fonológica é importante porque se estabelece uma relação recíproca entre ela e a escrita, na qual sempre pode haver influência de uma sobre a outra (Delgado-Martins 1996; Freitas 2004; Freitas et al. 2007; Correia 2009; Freitas et al. 2012). Por fim, a estimulação da consciência fonológica faz com que as crianças passem a reconhecer e poder manipular as unidades fonológicas. Posteriormente, isso vem a facilitar a manipulação dos grafemas, enquanto unidades discretas, na fase inicial da alfabetização.

Descrita uma das dimensões com mais influência na aprendizagem da escrita, passa-se agora à descrição da aquisição da sibilante /ʃ/ em ataque somente, uma vez que não dispomos de informação acerca da aquisição da africada.

### 3. A aquisição da sibilante /ʃ/

Neste ponto do trabalho, pretende-se apresentar o que está descrito na literatura a que tivemos acesso acerca da aquisição da sibilante /ʃ/, em Ataque silábico. Divide-se a descrição em duas partes, consoante o perfil

dos alunos no *corpus* utilizado: grupo de desenvolvimento típico e grupo de desenvolvimento atípico.

### 3.1 Desenvolvimento típico

Para as crianças com desenvolvimento fonológico típico, em Ataque não ramificado, as sibilantes são a segunda classe de segmentos consonânticos a emergir, ou seja, só são adquiridas depois das consoantes oclusivas e das consoantes nasais (Freitas 1997; Costa 2010; Amorim 2015; Matzenauer & Costa 2017). Segundo Costa (2010: 27-28), a ordem de aquisição das fricativas em posição de ataque tende a ser a seguinte:

(1) [+anterior] >> [-anterior]

Desta forma, as primeiras fricativas a serem adquiridas pelas crianças são as Labiais, ou seja, /f/ e /v/. Posteriormente, segue-se a aquisição das fricativas coronais, sendo as [-anteriores] adquiridas em momentos diferentes pelas duas crianças em que foi possível apurar uma ordem de aquisição, a Joana e a Inês. A Inês adquiriu o /s/ e o /ʃ/ aos 2;11 e a Joana só adquiriu /s/ aos 4;0 e o /ʃ/ aos 4;2. No trabalho de Mendes, Afonso, Lousada & Andrade (2009/2013; 2013), é descrita a seguinte ordem de aquisição dos segmentos fricativos coronais em posição de ataque silábico:

(2) /s >> ʃ >> (3;6) >> z >> ʒ (4;0 - 4;6)/

Contudo, a aquisição destes segmentos consonânticos parece depender de indivíduo para indivíduo, existindo uma grande variação individual na aquisição das sibilantes, em ataque, no PE. Podemos igualmente perceber isto quando observamos os diferentes resultados obtidos por Ramalho (2017):

(3) /s (3;0 - 3;11) >> ʃ >> z >> ʒ (4;00 - 4;11)/

Como é possível verificar, nos trabalhos referidos, as primeiras sibilantes a serem adquiridas são as não-vozeadas, logo, é possível concluir que o traço [± vozeado] influencia a ordem de aquisição destes segmentos. Esta evidência

foi também observada em Amorim & Matzeunaer (2014), num estudo sobre as fricativas coronais, onde se verificou que a assimetria na aquisição dos segmentos coronais é determinada pela sonoridade do segmento-alvo, sendo que as crianças começam por estabelecer o valor não marcado do traço [vozeado] e só mais tardiamente estabilizam o valor positivo desse traço. Relativamente ao /ʃ/, existe uma diferença significativa em relação ao seu período de aquisição, nos dois estudos, uma vez que, em Mendes et al. (2009/2013; 2013) este está estabilizado aos 3;6 anos de idade, mas, em Ramalho (2017), está estabilizado aos 4;00 – 4;11 anos de idade, ficando, assim, clara a influência da variação individual na estabilização da aquisição de /ʃ/ (embora as diferenças observadas possam também ser atribuídas às características da amostra ou às opções metodológicas utilizadas).

Por fim, Matzenauer & Costa (2017) descrevem que a estratégia de reconstrução mais produtiva para o fonema /ʃ/, em PE, é a substituição por [s]. Assim, *chapéu* é produzido como [sa'pɛw]. Por sua vez, Amorim (2015), num estudo sobre a aquisição de segmentos consonânticos em falantes do Grande Porto, verificou que as estratégias de reconstrução mais produtiva para o fonema /ʃ/ envolvem a substituição do traço [anterior] e do traço [vozeado]. Neste sentido, a autora verificou ocorrências do tipo: *chaminé* produzido como [ʒɐmi'ne], existindo uma substituição do traço [vozeado] de [-vozeado] para [+vozeado]; e *bolacha* produzido como [bu'lasɐ], existindo uma substituição do traço [anterior] de [-anterior] para [+anterior]. Por fim, no mesmo trabalho, foram observadas substituições pontuais do segmento /ʃ/ por segmentos oclusivos. Assim, registaram-se ocorrências como *chapéu* produzido como [ta'pɛw], substituindo desta forma o traço [+contínuo] por [-contínuo] e o traço [-anterior] por [+anterior].

Relativamente à aquisição do segmento consonântico /ʃ/ em coda silábica, Almeida, Costa & Freitas (2010) realçam o facto de certas fricativas poderem surgir primeiro em posição de Coda e só mais tarde em posição de Ataque. Na mesma pesquisa, a fricativa /ʃ/ já era produzida pela criança, em posição de Coda silábica, conforme o alvo, aos 1:9 anos: *meus* ['mewʃ]. Por sua vez, a fricativa /ʃ/ só mais tarde surge em posição de Ataque, aos 2:9 anos de idade: *chorar* [ʃu'rari]. Desta forma, é visível a existência de variação de comportamento para a aquisição das sibilantes consoante a posição ocupada na sílaba.

Posto isto, verificou-se que o segmento fricativo em estudo é adquirido antes dos 4 anos, logo, não deveria causar dificuldades na fase de aprendizagem da escrita. Se algumas dificuldades existem na fase de alfabetização, estão relacionadas com a existência de representações ortográficas múltiplas para um mesmo som. E que, por sua vez, só tem obrigatoriamente representações múltiplas nas regiões em que tenha havido neutralização da oposição entre /ʃ/ e /tʃ/.

Acabada a descrição da aquisição de /ʃ/, para as crianças com desenvolvimento fonológico típico, segue-se agora a descrição da aquisição de /ʃ/ em crianças com desenvolvimento fonológico atípico.

### 3.2 Desenvolvimento atípico

As crianças com perturbações no desenvolvimento fonológico seguem a mesma ordem de aquisição das sibilantes que é seguida pelas que têm um desenvolvimento fonológico típico. Porém, o período de estabilização destes segmentos consonânticos é mais tardio, quando comparado com as crianças com um desenvolvimento fonológico típico. Ramalho e Freitas (2019) apresenta a seguinte ordem de aquisição das consoantes sibilantes para o desenvolvimento fonológico com perturbações:

(4) /s, ʃ/ >> /z, ʒ/

Segundo o trabalho das autoras, os dados observados revelaram uma grande diversidade nos comportamentos verbais da sua amostra. Contudo, foi possível observar uma maior taxa de acerto nas fricativas dento-alveolares, quando comparadas com as fricativas palatais.

Quando considerado o traço [ $\pm$  vozeado], existe um paralelismo com as crianças de desenvolvimento fonológico típico, isto é, os segmentos não vozeados são os primeiros a serem produzidos conforme a estrutura alvo (Ramalho 2017; Ramalho & Freitas 2019).

Relativamente às estratégias de reconstrução das estruturas utilizadas pelas crianças, as autoras identificaram a substituição por [s] como a estratégia mais produtiva.

Em relação à Coda fricativa, Ramalho & Freitas (2019) verificaram um efeito na variável constituência silábica, observando a ordem de aquisição Ataque > > Coda.

Para finalizar esta parte do trabalho, conforme Preston, Hull & Edwards (2013) observaram na sua investigação, as crianças com um desenvolvimento fonológico atípico apresentam alterações ao nível da consciência fonológica. Como já foi dito, esta é uma das dimensões da consciência linguística mais importante para a aprendizagem da ortografia. Deste modo, crianças com perturbações no desenvolvimento fonológico podem evidenciar mais dificuldades em aprender a escrever. Assim, será expectável que as crianças sem perturbações no desenvolvimento fonológico tenham um melhor desempenho ortográfico, quando comparadas com as crianças com perturbações no desenvolvimento fonológico.

#### 4. Hipóteses de interpretação das FNCs

O objetivo deste trabalho é o de analisar o desempenho ortográfico das crianças do 2.º ano e do 4.º ano de escolaridade do 1.º Ciclo do Ensino Básico, em relação aos grafemas <x> e <ch>. Pretende-se, de igual forma, estudar e comparar o desempenho das crianças com perturbações no desenvolvimento fonológico com crianças sem perturbações no desenvolvimento fonológico. Posto isto, e tendo em conta a revisão literária apresentada, as hipóteses deste trabalho são as seguintes:

- (1) Os textos das crianças de Chaves presentes no *corpus* vão apresentar menos FNCs do que os textos das de Lisboa e do Porto, uma vez que, no seu dialeto, existe a distinção fonológica entre /tʃ/ e /ʃ/.
- (2) As crianças com desenvolvimento fonológico atípico irão apresentar um desempenho ortográfico inferior, quando comparadas com as crianças de desenvolvimento fonológico típico, porque é previsível que apresentem perturbações no desenvolvimento fonológico e alterações ao nível da consciência fonológica.

## 5. Metodologia

Para a realização deste trabalho, utilizou-se o *corpus* EFFE-On: Escreves como falas- falas como escreves? (Rodrigues et al. 2015). É um corpus de acesso online gratuito, que contém dados de fala e escrita de crianças no primeiro ciclo de escolaridade. As composições dos estudantes são de dois géneros: descritivo e narrativo. O primeiro tipo de texto foi obtido através de imagens-cenário, tendo diversos temas, por exemplo: a cidade, a cozinha, entre outros. O segundo género textual foi obtido através de imagens em série, sendo que o tópico principal era sempre a “História da Bruxinha”. Às crianças, o investigador pedia sempre que produzissem uma composição descritiva e uma composição narrativa. Todos os textos utilizados nesta amostra correspondem apenas a crianças cujos textos seguem esta metodologia.

Dos dados disponíveis no EFFE-On, foram selecionados os textos de crianças do 2.º ano e 4.º ano. Os dados foram recolhidos em dois momentos diferentes, ou seja, numa primeira fase, foram recolhidos os textos quando as crianças tinham sete anos de idade e, numa segunda fase, quando as crianças tinham nove anos de idade. Tendo em conta os objetivos deste trabalho – comparar as regiões onde permanece a distinção fonológica entre /tʃ/ e /ʃ/ com as regiões onde já não existe esta oposição – e os dados disponíveis no *corpus* selecionado, tomou-se como base de investigação uma amostra de textos produzidos por crianças da cidade de Chaves, da cidade de Lisboa e da cidade do Porto. Cada região foi dividida em dois grupos: o primeiro grupo compreende as crianças com desenvolvimento típico (GDT) e o segundo grupo às crianças com desenvolvimento atípico (GDA). No primeiro grupo, foram incluídas, para além das crianças sem necessidades médicas, os estudantes que já tinham sido acompanhados em consultas de alergologia, uma vez que se considerou que tal não comprometeria o estudo. Por sua vez, o segundo grupo inclui as crianças com necessidades médicas como terapia da fala, otorrino, psicólogo, entre outras. Decidiu-se incluir as crianças com reporte de acompanhamento psicológico neste grupo, por se ter em conta que é muitas vezes na consulta de psicologia que são descobertas algumas perturbações cognitivas. Por fim, importa ressaltar que os informantes bilingues foram excluídos da amostra.

Na tabela 2, é apresentado o número de crianças por ano de escolaridade, por grupo e por cidade.

Tabela 2 - Número de participantes do 2.º ano e 4.º ano, dos GDT e GDA, das cidades de Chaves, Lisboa e Porto.

Ano de escolaridade	Chaves		Lisboa		Porto	
	GDT	GDA	GDT	GDA	GDT	GDA
2.º	26	19	45	10	26	7
4.º	27	15	65	16	30	10

Como já foi referido, a amostra de participantes é muito semelhante nas duas fases de recolha de dados, ou seja, os estudantes do 4.º ano são aproximadamente as mesmas crianças que participaram na primeira fase de recolha de dados no 2.º ano de escolaridade. A única exceção é a da cidade de Lisboa, que, para além da turma de 4.º ano, que coincide com a turma de 2.º ano, conta, igualmente, com dados de uma outra turma do 4.º ano de uma escola diferente.

O facto de a quantidade de participantes ser praticamente a mesma, nos dois anos de escolaridade, nas três regiões, permite obter uma amostra longitudinal do desempenho na escrita destas crianças.

Posteriormente, os dados foram tratados e analisados no programa informático *Excel*.

## 6. Análise dos dados

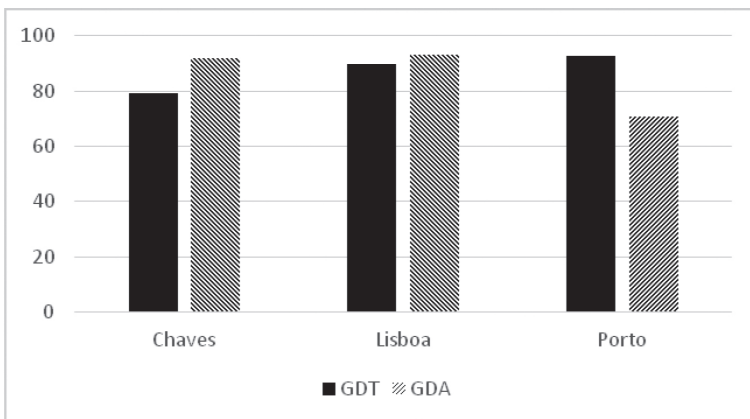
A presente secção tem como objetivo apresentar e discutir os resultados obtidos. Como tal, encontra-se dividida em três partes, a apresentação dos dados do 2.º ano e as respetivas taxas de sucesso, a comparação entre regiões e a comparação entre os grupos de desenvolvimento típico e desenvolvimento atípico. Na segunda parte desta secção, analisam-se e discutem-se as taxas de sucesso do 4.º ano de escolaridade, seguindo a organização já apresentada. Por fim, são tratadas as estratégias de reconstrução mais utilizadas, apresentando-se as possíveis razões que levaram as crianças a adotarem-nas.

### 6.1 Taxas de sucesso: 2.º ano

Nesta parte do trabalho, pretende-se expor as taxas de sucesso e discuti-las à luz das hipóteses já referidas e da informação existente na literatura. Para tal, importa introduzir a escala de desenvolvimento de Hernandorena (1990), pois, apesar de não ter sido criada para este fim, consideramos que ela nos pode ajudar a compreender o grau de domínio das estruturas ortográficas e identificar diferentes momentos do desenvolvimento infantil. A escala divide-se nas seguintes etapas de desenvolvimento em função da taxa de sucesso associada a cada etapa: (i) menos de 50% de correspondência produção/alvo: estrutura não adquirida; (ii) de 51% a 75% de correspondência produção/alvo: estrutura em aquisição; (iii) de 76% a 85% de correspondência produção/alvo: estrutura adquirida, mas não completamente estabilizada; (iv) de 86% a 100% de correspondência produção/alvo: estrutura adquirida e estabilizada.

Uma vez apresentada a escala da autora, iniciamos a análise com as taxas de sucesso para o dígrafo <ch>, dos dois grupos de crianças- GDT e GDA-, das três cidades em estudo.

Gráfico 1 - Produções corretas para o dígrafo <ch>, 2.º ano, dos GDT e GDA, das cidades de Chaves, Lisboa e Porto.



Como é possível observar, os estudantes das três regiões dialetais

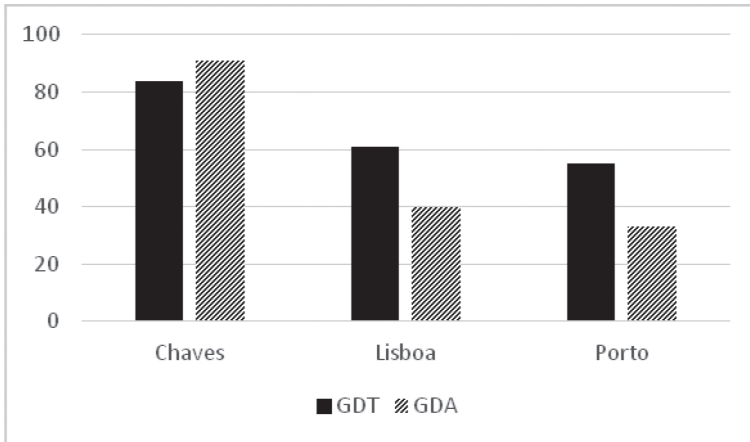
apresentam taxas de sucesso muito altas, isto é, todos os grupos apresentam taxas de sucesso superiores a 80%, para o dígrafo <ch>, sendo as únicas exceções o GDT de Chaves e o GDA do Porto. Nos grupos dos outros dialetos - Chaves e Lisboa - os GDA apresentam taxas de sucesso superiores, quando comparados com os GDT, registando-se a maior diferença na primeira cidade. Numa análise mais fina e recorrendo à escala anteriormente apresentada, pode-se constatar que a estrutura em análise se encontra estabilizada nos GDT das regiões dialetais de Lisboa e Porto, com taxas de sucesso de 89.9% e 92.9%, respetivamente. O mesmo se verifica para os GDA de Lisboa, bem como para o GDA da região de Chaves, com taxas de sucesso de 93% e 91.9%, respetivamente. Por fim, verifica-se que a estrutura se encontra adquirida, mas não completamente estabilizada para o GDT de Chaves e que se encontra em aquisição para o GDA da região do Porto.

Assim, contrariamente ao esperado, as crianças da cidade de Lisboa e do Porto do GDT registam um melhor desempenho, quando comparadas com as crianças da cidade de Chaves. A razão que cremos subjazer a este comportamento é a de as crianças destas duas regiões serem mais facilmente alertadas para a ortografia das estruturas <ch> e <x>, devido ao facto de elas serem pronunciadas da mesma maneira. Já as crianças do GDT de Chaves, onde pode haver ou não fusão dos dois segmentos fonológicos, apresentam mais FNCs porque estão em contacto com diferentes pronúncias para o dígrafo <ch>, na medida em que a simplificação da africada também se encontra presente no dialeto, sem que ninguém lhes chame a atenção para isso. Por outras palavras, o facto de, no seu dialeto, <ch> poder ser produzido com [tʃ] ou [ʃ], mas noutras variedades ser produzido como [ʃ], cria dúvidas aos estudantes de Chaves. Por conseguinte, esta instabilidade das realizações fonéticas do dígrafo <ch> é, segundo cremos, o principal motivo pelo qual este ainda não se encontra estabilizado na ortografia destes estudantes do GDT de Chaves<sup>3</sup>.

---

<sup>3</sup> Uma questão não muito discutida em trabalhos acerca da variação linguística do português prende-se com as possíveis diferenças registadas entre a oralidade nos centros urbanos e nos meios rurais. As recolhas dialetológicas são feitas preferencialmente em meios rurais, onde se espera que exista menos influência da variedade padrão da língua, e de modo diferente do que acontece nos meios urbanos.

Gráfico 2 - Produções corretas para o grafema &lt;x&gt;, 2.º ano, dos GDT e GDA, das cidades de Chaves, Lisboa e Porto.



No grafema <x> existe maior disparidade de resultados. Começando a análise pela cidade de Chaves, regista-se, tal como para o dígrafo <ch>, taxas de sucesso relativamente altas, com 83,6% para o GDT e 90,9% para o GDA. Deste modo, o GDA de Chaves mostra um melhor desempenho do que o GDT, estando a estrutura adquirida e estabilizada no primeiro grupo e adquirida, mas em estabilização, no segundo grupo. O facto de este GDA evidenciar taxas de sucesso mais altas do que o GDT pode ser explicado por haver uma intervenção específica do terapeuta da fala em relação ao /j/. Para além disso, as crianças do GDA parecem poder ser sensíveis exclusivamente à pronúncia.

Quanto às regiões de Lisboa e do Porto, estas apresentam taxas de acerto bastante mais baixas, sendo que, para nenhum destes dois grupos, a estrutura se encontra estabilizada, na ortografia. Os dados também mostram que, ao contrário de Chaves, nas duas outras regiões dialetais, o GDA regista um desempenho ortográfico inferior, quando comparado com o GDT. O facto de os estudantes de Chaves demonstrarem uma maior proficiência, quando comparados com os estudantes dos outros dialetos, mostra que, para este grupo de crianças, a correspondência entre <x> e /j/ é clara, não se verificando o mesmo para os outros estudantes. As taxas de sucesso de Lisboa e Porto para <x> correspondente a /j/ mostram que as crianças

apenas utilizam como estratégia de reconstrução o dígrafo nas duas cidades. Isto pode ser explicado, provavelmente, como refere Miranda (2010), pelo facto de serem pouco frequentes as palavras com <x>, comparativamente com as com <ch>. Assim, como as duas representações ortográficas correspondem ao fonema /ʃ/ em posição de Ataque, é natural que as crianças, em situações de dúvida, optem por representá-lo por intermédio do dígrafo <ch> por ser mais frequente no léxico. Dos estudantes destas duas cidades, o GDA do Porto é o que apresenta taxas de sucesso mais baixas. Isto parece evidenciar que para estas crianças a correspondência entre <x> e o segmento /ʃ/ não é totalmente transparente, o que pode ser explicado pelo facto de o grafema poder também ter uma correspondência com outros sons, nomeadamente, /z/ (e.g. *exame*). Desta forma, a taxa de sucesso relativamente baixa observada deve-se ao facto de existirem vários sons associados ao mesmo grafema que causam dúvidas aos estudantes. Alves (2012) já verificara que, quando os segmentos fricativos Coronal [+contínuo; -soante] mantêm uma relação grafémica multívoca, estes causam mais dificuldades na aprendizagem da escrita devido às suas relações grafémicas múltiplas. Em contrapartida, a autora verificou que os segmentos fricativos labiais com relações grafémicas transparentes, ou seja, <f> para [f] e <v> para [v], não apresentavam dificuldades às crianças, de um modo geral.

De seguida, são apresentados os dados para o 4.º ano de escolaridade.

## 6.2 Taxas de sucesso - 4.º ano

Nesta parte do trabalho, serão apresentadas e discutidas as taxas de sucesso para o 4.º ano, começando, uma vez mais, pelo dígrafo <ch> (veja-se o gráfico 3).

Como já se tinha verificado para os grupos do 2.º ano das regiões dialetais estudadas, as taxas de sucesso são muito altas para o dígrafo <ch>. Posto isto, e recuperando a escala anteriormente mencionada de Matzenauer (1990), observa-se que, para os GDT e GDA das regiões de Chaves e Lisboa, a estrutura está adquirida e estabilizada, bem como para o GDT da região do Porto. O único grupo que não apresenta a estrutura estabilizada é o GDA da região do Porto. Todavia, a sua taxa de sucesso (de 84.6%) não

deixa de ser elevada, estando próxima da estabilização nesse grupo. Se compararmos os dados deste grupo com os dados correspondentes do 2.º ano de escolaridade, observa-se uma maior maturidade por parte destes estudantes em relação à representação do dígrafo.

Gráfico 3 - Produções corretas para o grafema <ch>, 4.º ano, dos GDT e GDA, das cidades de Chaves, Lisboa e Porto.

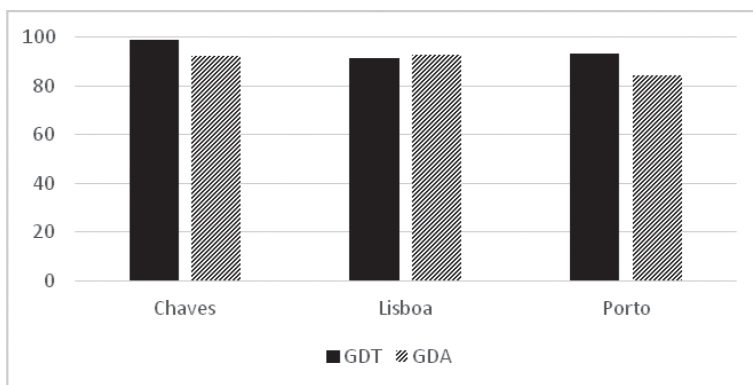
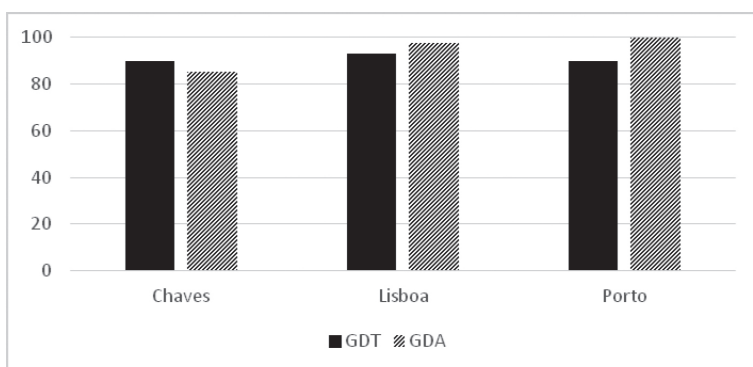


Gráfico 4 - Produções corretas para o grafema <x>, 4.º ano, dos GDT e GDA, das cidades de Chaves, Lisboa e Porto.



Ao contrário do 2.º ano de escolaridade, todos os grupos do 4.º ano de todas as regiões dialetais estudadas apresentam o grafema <x> adquirido e estabilizado, sendo a única exceção o GDA de Chaves. Contudo, a estrutura

encontra-se muito próxima da fase de estabilização mesmo nesse grupo. Isto demonstra que o segmento em análise, que era o que suscitava mais problemas no 2.º ano, deixou de os apresentar nos dados das três cidades. Em relação às taxas de sucesso das crianças dos diversos dialetos, não existem diferenças relevantes a registar. O mesmo acontece em relação à comparação entre os GDT e GDA, evidenciando, assim, que estas estruturas ortográficas já quase não são problemáticas para nenhum dos grupos. Isto significa que o número de FNCs diminui com o aumento da escolarização. Esta superioridade de desempenho já seria de esperar, naturalmente, uma vez que, como Santos, Freitas & Veloso (2014: 420) referem:

- (i) já não se esperam problemas de desenvolvimento fonológico que não tenham sido detetados e que não tenham sido acompanhados, neste nível etário;
- (ii) um percurso escolar mais longo resulta num conhecimento mais consolidado das regras ortográficas e num património lexical mais alargado.

Por último, importa dizer que, apesar das elevadas taxas de sucesso observadas, ainda há margem para evolução, uma vez que ainda se registam desvios nas produções escritas dos estudantes de todas as cidades no 4.º ano de escolaridade.

### 6.3 Estratégias de reconstrução

As estratégias apresentadas, nesta secção, referem-se aos três dialetos compreendidos neste estudo. Optou-se por organizar os dados desta forma, porque foram registadas FNCs dos mesmos tipos em todas as regiões. Desta forma, pode-se concluir que, quando existem dúvidas na representação gráfica de /ʃ/ em Ataque, as estratégias de reconstrução utilizadas pelas crianças são as mesmas.

#### 6.3.1 Estratégias de reconstrução: 2.º ano

A tabela 3 apresenta as estratégias de reconstrução que mais foram

recrutadas pelos estudantes do 2.º ano de escolaridade, em todos os dialetos, tanto para o dígrafo <ch> como para o grafema <x> (os exemplos podem ser consultados no apêndice A).

Tabela 3 - Estratégias de reconstrução para o dígrafo <ch> e grafema <x>, no 2.º ano, dos GDT e GDA, das regiões de Chaves, Lisboa e Porto.

Estratégia de reconstrução	Tipo de estrutura			
	<ch>		<x>	
	n	%	n	%
<ch>			176	87.5
<x>	30	46.1		
<j>	13	20	5	2.5
<s>	12	18.5	7	3.5
Outras produções <sup>4</sup>	10	15.4	13	6.5
<b>Total de FNCs</b>	65	100	201	100

Começando por uma análise geral, é visível que o grafema <x> apresenta um maior número de FNCs, no 2.º ano. Estes resultados são explicados, como já foi referido, pela maior frequência lexical de palavras com <ch> e pelo facto de o grafema <x> corresponder a outros sons, enquanto o dígrafo só tem correspondência com /ʃ/. Como tal, é natural que, num contexto de dúvida, a criança opte pelo dígrafo. Assim, registam-se 201 FNCs para o grafema <x> e 65 para o dígrafo <ch>, num total de 634 e 583 ocorrências, respetivamente.

Analisando agora as estratégias de reconstrução, pode-se observar que, para <ch>, a estratégia mais recrutada pelas crianças foi o <x>, com 30 ocorrências. Isto já seria de esperar, visto que ambos podem representar

<sup>4</sup> Neste ponto, foram incluídas, para o <ch>, estratégias como: <clh>, <f>, <h>, entre outras. Para o grafema <x>, foram incluídas estratégias como: <lh>, <xh>, <cl>, entre outras representações.

o fonema /ʃ/, em posição de Ataque. De seguida, o uso de <j> é a estratégia de substituição com maior expressividade, com 13 ocorrências, seguido pelo uso de <s>, com 12 ocorrências. Em relação ao grafema <x>, a estratégia de reconstrução que ocorre com mais frequência é o dígrafo <ch>, sendo que, das 201 FNCs registadas, 176 correspondem a esta representação ortográfica. No grafema <x>, também se verificam substituições por <s>, com 7 registos, e por <j>, com 5 registos, apesar de estas estratégias não serem tão recrutadas como são para o dígrafo <ch>.

A substituição de <x> e <ch> por <j> pode dever-se a uma falta de deteção da presença/ausência de vozeamento, como identificam os autores Schier *et al.* (2013) por parte de algumas crianças. Se as crianças não distinguirem corretamente os sons vozeados e não vozeados, ou seja, não discriminarem com precisão o traço [-vozeado] de /ʃ/, poderão associar o grafema <j> aos sons não vozeados, e não apenas a /ʒ/. A estratégia de reconstrução com o grafema <s> pode dever-se a uma generalização, efetuada pela criança, do uso deste grafema para representação gráfica do som [ʃ] em posição de Coda, quando é seguido de uma consoante [-vozeada] ou de pausa. Se, por outro lado, tivermos em conta que a mesma estratégia de reconstrução ocorre também no processo de aquisição da língua com alguma frequência (Oliveira 2004; Amorim 2015; Matzenauer & Costa 2017), é possível entendê-la na escrita destas crianças como a reutilização de uma estratégia disponível desde a aquisição da língua.

### 6.3.2 Estratégias de reconstrução: 4.º ano

Por sua vez, a tabela 4 regista as estratégias de reconstrução que mais foram ativadas para o 4.º ano de escolaridade, para todos os dialetos, tanto para o dígrafo <ch> como para o grafema <x> (os exemplos podem ser consultados no apêndice B).

Em relação ao 4.º ano de escolaridade, registou-se um número inferior de FNCs, face ao observado no 2.º ano de escolaridade. Analisando de forma detalhada, observam-se 39 FNCs, para o dígrafo <ch>, e 132 FNCs, para o grafema <x>, num total de 662 e 1585 ocorrências, respetivamente. Quanto às estratégias mais utilizadas pelos estudantes, verifica-se que estas são as mesmas que foram registadas para o 2.º ano de escolaridade. Contudo, ocorrem em menor número.

Tabela 4 - Estratégias de reconstrução para o dígrafo &lt;ch&gt; e grafema &lt;x&gt;, no 4.º ano, dos GDT e GDA, das regiões de Chaves, Lisboa e Porto.

Estratégia de reconstrução	Tipo de estrutura			
	<ch>		<x>	
	n	%	n	%
<ch>			104	78.8
<x>	19	48.7		
<j>	6	15.4	11	8.3
<s>	4	10.3	1	0.8
Outras produções <sup>5</sup>	10	15.4	16	7.5
<b>Total de FNCs</b>	39	100	132	100

Começando a descrição das estratégias de reconstrução pelo dígrafo <ch>, verificou-se, mais uma vez, que o recurso ao grafema <x> é a estratégia de substituição mais utilizada, pelos estudantes, com 19 ocorrências, registando-se assim uma diminuição em relação ao 2.º ano. Posteriormente, ocorrem as FNCs <j> e <s>, com 6 e 4 ocorrências, respetivamente.

No que se refere ao grafema <x>, a estratégia de reconstrução mais utilizada no 4.º ano é, por sua vez, o dígrafo <ch>, com 104 realizações. De igual modo, foi registada uma menor frequência da forma não convencional <ch> como estratégia de reconstrução face ao 2.º ano. Seguidamente, aparecem os grafemas <j>, com 11 ocorrências, e o grafema <s>, com apenas 1 ocorrência. No que se refere à estratégia de reconstrução <s>, que, de acordo com os dados de oralidade observados na aquisição, se esperaria ser frequentemente utilizada, os dados de escrita aqui analisados não corroboram essa expectativa, uma vez que <s> só ocorre uma vez.

<sup>5</sup> Neste ponto, foram incluídas, para o <ch>, estratégias como: <clh>, <g>, <z>, entre outras. Para o grafema <x>, foram incluídas estratégias como: <sc>, <c>, <ss>, entre outras representações.

## 7. Conclusão

Os dados apresentados neste artigo relativos ao 2.º ano de escolaridade evidenciaram que o dígrafo <ch> registou taxas altas de sucesso em todos os grupos, com exceção do GDA da região do Porto. Uma vez que existe alta frequência de palavras com <ch> no léxico do PE, é expetável que as crianças regularizem mais cedo a representação do dígrafo do que a de <x>. Quanto ao grafema <x>, as taxas de sucesso são consideravelmente mais baixas em todas as regiões dialetais consideradas, com a exceção de Chaves. Desta forma, pode-se concluir que, para este último grupo de estudantes (o de Chaves), a correspondência entre <x> e /ʃ/ é clara. Assim, e retomando a primeira hipótese - *os textos dos estudantes de Chaves presentes no corpus vão apresentar menos FNCs do que os textos de Lisboa e do Porto, uma vez que, no seu dialeto, existe a distinção fonológica entre /tʃ/ e /ʃ/* -, verificou-se que esta é confirmada para as crianças da região dialetal de Chaves, ou seja, estas apresentam um melhor desempenho ortográfico, em relação aos grafemas <x> e <ch>, quando comparados com os estudantes de Lisboa e do Porto. Isto é explicado pelo facto de os estudantes desta região estarem expostos ao contraste de pronúncia correspondente à diferença da grafia, mesmo que não o reproduzam na sua oralidade, como apontado por Freitas et al. (2012: 150). Consequentemente, é possível concluir que os estudantes da região de Chaves desenvolvem uma maior sensibilidade e consciência destas unidades segmentais, refletindo-se, desta forma, num melhor desempenho na ortografia nesta questão específica.

Relativamente à segunda hipótese - *as crianças do GDA irão apresentar um desempenho ortográfico inferior, quando comparadas com as crianças do GDT* -, observa-se que o dígrafo <ch> só é problemático para o GDA do Porto. Em relação ao grafema <x>, os GDAs apresentam sempre um desempenho inferior, se comparados aos GDTs. A única exceção são as crianças de Chaves, uma vez que o seu GDA regista melhores resultados do que o GDT. Posto isto, e retomando a segunda hipótese, esta é parcialmente confirmada para o grafema <x>, pois é a única estrutura que apresenta ainda dificuldade aos GDAs. Contudo, verifica-se que <x> é igualmente problemático para os GDTs, apesar de se registarem taxas de sucesso mais altas.

No que diz respeito aos dados do 4.º ano de escolaridade, as taxas de sucesso são mais altas, tanto para o grafema <x> como para o dígrafo <ch>, para todos os grupos dialetais. Como já foi referido, isto deve-se ao mais longo percurso escolar destas crianças e, como consequência, a uma maior consolidação das regras ortográficas do PE.

Quanto às estratégias de reconstrução, observou-se um comportamento uniforme, nas crianças dos diferentes dialetos. Por conseguinte, para o dígrafo <ch>, a estratégia de reconstrução mais utilizada foi o <x>, sendo que, para o grafema <x>, a estratégia de reconstrução mais utilizada foi o <ch>. Esta dúvida na forma como se deve representar graficamente o /ʃ/, em posição de Ataque, deve-se ao facto de ambos os símbolos gráficos poderem representar, ortograficamente, este fonema. Posteriormente, verificou-se que as FNCs mais registadas foram <j> e <s>. A forma não convencional <j> poderá dever-se ao facto de existir uma confusão em relação ao traço [± vozeado], contudo, esta hipótese de interpretação só poderia ser confirmada com observação de dados orais das crianças, o que aqui não foi possível apresentar. Desta forma, é possível que as crianças confundam /ʃ/ com /z/ e, como consequência, optem pelas formas gráficas que representam o segundo fonema, ou seja, <j>. A forma não convencional <s> é explicada pelo facto de este grafema representar a marca do plural em PE, entre outros [ʃ] lexicais (*rapaz, noz*). Assim, <s> representa muitas vezes o som [ʃ] em coda, podendo ser responsável pelas FNCs relacionadas com os símbolos gráficos <ch> e <x>. Em trabalhos futuros, seria vantajoso analisar dados da fala espontânea de crianças da região de Chaves, de forma a verificar se a distinção fonológica analisada neste estudo ainda se mantém nas novas gerações como, de certo modo, os dados aqui apresentados indicam que pode suceder. Além disso, seria ainda relevante estudar mais dados de escrita e dados de fala da região do Porto com a finalidade de avaliar a possível presença de semivogal antes dos sons estudados, uma vez que a inserção de <i>, aí, se registou, esporadicamente, como em *mexam* <\*meilhe> da criança 93 e *fecha* <\*feixa> da criança 115.

## REFERÊNCIAS

- Almeida, L., Costa, T. & Freitas, M. J. 2010. Estas portas [ˈɛtɐs ˈpɔtɐs] e janelas [nɛˈnɛs]: O caso das sibilantes na aquisição do Português europeu. In: A. M. Brito, F. Silva & A. Fiéis (Eds.). *Textos Seleccionados do XXV Encontro Nacional da Associação Portuguesa de Linguística*. Porto: APL, 153-168.
- Amorim, C. 2015. *Padrão de aquisição de contrastes do PE: a interacção entre traços, segmentos e sílabas*. Tese de Doutoramento. Faculdade de Letras da Universidade do Porto.
- Amorim, C. & Matzenauer, C. L. B. 2014. A aquisição das fricativas coronais no português europeu. *Cadernos de Estudos Linguísticos*. 56(2): 235-246.
- Archangeli, D. B. 1988. Aspects of Underspecification Theory. *Phonology Yearbook*. 5: 183-207.
- Baptista, A., Viana, F. L. & Barbeiro, L. F. 2011. *O ensino da escrita: Dimensões gráfica e ortográfica*. Lisboa: Ministério da Educação, Direcção-Geral de Inovação e de Desenvolvimento curricular.
- Carraher, T. N. 1985. Explorações sobre o desenvolvimento da competência em ortografia em português. *Psicologia: Teoria e Pesquisa*. 1(3): 269-285.
- Cassar, M. & Treiman, R. 1997. The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*. 89(4): 631-644.
- Castro, I. 2006. *Introdução à História do Português*. Lisboa: Colibri.
- Castro, S. L. & Gomes, I. 2000. *Dificuldades de aprendizagem da língua materna*. Lisboa: Universidade Aberta.
- Cintra, L. 1971. Nova Proposta de classificação dos dialectos galego-portugueses. *Boletim de Filologia*. 22: 81-116.
- Clements, G.N. & Hume, E. 1995. The internal organization of speech sounds. In: J. Goldsmith (Ed.). *The Handbook of Phonological Theory*, 7. Oxford/Cambridge: Basil Blackwell: 245-306.
- Correia, I. S. C. 2009. "Isso não Soa Bem". A Consciência Fonológica do lado de Lá - Reflexão em torno exercícios de Consciência Fonológica no Primeiro Ciclo. *Exedra: Revista Científica*. 1: 119-132.
- Costa, T. 2010. *The acquisition of the consonantal system in European Portuguese: focus on place and manner features*. Tese de Doutoramento. Faculdade de Letras da Universidade de Lisboa.

- Delgado-Martins, R. 1996. Representações da Linguagem Verbal. In: I. Faria, E. Pedro, I. Duarte & C. Gouveia (Orgs.). *Introdução à Linguística Geral Portuguesa*. Lisboa: Editorial Caminho, 85-102.
- Freitas, G. C. M. 2004. *Consciência fonológica e aquisição da escrita: um estudo longitudinal*. Tese de Doutoramento. Faculdade de Letras da Pontifícia Universidade Católica do Rio Grande do Sul.
- Freitas, M. J., Rodrigues, C., Costa, T. & Castelo, A. 2012. *Os sons que estão dentro das palavras. Descrição e Implicações para o Ensino do Português como Língua Materna*. Lisboa: Colibri.
- Freitas, M. J., Alves, D. & Costa, T. 2007. *O conhecimento da língua: Desenvolver a consciência fonológica*. Lisboa: Ministério da Educação, Direção-Geral de Inovação e Desenvolvimento Curricular.
- Freitas, M. J. 1997. *Aquisição da estrutura silábica do português europeu*. Tese de Doutoramento. Faculdade de Letras da Universidade de Lisboa.
- Hernandorena, C. M. 1990. *Aquisição da fonologia do Português: estabelecimento de padrões com base em traços distintivos*. Tese de Doutoramento. Universidade Católica do Rio Grande do Sul.
- Lao, Sh. 2020. *Variação da nasalização regressiva heterossilábica em Portugal Continental*. Dissertação de Mestrado em Linguística. Faculdade de Letras da Universidade de Lisboa.
- Lourenço-Gomes, M. C., Rodrigues, C. & Alves, I. 2016. EFFE-Escreves como falas - falas como escreves. *Revue Romane*. 51(1): 36-69.
- Mateus, M. H. M. & d'Andrade, E. 2000. *The Phonology of Portuguese*. Oxford: University Press.
- Matzenauer, C. & Costa, T. 2017. Aquisição da fonologia em língua materna: os segmentos. In: M. J. Freitas & A. L. Santos (Eds.). *Aquisição de língua materna e não materna: Questões gerais e dados do português*. Berlin: Language Science Press, 51-70.
- Meireles, E. & Correa, J. 2005. Regras contextuais e morfossintáticas na aquisição da ortografia da língua portuguesa por criança. *Psicologia: Teoria e Pesquisa*. 21(1): 77-84.
- Mendes, A., Afonso, E., Lousada, M. & Andrade, F. 2009/2013. *Teste fonético fonológico – Avaliação da linguagem pré-escolar*. Aveiro: Designed, Lda.
- Mendes, A., Afonso, E., Lousada, M. & Andrade, F. 2013. *Teste Fonético-Fonológico ALPE*. Aveiro: Edubox.
- Miranda, A.R.M. 2009. A grafia de estruturas silábicas complexas na escrita de crianças das séries iniciais. In: Sheila Z. de Pinho (Org.). *Formação de Educadores: o papel do educador e sua formação*. São Paulo: Unesp, 409-426.

- Miranda, A. R. M. & Matzenauer, C. 2010. Aquisição da fala e da escrita: relações com a fonologia. *Cadernos de Educação*. 35: 359-405.
- Miranda, A. R. M. 2010. Um estudo sobre o erro ortográfico. In: O. L. Heining & C. A. Fronza (Orgs.). *Diálogos entre linguística e educação*. 1. Blumenau: EDIFURB, 141-162.
- Oliveira, C. C. 2004. Sobre a aquisição das fricativas. In: R. Lamprecht (Eds.). *Aquisição fonológica do português: Perfil de desenvolvimento e subsídios para a terapia*. Porto Alegre: Artmed, 83-94.
- Pérez, X. A. Á. 2015. Isoglossas portuguesas nos materiais do *Atlas Lingüístico de la Península Ibérica*: análise crítica da *Nova Proposta* de Lindley Cintra. *Zeitschrift für romanische Philologie*. 131(1): 185-223.
- Pinto, M. G. 1998. *Saber Viver a Linguagem. Um Desafio aos Problemas de Literacia*. Porto: Porto Editora.
- Preston, J., Hull, M. & Edwards, M. L. 2013. Preschool speech error patterns predict articulation and phonological awareness outcomes in children with histories of speech sound disorders. *American Journal of Speech-Language Pathology*. 22: 173-183.
- Ramalho, A. M. 2017. *Aquisição fonológica na criança: tradução e adaptação de um instrumento de avaliação interlinguístico para o PE*. Tese de Doutoramento. Universidade de Évora.
- Ramalho, A. M. & Freitas, M. J. 2019. Ainda o ponto de articulação das sibilantes na alteração fonológica primária: dados de crianças portuguesas. *Fórum Linguístico*. 16(2): 3790-3808.
- Rio-Torto, G. 2000. Para uma pedagogia do erro. *Actas do V Congresso Internacional de Didáctica da Língua e da Literatura*. 1. Coimbra: Almedina, 595-618.
- Rodrigues, C. 2016. Variação sociolinguística. In: E. Carrilho & A. M. Brito (Eds.). *Manual de Linguística Portuguesa*. Berlim: Walter de Gruyter GmbH & Co KG, 16-98.
- Rodrigues, C., Lourenço-Gomes, M. C., Alves, I., Janssen, M., Lourenço-Gomes, I. 2015. EFFE-On – Escreves como falas – Falas como escreves? (Online corpus of writing and speech of children in the early years of schooling). Lisboa: CLUL.
- Rodrigues, C. & Lourenço-Gomes, M. C. 2016. Estudo longitudinal da proficiência ortográfica no 2.º e 4.º anos de escolaridade - estruturas /e/, /el/ e /oU/. *Revista Diacrítica*. 30(1): 115-36.
- Santos, R., Freitas, M. J. & Veloso, J. 2014. Grupos consonânticos na escola: desenvolvimento fonológico e conhecimento ortográfico. *Revista Diacrítica*. 28(1): 407-436.
- Scliar-Cabral, L. 2003. *Guia prático de alfabetização*. São Paulo: Contexto.

- Schier, A. C., Bert, L. C. & Chacon, L. 2013. Desempenho perceptual-auditivo e ortográfico de consoantes fricativas na aquisição da escrita. *CoDAS* 25: 45-51
- Segura, L. 2013. Variedades dialetais do português europeu. In: E. Paiva Raposo, M. F. Bacelar do Nascimento, M. A. Mota, L. Segura & A. Mendes (Eds.). *Gramática do português*. Lisboa: Fundação Calouste Gulbenkian, I, 85-142.
- Seymour, P. 1997. Foundations of orthographic development. In C. Perfetti, L. Rieben & M. Fayol (Eds.). *Learning to Spell*. Hillsdale, NJ: Erlbaum, 319-337.
- Silva, P. F. 2017. *Palatalização de laterais por harmonização de elementos*. Dissertação de Mestrado. Universidade do Porto.
- Teysier, P. 1982. *História da língua portuguesa*. Lisboa: Sá da Costa, 45-46.
- Treiman, R. 1998. Why spelling? The benefits of incorporating spelling into beginning reading instruction. In: J. L. Metsala & L. C. Ehri (Eds.). *Word recognition in beginning literacy*. Mahwah, NJ: Erlbaum, 289-313.
- Vale, A. P. S. & Sousa, O. 2017. Conhecimento ortográfico e escrita. *Da investigação às práticas. Estudos de Natureza Educacional*. 7(3): 3-7.
- Veloso, J. 2005. A língua na escrita e a escrita da língua. Algumas considerações gerais sobre transparência e opacidade fonémicas na escrita do português e outras questões. *Da Investigação às Práticas. Estudos de Natureza Educacional*. 6(1): 49-69.
- Veloso, J. 2010. Primeiras produções escritas e operações metafonológicas explícitas como pistas para a caracterização inferencial do conhecimento fonológico. *Cadernos de Educação*. 35: 19-50.
- Veloso, J. 2019. Complex Segments in Portuguese: The Unbearable Heaviness of Being Palatal. In: I. E. Zendoia & O. J. Nazabal (Eds.). *Bihotz ahots. M. L. Oñederra iraklearen omenez*. Bilbao: Universidad del País Vasco, Euskal Herriko Unibertsitatea, 513-526.

## Apêndice A

Tabela 5 - Exemplos de FNCs com o dígrafo <ch> no 2.º ano, dos GDT e GDA, da região de Chaves.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	E ela dantes não conseguia	jegar	as maçãs vermelhas	<j>
	O sol apares e a	clhuva.	O trigue faz mal	<clh>

GDA	a moca preta e o	sapeu	amarelo e a rabata amarela	<s>
	masã, mas não corsegia	segar	à masã. E viu um	<s>

Tabela 6 - Exemplos de FNCs com o dígrafo <ch> no 2.º ano, dos GDT e GDA, da região de Lisboa.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	ves uma broxa que	xamada	Lili que foi dar um	<x>
	os convite e ainda não	gegarão	todos e quen é que	<g>
GDA	lenbrai! uma	xavena	de xá sositivese	<x>
	A mamã quando	sego	a casa viu o papá	<s>

Tabela 7 - Exemplos de FNCs com o dígrafo <ch> no 2.º ano, dos GDT e GDA, da região do Porto.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	comvidaram o dara-gão mas todos	assaram	diferente por-que o dra-gaio elefante	<ss>
	vai a corer atras do	xapoeu	.	<x>
GDA	cabesa e fitu	japou	do voeu.   E ão	<j>
	ficou cotet e ela tirou	hapeu	veiu e ficou e vuou	<h>

Tabela 8 - Exemplos de FNCs com o grafema <x> no 2.º ano, dos GDT e GDA, da região de Chaves.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	estava a escon-der atrás da	bruchinha	estava a ser persegui-do por	<ch>
	o bolo mes a menina	deizou	sai quemre.   A menina	<z>
GDA	chegar la, e a	frucha	ficau contente.	<ch>
	lá. E olhou para	vacho	e viu uma tartaruga e	<ch>

Tabela 9 - Exemplos de FNCs com o grafema &lt;x&gt; no 2.º ano, dos GDT e GDA, da região de Lisboa.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	o gata oulo para a	brulha	.	<lh>
	la um passaro e a	bruza	estava a olar para ele	<z>
GDA	fazer chichi fes e depois	pochou	o otocolismo e fechou a	<ch>
	e um tigre e um	pais	e uma serpete e seis	<s>

Tabela 10 - Exemplos de FNCs com o grafema &lt;x&gt; no 2.º ano, dos GDT e GDA, da região do Porto.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	brocha vio um pasaro   A	bocla	istoeфонou o pasaro nom	<cl>
	O chapéu   “ O o A	borsa	extaba sentada no bamco e	<s>
GDA	O chapéu   Ea um	vuha	cafoi a pasa e	<h>
	vio um passaro e a	proja	rezou veu en teranformar o	<j>

### Apêndice B

Tabela 11 - Exemplos de FNCs com o dígrafo &lt;ch&gt; no 4.º ano, dos GDT e GDA, da região de Chaves.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	a se gastar, os	senelas	dela também era a	<s>
	comesou a abrir e a	fexar	a porta.   Os dois	<x>
GDA	uma rato a roubar qe-íjojo	ceio	de felicidade ou lando um	<c>
	a parser um gato. Upes	aclho	que este ficou muito feio	<clh>

Tabela 12 - Exemplos de FNCs com o dígrafo <ch> no 4.º ano, dos GDT e GDA, da região de Lisboa.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	por causa da	clave	.	<cl>
	a bola e um menino	jotou	a bola inves de	<j>
GDA	e a bruxa pensou eu	ajo	que tu pares- ses um macaco	<j>
	Sim   E ela ficou tão	satia	que disse um feitiço.	<s>

Tabela 13 - Exemplos de FNCs com o dígrafo <ch> no 4.º ano, dos GDT e GDA, da região do Porto.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	escola e quando	esgamos	a casa vio todo dezaro- mado	<>
	E um dado do	lhau	vaco e peto e um	<lh>
GDA	rapaz! E o meu	gegou	e ele tinha vito todo	<g>
	Nuno.   A diretora	jamou	o Miguel mas antes ele	<j>

Tabela 14 - Exemplos de FNCs com o grafema <x> no 4.º ano dos GDT e GDA da região de Chaves.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	lapis de cor.	Devàcho	da mesa tinha dois livlos	<ch>
	Era uma vez uma	brusca	que fazia muitas festinhas	<sc>
GDA	, laranja, castanho,	cocho	, preto, branco,	<ch>
	gigante da abóbora.   A	brucha	até ficou assim, será	<ch>

Tabela 15 - Exemplos de FNCs com o grafema &lt;x&gt; no 4.º ano dos GDT e GDA da região de Lisboa.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	sofá a comer, todo	relachado	.   Passado algum tempo,	<ch>
	salto da gangorra. A	bocla	disse.   - Meu deus	<cl>
GDA	o cabelo ruivo, a	mecher	numa esponcha e na	<ch>
	pior credo e o João	pojava	o tapete para a Inés	<j>

Tabela 16 - Exemplos de FNCs com o grafema &lt;x&gt; no 4.º ano dos GDT da região do Porto.

Grupo	Contexto precedente	Palavra-Alvo	Contexto seguinte	FNCs
GDT	mão gotu que	meilhe	mas milhas coisas	<lh>
	burxa começou a	micere	na tezoura	<c>

# Questionário sociolinguístico parental para famílias emigrantes bilingues (QuesFEB): uma ferramenta de recolha de dados sociolinguísticos de crianças falantes de herança<sup>1</sup>

Liliana Correia

[lilianamocorreia@gmail.com](mailto:lilianamocorreia@gmail.com)

*Centro de Estudos Humanísticos da Universidade do Minho (Portugal)*

Cristina Flores

[cflores@elach.uminho.pt](mailto:cflores@elach.uminho.pt)

*Centro de Estudos Humanísticos, Departamento de Estudos Germanísticos e Eslavos, Escola de Letras, Artes e Ciências Humanas, Universidade do Minho (Portugal)*

## ABSTRACT

Empirical research in the field of bi-/multilingualism has shown that the acquisition of two (or more) languages during childhood is significantly influenced by the sociolinguistic experience of each individual, namely by the quantity and the quality of language exposure to the target languages (Unsworth 2016a). In fact, the heterogeneity of sociolinguistic contexts in which bilingual acquisition takes place leads to variation in the quantity and quality of the input to which bilingual children are exposed on a daily basis, which, in turn, originates individual variation in the levels of language development in the languages under acquisition, mainly, in the minority language (also known as heritage language/HL; cf. Montrul 2016). In order to assess the effect of language experience on bilingual development, studies usually resort to sociolinguistic questionnaires, which allow the researcher to outline the sociolinguistic profile of the subjects under analysis, as well as to obtain crucial information about predictive variables of bilingual development (see Unsworth 2019). In this paper, we present a sociolinguistic questionnaire, in Portuguese, developed for the collection of data on the sociolinguistic experience of bilingual children, between the ages of six and ten, with a migration background – the *Questionário Sociolinguístico Parental para Famílias Emigrantes Bilingues* (QuesFEB). This parental questionnaire, intended for researchers who conduct studies in the field of heritage bilingualism, has as its main objective the collection of biographical and

---

<sup>1</sup> O questionário foi desenvolvido no âmbito do projeto de investigação “A influência de fatores extralinguísticos sobre o desenvolvimento lexical de crianças bilingues. Um estudo sobre falantes de herança do Português Europeu residentes em diferentes contextos migratórios”, o qual é financiado pela Fundação para a Ciência e a Tecnologia (FCT), pelo Fundo Social Europeu (FSE) e pelo Programa Operacional Capital Humano (POCH), através da bolsa de doutoramento com a referência SFRH/BD/138397/2018.

sociolinguistic information not only for the detailed characterisation of the context in which bilingual children acquire the heritage language, but also, and mainly, for the quantification of their language experience in the target language, enabling the assessment of the effect that variables related to input quantity and quality have on that language. Besides providing a detailed description of the content of the sections that compose the QuesFEB, we will present, in detail, the method of codification and calculation of four key variables that have been found to be predictive of HL development, namely: (i) current quantity of HL use (i.e., input and output) in the household; (ii) quantity of cumulative exposure to the HL in the household; (iii) quantity of HL use (i.e., input and output) with migrant grandparents who are native speakers of the language of origin; and (iv) richness of the language exposure to the HL.

#### KEY-WORDS

parental sociolinguistic questionnaire, child bilingualism, heritage language, child heritage speakers, sociolinguistic experience

#### RESUMO

A investigação empírica na área do bi-/multilinguismo tem revelado que a aquisição de duas (ou mais) línguas durante a infância é significativamente condicionada pela experiência sociolinguística de cada indivíduo, designadamente pela quantidade e pela qualidade da exposição linguística às línguas-alvo (Unsworth 2016a). De facto, a heterogeneidade de contextos sociolinguísticos em que a aquisição bilingue ocorre leva a variação na quantidade e na qualidade do *input* a que as crianças bilingues são expostas no dia-a-dia, o que origina, por sua vez, variação individual nos níveis de desenvolvimento linguístico das línguas em aquisição, sobretudo no da língua minoritária (também designada por língua de herança/LH; cf. Montrul 2016). Para avaliação do efeito da experiência linguística sobre o desenvolvimento bilingue, os estudos recorrem, geralmente, a questionários sociolinguísticos, os quais permitem ao investigador traçar o perfil sociolinguístico dos sujeitos em análise, bem como obter informações cruciais sobre variáveis preditivas do desenvolvimento bilingue (veja-se Unsworth 2019). No presente artigo, apresentamos um questionário sociolinguístico, em língua portuguesa, desenvolvido para a recolha de dados sobre a experiência sociolinguística de crianças bilingues, entre os seis e os dez anos de idade, com *background* migratório – o Questionário Sociolinguístico Parental para Famílias Emigrantes Bilingues (QuesFEB). Este questionário parental, destinado a investigadores que desenvolvem estudos na área do bilinguismo de herança, tem como principal objetivo recolher informações biográficas e sociolinguísticas que permitam não só caracterizar detalhadamente o contexto em que as crianças bilingues adquirem a língua de herança, mas também, e principalmente, quantificar a sua experiência linguística na língua-alvo, de modo a tornar possível a avaliação do efeito que variáveis relacionadas com a quantidade e a qualidade do *input* exercem sobre a mesma. Além da descrição detalhada do conteúdo das secções que compõem o QuesFEB, apresentar-se-á, pormenorizadamente, o método de codificação e de cálculo de quatro variáveis-chave que têm sido consideradas preditivas do desenvolvimento de uma LH, designadamente: (i) a quantidade de uso (i.e., *input* e *output*) da LH no contexto doméstico na atualidade; (ii) a quantidade de exposição acumulada à LH no agregado familiar; (iii) a quantidade de uso (i.e., *input* e *output*) da LH com avós emigrantes falantes nativos da língua de origem; e (iv) a riqueza da exposição linguística à LH.

## PALAVRAS-CHAVE

questionário sociolinguístico parental; bilinguismo infantil; língua de herança; crianças falantes de herança; experiência sociolinguística

## 1. Introdução

Nas últimas décadas, inúmeros estudos na área da aquisição de duas (ou mais) línguas durante a infância têm-se centrado no desenvolvimento linguístico das segundas e terceiras gerações de emigrantes de diferentes países de origem, isto é, no desenvolvimento linguístico de falantes de herança (FH) de diversas línguas-alvo (veja-se Gathercole 2014; Montrul 2016).

Os FH correspondem a um perfil específico de falantes bilingues precoces que crescem expostos à língua de origem da família emigrante – a língua de herança (LH) – e, simultânea ou sequencialmente, à língua do país de acolhimento, também designada por língua maioritária (Benmamoun, Montrul & Polinsky 2013). Sendo expostos à LH desde o nascimento, adquirindo-a, portanto, como língua materna (L1), os FH podem ser categorizados entre falantes bilingues *simultâneos* ou *sequenciais* segundo a idade de início de exposição à língua da sociedade maioritária (Rothman 2009). Assim, se os FH entrarem simultaneamente em contacto com a língua da família emigrante e a língua do país de acolhimento desde a nascença, ou até aos três/quatro anos (Meisel 2008), estamos a referir-nos a falantes bilingues simultâneos; se, por outro lado, estes começam a receber *input* da língua maioritária após os alicerces da L1 estarem estabelecidos, estamos perante falantes bilingues sequenciais.

Por ser sociopoliticamente minoritária, a LH é tipicamente adquirida sob condições de *input* linguístico reduzido, estando o seu uso geralmente limitado às interações comunicacionais no seio da família e da comunidade emigrante existente no país de acolhimento, a qual, por sua vez, pode ser demograficamente bastante significativa (e.g., a comunidade emigrante turca na Alemanha; a hispânica, nos EUA; a russa, em Israel; e a portuguesa, em França) ou, então, pouco expressiva (e.g., a comunidade alemã em Portugal; a comunidade portuguesa na Finlândia). Já a língua maioritária, usufruindo geralmente de estatuto oficial, é adquirida em contextos de

comunicação muito diversificados. É a língua principal da escolarização, da administração pública, dos *media* e das relações sociais fora do núcleo familiar (Benmamoun *et al.* 2013), podendo também ser uma das línguas faladas em contexto doméstico, não só em casos de bilinguismo simultâneo, como também em famílias que, por opção, ou pressionadas socialmente a assimilarem-se à cultura dominante, começam progressivamente a introduzi-la no quotidiano familiar (De Houwer 2017; Rothman 2009).

Para além das limitações impostas pelo estatuto sociopolítico que a LH detém na sociedade de acolhimento, a experiência sociolinguística dos falantes bilingues relativamente ao uso desta língua no seio familiar e na comunidade emigrante é ainda caracterizada por variação (Unsworth 2019). Há variação não só no que diz respeito à quantidade de interlocutores falantes nativos da LH com os quais as crianças bilingues contactam no dia-a-dia (i.e., um ou ambos os progenitores, membros da família alargada residentes no país de acolhimento, membros da comunidade emigrante, etc.), como também no que se refere às oportunidades que as crianças têm para usar a LH em atividades culturais e de lazer oferecidas pela comunidade (i.e., serviços religiosos, convívios com compatriotas, grupos folclóricos, etc.) (Correia, Flores & Melo-Pfeifer submetido). Concomitantemente, a experiência sociolinguística dos falantes bilingues varia ainda em relação à frequência do ensino formal da LH, isto é, há crianças bilingues que têm a oportunidade de frequentar aulas extracurriculares desta língua ou, inclusivamente, uma escola de ensino bilingue, ao passo que outras crianças não têm acesso a instrução formal na LH na sua área de residência.

Naturalmente, a multiplicidade de contextos sociolinguísticos em que a aquisição bilingue ocorre origina variação individual não só relativamente à quantidade e à qualidade do *input* linguístico a que os falantes bilingues são expostos, mas também (e conseqüentemente) no que diz respeito ao grau de proficiência linguística que os mesmos alcançam na LH, o qual se pode encontrar mais próximo ou mais afastado dos padrões de desenvolvimento monolíngue (Montrul 2016; Polinsky 2015).

Assim, ante a variação frequentemente observada na experiência sociolinguística e no nível de desenvolvimento linguístico dos falantes bilingues na LH, a investigação centrada na área do bi-/multilinguismo de herança tem investido na criação de instrumentos que permitam, por um

lado, avaliar o conhecimento linguístico destes falantes (veja-se Montrul 2016:180-188) e, por outro lado, caracterizar detalhadamente o seu perfil sociolinguístico e quantificar a sua experiência linguística (veja-se Kaščelan, Prevost, Serratrice, Tuller, Unsworth & De Cat 2020; Unsworth 2019).

Entre os instrumentos que têm sido desenvolvidos e utilizados para a recolha de informações sobre a experiência sociolinguística de crianças inseridas em contextos de aquisição bilingue, encontram-se os questionários sociolinguísticos parentais, os quais têm sido considerados ferramentas eficientes não só para a recolha de dados biográficos sobre os sujeitos bilingues, como também para a obtenção de dados quantificáveis sobre as suas práticas e experiências linguísticas (Codó 2008).

No presente artigo, apresentar-se-á um questionário sociolinguístico parental, o *Questionário Sociolinguístico Parental para Famílias Emigrantes Bilingues* (QuesFEB), desenvolvido para a avaliação da experiência sociolinguística de crianças falantes de herança entre os seis e os dez anos de idade. O QuesFEB distingue-se de outros questionários sociolinguísticos presentes na literatura por permitir recolher, num único instrumento de recolha de dados e em língua portuguesa, informações quantificáveis sobre diferentes componentes da experiência sociolinguística de falantes bilingues, em específico sobre: a quantidade de *input* e de *output* (na LH) no agregado familiar no presente; a quantidade de exposição à LH, no agregado familiar, desde o nascimento até ao presente; a qualidade da exposição linguística à LH (i.e., a riqueza do *input*); e a quantidade de uso (i.e., *input* e *output*) da LH entre as crianças bilingues e avós emigrantes falantes nativos da LH, uma componente sociolinguística raramente quantificada nos questionários parentais disponíveis na literatura. Para além disso, o QuesFEB destaca-se, ainda, por permitir recolher, de forma extensiva, informações sobre o cuidador da criança, designadamente sobre: as suas escolhas linguísticas no país de origem e em diversas esferas da sociedade maioritária; as suas vivências sociolinguísticas no país de acolhimento; as suas motivações, ou inibições, para a transmissão intergeracional da LH; e as estratégias implementadas nesse processo de transmissão.

O artigo encontra-se organizado da seguinte forma: na secção 2, abordar-se-ão, de forma breve, os principais fatores ambientais relacionados com a exposição linguística que têm sido considerados preditivos do

desenvolvimento linguístico de crianças falantes de herança e que têm estado na génese da criação de questionários sociolinguísticos direcionados para populações bilingues; na secção 3, focar-nos-emos num tipo específico de questionários sociolinguísticos, os questionários parentais, discutindo as suas valências e potencialidades na investigação sobre o desenvolvimento de LH; e, por fim, na secção 4, será feita uma caracterização detalhada do QuesFEB. Em específico, apresentaremos: (i) a sua origem e objetivo; (ii) a população-alvo; (iii) o modo de administração e duração; (iv) a constituição e o conteúdo; (v) e o método de codificação e de cálculo das variáveis de *input* preditivas do desenvolvimento linguístico.

## 2. Fatores ambientais preditivos do desenvolvimento de línguas de herança

A investigação empírica realizada com falantes de herança de diversas línguas tem demonstrado que, embora estes possam vir a alcançar, em idade adulta, níveis de proficiência linguística idênticos aos dos seus pares monolíngues do país de origem, o ritmo de desenvolvimento linguístico da LH é fortemente condicionado pela quantidade e pela qualidade da exposição linguística à mesma.

Os efeitos da quantidade de *input* sobre a aquisição da LH têm sido constatados em vários domínios do saber linguístico, tais como o do léxico e o da morfossintaxe. No léxico, por exemplo, tem sido observado que os falantes de herança tendem a apresentar vocabulários produtivos e recetivos quantitativamente inferiores aos dos seus pares monolíngues (Bialystok, Luk, Peets & Yang 2010; Cobo-Lewis, Pearson, Eilers & Umbel 2002; Correia & Flores 2017; Cote & Bornstein 2014, entre outros). Já no que diz respeito ao domínio da morfossintaxe, tem-se verificado que o processo de aquisição de algumas propriedades gramaticais das línguas-alvo tende a ser mais moroso (Flores & Barbosa 2014; Flores, Santos, Jesus & Marques 2017; Gathercole 2002; Rodina & Westergaard 2017, entre outros), podendo estas, inclusivamente, vir a não estar completamente estabilizadas em idade adulta.

Dos vários fatores que contribuem para a variação na quantidade de *input* que as crianças recebem no quotidiano e, conseqüentemente, para

o seu enquadramento ou afastamento dos padrões de desenvolvimento linguístico dos falantes monolíngues da língua-alvo, destacam-se a constituição do agregado familiar e a presença de avós falantes nativos da LH no país de acolhimento. Quanto ao agregado familiar, tem sido documentado na literatura que, embora haja variação nos padrões de uso linguístico no ambiente doméstico, o desenvolvimento linguístico da LH é mais consistente quando os dois progenitores são falantes nativos da língua minoritária (Correia & Flores 2017; Rodina & Westergaard 2017). Também a existência/ausência de irmãos mais velhos no agregado familiar é apontada como variável preditiva do desenvolvimento na LH (Bridges & Hoff 2014). Assim, se, por um lado, o uso da LH tende a ser mais intensivo em lares em que os dois pais têm a LH como língua materna, por outro, a presença de irmãos mais velhos nos agregados familiares tende a promover o uso da língua maioritária, visto que os irmãos mais velhos costumam recorrer a esta língua com mais frequência nas interações diárias e são, em regra, os seus introdutores em famílias cuja única língua de comunicação costumava ser a de origem. Por sua vez, a presença de avós falantes nativos da LH no país de acolhimento tem sido apontada como um fator positivo para a transmissão da mesma (Ishizawa 2004; Zhang 2008). O facto de os avós emigrantes frequentemente apresentarem níveis de proficiência baixos na língua do país de acolhimento leva a que as interações intergeracionais ocorram sobretudo com recurso à língua de origem, aumentando, assim, as oportunidades de as crianças serem expostas a esta língua. Para além disso, em algumas famílias emigrantes, os avós fazem parte do agregado familiar ou, então, colaboram na educação dos netos, o que contribui significativamente para um uso mais regular da língua minoritária.

Outra variável importante da experiência linguística dos falantes bilingues que, estando também correlacionada com a quantidade de *input* (veja-se Pearson 2007), origina variação no desenvolvimento linguístico da LH é a própria produção linguística nesta língua por parte das crianças bilingues, isto é, o seu *output* na LH. Alguns estudos têm mostrado que a quantidade de *output* na LH está positiva e significativamente associada ao desempenho destes falantes em diversos testes de avaliação linguística (Bohman, Bedore, Peña, Mendez-Perez & Gillam 2010; Correia & Flores 2017; Unsworth 2012), isto é, os falantes de herança que fazem um uso

mais ativo da LH no quotidiano apresentam um desenvolvimento linguístico mais avançado do que aqueles que usam esta língua de forma menos ativa.

E, por fim, quanto à qualidade da exposição linguística à LH, a variação na diversidade e na complexidade do *input* a que os falantes de herança têm acesso no quotidiano também determina o nível de proficiência linguística na língua-alvo. De facto, vários estudos têm revelado que o desenvolvimento linguístico destes falantes na LH é significativamente influenciado não só pelo contacto com diferentes interlocutores falantes da LH (Gollan, Starr & Ferreira 2015; Place & Hoff 2016) e pela frequência do ensino formal desta língua (Bylund & Diaz 2012; Cobo-Lewis *et al.* 2002), como também pelo contacto com a LH através de atividades de lazer (Makarova, Terekhova & Mousavi 2019) e socioculturais (Gathercole & Thomas 2007).

### 3. Questionários parentais

Sendo o perfil dos falantes de herança determinado pelo contexto sociolinguístico em que a aquisição da língua de herança ocorre, os questionários sociolinguísticos tornam-se instrumentos indispensáveis para a realização de investigação na área do bi-/multilinguismo (Montrul 2016).

No caso específico de estudos centrados no desenvolvimento linguístico de crianças falantes de herança, os questionários sociolinguísticos tendem a ser aplicados, de forma direta ou indireta, ao cuidador da criança, sendo-lhe solicitado que forneça informações biográficas sobre o agregado familiar, bem como sobre a experiência sociolinguística não só da criança, mas também, de acordo com os objetivos de cada estudo, de outros indivíduos que contactem regularmente com a mesma.

Do leque de potenciais informações biográficas e sociolinguísticas que podem ser requisitadas através deste tipo de instrumentos de recolha de dados, destacam-se as que se encontram listadas na Tabela 1.

TABELA 1 – Informações biográficas e sociolinguísticas passíveis de serem recolhidas em questionários sociolinguísticos (adaptado de Codó 2008).

<i>Informação biográfica</i>	<ul style="list-style-type: none"><li>▪ idade (cronológica, de início de exposição às línguas-alvo, de emigração)</li><li>▪ local de nascimento</li><li>▪ género</li><li>▪ escolarização (número de anos, tipo de escola, língua(s) de instrução)</li><li>▪ profissão</li><li>▪ problemas de desenvolvimento (audição e/ou fala)</li></ul>
<i>Uso linguístico</i>	<ul style="list-style-type: none"><li>▪ línguas faladas com diferentes interlocutores (<i>input</i> e <i>output</i>)</li><li>▪ línguas faladas nas diversas esferas da sociedade</li><li>▪ línguas faladas em diferentes fases da vida dos participantes</li><li>▪ línguas utilizadas em diferentes atividades de lazer e socioculturais</li><li>▪ razões para as escolhas linguísticas</li></ul>
<i>Proficiência linguística</i>	<ul style="list-style-type: none"><li>▪ auto-/heteroavaliação da competência linguística na(s) língua(s)-alvo (compreensão oral/escrita; produção oral/escrita)</li></ul>
<i>Atitudes e preferências linguísticas</i>	<ul style="list-style-type: none"><li>▪ atitudes face aos falantes de uma determinada língua</li><li>▪ atitudes face ao uso das línguas-alvo nos vários domínios sociais</li><li>▪ grau de identificação com as línguas-alvo e culturas-alvo</li><li>▪ línguas preferidas</li></ul>
<i>Transmissão/Aquisição linguística</i>	<ul style="list-style-type: none"><li>▪ motivações para a transmissão/aquisição das línguas-alvo</li><li>▪ fatores que fomentam/inibem a transmissão/aquisição das línguas-alvo</li></ul>

Em relação às questões para a recolha destes dados, estas podem ser apresentadas em formato de questão fechada, aberta ou semiaberta/semifechada (Christensen, Johnson & Turner 2015). As questões de resposta fechada não permitem que o participante forneça informações para além do conjunto de opções de respostas pré-estabelecidas e variam geralmente entre questões de resposta única, de escolha múltipla e de escala. Já as questões de resposta aberta dão a oportunidade ao participante de responder de forma livre, não impondo limitações ao conteúdo das respostas. Por sua vez, as questões de resposta semiaberta/semifechada caracterizam-se por combinarem propriedades dos dois tipos de questões anteriores.

Em regra, a generalidade dos questionários parentais que são usados na área de investigação do bilinguismo recorre maioritariamente a questões fechadas, uma vez que, devido ao facto de as possibilidades de respostas serem padronizadas (i.e., limitadas a categorias predeterminadas), estas facilitam a codificação numérica dos dados recolhidos, sendo, portanto, adequadas para o tratamento quantitativo e estatístico dos mesmos (Dörnyei & Taguchi 2010).

Desta forma, através do recurso a este tipo de questões, o investigador consegue recolher dados quantificáveis que podem ser utilizados, por um lado, para a análise descritiva da população-alvo e do contexto sociolinguístico em que está inserida e, por outro lado, para as análises correlacionais entre as variáveis relacionadas com a experiência linguística dos participantes e os seus resultados em testes de avaliação de competências linguísticas. Salienta-se, contudo, que, embora não seja muito frequente, o recurso a questões abertas permite uma exploração mais aprofundada de determinadas temáticas (e.g., motivações para um determinado comportamento, razões para preferências linguísticas, etc.), podendo inclusivamente levar o investigador a identificar temas que não tinha antecipado (Codó 2008; Dörnyei & Taguchi 2010).

No que diz respeito à avaliação e quantificação da experiência linguística dos falantes bilingues, os diversos questionários parentais disponíveis variam não só quanto às variáveis-alvo que procuram aferir, como também relativamente ao método adotado para o cálculo das mesmas (veja-se Kaščelan *et al.* 2020).

De seguida, apresentar-se-ão alguns dos métodos adotados para a recolha e quantificação de componentes-chave da experiência sociolinguística de falantes bilingues, tendo por base dois questionários parentais, o *Alberta Language Environment Questionnaire* (ALEQ) de Paradis (2011) e o *Bilingual Language Experience Calculator* (BiLEC) de Unsworth (2013, 2016b), que têm sido amplamente usados em estudos centrados no bi-/multilinguismo (Chondrogianni & Schwartz 2020; Garraffa, Beveridge & Sorace 2015; Rezzonico, Goldberg, Milburn, Belletti & Girolametto 2017; Rodina & Westergaard 2017; Sorenson Duncan & Paradis 2020; entre outros) e que deram origem ao questionário sociolinguístico que será apresentado na secção 4.

### 3.1 Recolha e quantificação da experiência sociolinguística de falantes de herança

A componente da experiência linguística mais visada nos diferentes questionários parentais é, naturalmente, a exposição linguística da criança à(s) língua(s)-alvo no presente. No entanto, apesar de esta componente ser comum à generalidade destes questionários, o modo como a sua avaliação é operacionalizada varia de instrumento para instrumento. Por exemplo, no *Alberta Language Environment Questionnaire* de Paradis (2011), um questionário parental direcionado para cuidadores de crianças bilingues sequenciais em idade escolar, a autora conjuga a quantidade de *input* da língua-alvo que as crianças recebem dos membros do agregado familiar com a quantidade de *output* que as mesmas produzem para estes membros para criar uma medida composta de uso linguístico no presente. Para a quantificação da exposição e da produção linguística no agregado familiar, Paradis (2011) recorre às estimativas parentais fornecidas em escalas de 5-pontos (e.g., *Mother tongue always/English never* = 0; *Mother tongue usually/English seldom* = 1; *Mother tongue 50%/English 50%* = 2; *Mother tongue seldom/English usually* = 3; *Mother tongue almost never/English almost always* = 4). Num outro questionário, o *Bilingual Language Experience Calculator* de Unsworth (2013, 2016b), destinado a cuidadores de falantes bilingues simultâneos ou sequenciais entre os dois e os dezoito anos de idade, a autora avalia a quantidade de *input* e de *output* na língua-alvo em separado. Para a quantidade de exposição linguística à língua-alvo, este questionário, que se faz acompanhar de um ficheiro *Excel* de cálculo automático, disponibiliza quatro medidas: (i) quantidade de *input* apenas no agregado familiar; (ii) quantidade de *input* no agregado familiar e em instituições (pré-)escolares; (iii) quantidade de *input* no agregado familiar, em instituições (pré-)escolares e de outras fontes, tais como atividades extracurriculares, *media* e amigos; e (iv) quantidade de *input* proveniente de todas as fontes anteriores, bem como a quantidade de exposição linguística durante o período de férias. Já a quantidade de *output* na língua-alvo é avaliada apenas no contexto doméstico. Salienta-se ainda que, embora quer o ALEQ quer o BiLEC utilizem os valores obtidos em escalas de 5-pontos para quantificar a exposição e a produção linguística dos participantes, o

BiLEC inclui ainda, nas fórmulas de cálculo, o número de horas diárias em que o participante está acordado, o número de horas que cada interlocutor passa com o participante, bem como o número de horas de exposição a fontes de *input* externas ao ambiente doméstico e escolar.

Devido à variação na quantidade de *input* que pode ocorrer ao longo da vida das crianças bilingues, alguns questionários parentais procuram também avaliar a quantidade de exposição acumulada à(s) língua(s)-alvo desde o nascimento até à idade atual das crianças (e.g., Cohen, Bauer & Minniear 2021, Gutiérrez-Clellen & Kreiter 2003; Unsworth 2013, 2016b). Para o efeito, questionários como o BiLEC solicitam, para cada ano de vida da criança, informações sobre a quantidade de *input* da língua-alvo que a criança recebeu de cada membro do agregado familiar, bem como sobre a sua frequência de creches ou instituições (pré-)escolares e a(s) língua(s) usada(s) nestes contextos. Com base nestes dados, torna-se possível, então, determinar a proporção de uso de uma dada língua em cada ano etário e calcular a exposição acumulada a essa língua ao longo da vida da criança.

Além da quantidade de exposição no presente e no passado, outra componente da experiência linguística dos falantes bilingues que tem vindo a ser avaliada nos questionários parentais é a qualidade da exposição linguística (Paradis 2011; Tuller 2015; Unsworth 2016b). No ALEQ, a qualidade da exposição linguística é quantificada através de uma medida composta – *richness score* – que avalia não só a frequência semanal com que a criança contacta com a língua-alvo em atividades relacionados com os *media* (e.g., ler livros/revistas, usar o computador, ver televisão, etc.) e em atividade extracurriculares, mas também a quantidade de uso desta língua entre a criança e os amigos com quem brinca regularmente. Cumulativamente, no caso de a língua-alvo corresponder à língua minoritária, esta variável composta avalia ainda a frequência do ensino formal nesta língua. Já o BiLEC, por sua vez, apresenta várias medidas para a avaliação da qualidade do *input* (para uma descrição detalhada das medidas geradas e das fórmulas de cálculo, veja-se Unsworth 2016b), destacando-se a que conjuga a quantidade de exposição linguística à(s) língua(s)-alvo com a proficiência linguística dos seus provedores, sendo esta avaliada, pelo cuidador da criança, em escalas de 6-pontos que variam entre *nenhuma fluência* (0) e *fluência nativa* (5).

#### 4. Questionário sociolinguístico parental para famílias emigrantes bilingues

Na presente secção, será apresentado, de forma detalhada, o QuesFEB - Questionário Sociolinguístico Parental para Famílias Emigrantes Bilingues, sendo descrita a sua origem e objetivo, a população-alvo a que se destina, os procedimentos de administração, assim como a sua estrutura e metodologia de quantificação.<sup>2</sup>

##### 4.1 Origem e objetivo

O QuesFEB foi desenvolvido no âmbito de um projeto de investigação centrado na influência de fatores extralinguísticos sobre o desenvolvimento lexical de crianças falantes de herança do Português Europeu (Correia, em preparação). Trata-se de um questionário parental detalhado, elaborado a partir do trabalho de Paradis (2011), Schmid & Dusseldorp (2010) e Unsworth (2013)<sup>3</sup>, que tem como objetivo primordial a recolha de dados biográficos e sociolinguísticos que permitam, por um lado, realizar uma caracterização pormenorizada do contexto em que as crianças bilingues adquirem a língua de herança e, por outro lado, quantificar a experiência linguística destes falantes de modo a ser possível correlacioná-la, através da constituição de diversas variáveis-chave, com o seu desenvolvimento linguístico na língua de origem.

##### 4.2 População-alvo

O QuesFEB destina-se a investigadores que pretendam conduzir estudos centrados na aquisição de línguas de herança por crianças bilingues entre os seis e os dez anos de idade. Salienta-se ainda que, embora tenha sido concebido para a recolha de dados sociolinguísticos de crianças falantes

---

<sup>2</sup> O questionário está disponível, em acesso aberto, no Repositório de Dados da Universidade do Minho (<https://doi.org/10.34622/datarepositorium/1|SLNQ>).

<sup>3</sup> As questões e fórmulas de cálculo para as variáveis INPUT.LH.AF, OUTPUT.LH.AF e RIQUEZA.INPUT (subsecções 4.5.1 e 4.5.4) foram adaptadas de Paradis (2011); já as questões relacionadas com a variável QEA (subsecção 4.5.2) foram adaptadas de Unsworth (2013), tendo-se simplificado a fórmula de cálculo da mesma; por fim, do questionário sociolinguístico de Schmid & Dusseldorp (2010), adaptaram-se as questões relacionadas com os padrões de uso linguístico do cuidador da criança no país de acolhimento, a valorização da transmissão da LH e a identificação cultural com a cultura de origem e a da sociedade de acolhimento (primeira secção do questionário parental).

de herança do Português Europeu residentes na Alemanha e em França, o QuesFEB pode ser adaptado para famílias emigrantes a viver em outros países de acolhimento e com outras línguas de origem.

#### 4.3 Modo de administração e duração

O QuesFEB é um questionário de autoadministração, i.e., para ser preenchido por um dos pais da criança, e demora cerca de 20 minutos a ser completado.

Por não ser preenchido na presença do investigador, este deve manter o contacto dos participantes para que seja possível completar informação em falta, no caso de o questionário ser devolvido com questões não respondidas.

É ainda importante destacar que o facto de o QuesFEB ter sido desenhado para ser preenchido de forma autónoma pelos cuidadores das crianças bilingues (contrariamente, por exemplo, ao ALEQ e ao BiLEC) pode torná-lo vantajoso não só para o cuidador da criança, que o pode preencher de acordo com a sua disponibilidade, como também para o próprio investigador, que, não tendo de agendar entrevistas com o cuidador, não necessita de despende de muito tempo para a sua aplicação.

#### 4.4 Constituição e conteúdo

O QuesFEB é constituído por três secções: as duas primeiras partes direccionadas à recolha de informações biográficas e sociolinguísticas sobre os pais da criança e uma última parte centrada na recolha de dados sobre a criança.

Em relação à primeira secção, focada no cuidador que está a preencher o questionário, solicitam-se não só dados biográficos sobre o mesmo (e.g., idade, país de nascimento, duração da estadia no país de acolhimento, país e região de origem dos progenitores, nível de escolaridade, entre outros), como também informações sociolinguísticas relacionadas com os padrões de uso da língua de herança e da língua maioritária no país de acolhimento e no país de origem, a importância atribuída à transmissão intergeracional da língua de origem e a existência de atitudes sociais desfavoráveis a este processo. Para além disso, o cuidador é ainda questionado sobre as

estratégias implementadas na transmissão da língua de origem, bem como sobre as razões que motivam/inibem este processo.

No que se refere à segunda secção, centrada no progenitor da criança que não está a preencher o questionário, solicitam-se apenas informações biográficas idênticas às requeridas na primeira secção. Uma vez que este progenitor poderá não ter um *background* migratório, ou, inclusivamente, não estar inserido no agregado familiar da criança, esta parte do questionário não recolhe dados sociolinguísticos relacionados com os padrões de uso da língua de herança, nem com a sua transmissão.

Por fim, no que diz respeito à última secção do questionário, além de se requisitar dados biográficos sobre a criança e eventuais irmãos da mesma (e.g., data e país de nascimento, idade de emigração), bem como sobre a existência de problemas de desenvolvimento relacionados com a audição e/ou a fala, solicita-se que o cuidador forneça informações sobre a experiência sociolinguística da criança, nomeadamente sobre: a idade e o contexto em que esta teve contacto pela primeira vez com cada uma das línguas em aquisição; a quantidade de *input*, na LH, que a criança recebe de cada membro do agregado familiar e a quantidade de *output*, nesta língua, que a criança produz para os mesmos na atualidade; a quantidade de *input* da LH que a criança recebeu de cada membro do agregado familiar no passado; os padrões de uso da LH e da língua maioritária com falantes nativos da língua de origem residentes no país de acolhimento (e.g., avós e amigos); a frequência de uso da LH através de atividades relacionadas com os *media* (e.g., ver televisão, ler, etc.); a participação em atividades culturais e de lazer, na LH, no seio da comunidade emigrante; o acesso ao ensino formal da LH; a periodicidade das visitas ao país de origem e a frequência com que a criança usa a LH neste contexto. Adicionalmente, solicita-se ainda que o cuidador avalie a proficiência linguística da criança quanto à produção e à compreensão oral na LH.

Deste modo, com os dados fornecidos pelo cuidador, torna-se possível fazer a caracterização detalhada do contexto sociolinguístico em que a aquisição da LH ocorre, bem como obter os dados que permitem quantificar alguns dos principais fatores de *input* que têm sido considerados preditivos do desenvolvimento bilingue. Para além disso, a recolha dos dados centrados no cuidador da criança que preenche o questionário torna ainda

possível realizar uma exploração das suas escolhas linguísticas e vivências sociolinguísticas que possam ter um impacto, negativo ou positivo, no processo de transmissão/manutenção da língua de origem.

#### 4.5 Método de codificação e cálculo das variáveis de *input* preditivas do desenvolvimento linguístico

##### 4.5.1 Quantidade de uso da língua de herança no contexto doméstico na atualidade

O uso da LH no contexto doméstico é avaliado através de duas variáveis: (i) quantidade de *input* da LH no agregado familiar (INPUT.LH.AF) e (ii) quantidade de *output* na LH no agregado familiar (OUTPUT.LH.AF).

Para o cálculo dos valores destas variáveis-alvo, o investigador deve recorrer aos dados das questões n.º 12 e n.º 13 da terceira secção do questionário, nas quais se solicita ao cuidador que avalie, numa escala de 5-pontos, não só a frequência com que cada um dos membros do agregado familiar usa a LH para comunicar com a criança (questão n.º 12), como também a frequência com que a criança utiliza esta língua para interagir com cada um desses membros (questão n.º 13).

De seguida, o investigador deve atribuir, às respostas assinaladas na escala de 5-pontos, os valores das medidas quantitativas que se encontram na Tabela 2.

TABELA 2 – Conversão das categorias da escala de 5-pontos das questões n.º 12, n.º 13, n.º 28, n.º 29 e n.º 30 em medidas quantitativas.

<i>Categorias da escala</i>	<i>Medidas quantitativas</i>
"Nunca (0%)"	0.00
"Raramente (25%)"	0.25
"Às vezes (50%)"	0.50
"Frequentemente (75%)"	0.75
"Sempre (100%)"	1.00

Após a atribuição dos valores quantitativos às respostas fornecidas, a quantidade de *input* da LH no agregado familiar é calculada através da Fórmula 1, em que *a* é a soma dos valores de *input* de cada membro do agregado familiar com idade superior a dois anos (segundo Paradis 2011) e *b* é o número de membros do agregado familiar, sem se contabilizar a criança em estudo, com idades também superiores a dois anos.

$$(1) \text{ INPUT.LH.AF} = \frac{a}{b}$$

A quantidade de *output* na LH no agregado familiar é, por seu turno, determinada através da Fórmula 2, em que *a* representa a soma dos valores de *output* para cada membro do agregado familiar com idade superior a dois anos e *b* corresponde ao número de membros do agregado familiar, excluindo-se a criança em análise, com idade superior a dois anos.

$$(2) \text{ OUTPUT.LH.AF} = \frac{a}{b}$$

Para as duas variáveis-alvo, os valores calculados podem variar entre o mínimo de 0 (i.e., 0% de *input/output* da LH) e o máximo de 1 (i.e., 100% de *input/output* da LH).

#### 4.5.2 Quantidade de exposição acumulada à língua de herança no agregado familiar

O cálculo da quantidade de exposição acumulada (QEA) à LH, isto é, da quantidade de *input* da LH, medida em meses, a que a criança foi exposta no agregado familiar desde o nascimento até à sua idade atual, é realizado em três etapas.

Na primeira etapa, o investigador deve recorrer aos dados fornecidos nas questões n.º 28, n.º 29 e n.º 30 da terceira secção do questionário, nas quais se solicita ao cuidador que avalie, numa escala de 5-pontos semelhante à da Tabela 2, a frequência com que cada membro do agregado familiar usava a LH para comunicar com a criança em três períodos temporais anteriores ao atual, designadamente: (a) desde o nascimento da criança até esta fazer três anos (questão n.º 28); (b) desde os três anos até a criança fazer seis

anos (questão n.º 29); e (c) desde os seis anos até à idade anterior à atual (i.e., anterior ao presente ano etário), caso a criança tenha idade superior a seis anos (questão n.º 30). Com os dados fornecidos nestas três questões, o investigador deve, num primeiro momento, determinar, para cada ano de vida da criança, a quantidade de *input* da LH a que a criança foi exposta no agregado familiar, através de uma fórmula idêntica à utilizada com a variável INPUT.LH.AF (i.e., Fórmula 1). De seguida, partindo do pressuposto de que a quantidade máxima de *input* da LH num ano (i.e., 1.00) equivale a 12 meses de exposição, o investigador deve converter a quantidade de *input* recebido em cada ano em meses de exposição, através da Fórmula 3, em que *a* representa a quantidade de *input* apurada para cada ano de vida, *b* corresponde à quantidade máxima possível de meses de exposição num ano (i.e., 12 meses), e *c* é a quantidade máxima de *input* possível (i.e., 1.00) de acordo com os valores apresentados na Tabela 2.

$$(3) \text{ meses de exposição por ano (no passado)} = \frac{a \times b}{c}$$

Por fim, e ainda dentro desta primeira etapa, o investigador deve proceder ao somatório dos meses de exposição linguística apurados para cada ano de vida da criança anterior à sua idade atual.

Na segunda etapa, o investigador deve recorrer aos valores da variável INPUT.LH.AF, relativa à quantidade de *input* da LH na atualidade, para determinar a quantidade de meses de exposição linguística à LH no presente ano etário e, posteriormente, adicioná-los aos meses de exposição já apurados. Para tal, e partindo da premissa de que 100% do *input* da LH na atualidade equivale a um número de meses de exposição igual à quantidade de meses de vida da criança no ano etário em que se encontra, o investigador deve proceder ao cálculo dos meses de exposição em falta através da Fórmula 4, em que *a* é o valor apurado na variável INPUT.LH.AF, *b* corresponde à quantidade máxima possível de meses de exposição no presente ano etário (i.e., a quantidade de meses de vida da criança no ano etário em que se encontra), e *c* representa a quantidade máxima de *input* possível (i.e., 1.00), segundo os valores expostos na Tabela 2.

$$(4) \text{ meses de exposição (no presente ano etário)} = \frac{a \times b}{c}$$

Na última etapa, o investigador deve somar os meses de exposição linguística do passado aos do presente. O resultado desta operação constitui o valor da variável QEA, isto é, a quantidade de *input* da LH, em meses, a que a criança foi exposta ao longo da sua vida por meio dos membros do agregado familiar. Para uma melhor clarificação do cálculo da variável QEA, apresenta-se, na Tabela 3, um exemplo do valor obtido para uma criança com a idade cronológica de 8;04.

TABELA 3 – Resultado do cálculo da variável QEA para uma criança com a idade cronológica de 8;04.

	Período I			Período II			Período III		Presente	
Anos	1	2	3	4	5	6	7	8		
Me- ses	0-12	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-100	
(1)	1.00	1.00	1.00	0.85	0.85	0.85	0.85	0.85	0.80	<b>QEA</b>
(2)	12.00	12.00	12.00	10.20	10.20	10.20	10.20	10.20	3.20	90.2

Nota. (1) = quantidade de input da LH; (2) = quantidade de exposição linguística à LH em meses.

#### 4.5.3 Quantidade de uso da língua de herança com avós emigrantes falantes nativos da língua de origem

O uso da LH entre a criança e avós emigrantes falantes nativos desta língua é avaliado através de duas variáveis: (i) quantidade de *input* da LH que a criança recebe destes familiares no dia-a-dia (INPUT.LH.AVÓS) e (ii) quantidade de *output* na LH que a criança produz para estes interlocutores (OUTPUT.LH.AVÓS).

Para calcular os valores para estas duas variáveis, o investigador deve recorrer aos dados das questões n.º 14, n.º 15, n.º 16 e n.º 17 da terceira secção do questionário, nas quais se solicita ao cuidador não só que indique os avós falantes nativos da LH que residem no país de acolhimento (fora do agregado familiar) e a frequência semanal com que os mesmos contactam

com a criança (questões n.º 14 e n.º 15, respetivamente), como também que especifique, numa escala de 5-pontos, a quantidade relativa de *input* da LH e da língua maioritária que a criança recebe de cada um desses avós (questão n.º 16) e a quantidade relativa de *output*, nestas línguas, que a criança produz para os mesmos (questão n.º 17). Os valores que serão atribuídos às variáveis INPUT.LH.AVÓS e OUTPUT.LH.AVÓS resultam da conjugação dos dados fornecidos nestas quatro questões e o seu cálculo realiza-se em três etapas.

Na primeira etapa, o investigador deve pontuar as respostas assinaladas na escala de 5-pontos de acordo com as medidas quantitativas que se encontram na Tabela 4.

TABELA 4 – Conversão das categorias da escala de 5-pontos das questões n.º 16 e n.º 17 em medidas quantitativas.

<i>Categorias da escala</i>	<i>Medidas quantitativas</i>
“Apenas X”	0.00
“Mais X do que Y”	0.25
“Y e X em quantidades iguais”	0.50
“Mais Y do que X”	0.75
“Apenas Y”	1.00

Nota. X = língua maioritária; Y = língua de herança.

De seguida, o investigador deve proceder ao cálculo da quantidade de *input* que a criança recebe de cada avô (INPUT.AVÔ.IND) e da quantidade de *output* que a criança produz também para cada um (OUTPUT.AVÔ.IND), semanalmente. O cálculo da variável INPUT.AVÔ.IND é efetuado através da Fórmula 5, em que *a* corresponde à quantidade de *input* avaliada segundo os parâmetros da Tabela 4, *b* representa o número de dias por semana em que a criança convive com o avô e *c* é o número de dias possíveis numa semana (i.e., sete dias).

$$(5) \text{ INPUT.AVÔ.IND} = a \times \left( \frac{b}{c} \right)$$

Por sua vez, o cálculo da variável OUTPUT.AVÔ.IND realiza-se através da Fórmula 6, em que  $a$  corresponde, desta vez, à quantidade de *output* avaliada segundo os critérios da Tabela 4,  $b$  representa o número de dias por semana em que a criança convive com o avô e  $c$  é o número de dias possíveis numa semana (i.e., sete dias).

$$(6) \text{ OUTPUT.AVÔ.IND} = a \times \left( \frac{b}{c} \right)$$

Na última etapa, o investigador determina os valores para as variáveis INPUT.LH.AVÓS e OUTPUT.LH.AVÓS através das Fórmulas 7 e 8, respetivamente. Na Fórmula 7,  $a$  corresponde à soma dos valores da quantidade de *input* proveniente de cada avô e  $b$  representa o número de avós falantes nativos da LH residentes no país de acolhimento. Na Fórmula 8, por seu turno,  $a$  equivale à soma dos valores da quantidade de *output* para cada avô e  $b$  é o número de avós falantes nativos da LH residentes no país de acolhimento.

$$(7) \text{ INPUT.LH.AVÓS} = \frac{a}{b}$$

$$(8) \text{ OUTPUT.LH.AVÓS} = \frac{a}{b}$$

Para as duas variáveis-alvo, os valores podem variar entre o mínimo de 0 e o máximo de 1, sendo que, quanto mais próximo de 1 está o valor, maior é o contacto semanal e a interação linguística na LH entre os avós e a criança.

#### 4.5.4 Riqueza da exposição linguística à língua de herança

A riqueza da exposição linguística da criança à língua de herança (RIQUEZA.INPUT), i.e., a qualidade da exposição linguística, é avaliada através dos dados recolhidos nas questões n.º 18, n.º 19, n.º 21, n.º 22, n.º 24 e n.º 25 da terceira secção do questionário, as quais solicitam informações sobre a experiência sociolinguística da criança em diversos domínios do quotidiano, nomeadamente sobre: (i) a frequência de contacto semanal com a língua de herança por meio de atividades relacionadas com os *media* (i.e.,

ver televisão, ler livros, etc.); (ii) a participação em atividades culturais e de lazer promovidas pela comunidade emigrante; (iii) o uso da língua de herança na convivência com amigos falantes desta língua residentes no país de acolhimento; e (iv) a frequência do ensino formal da língua de herança.

Para calcular o valor desta variável composta, o investigador deve, primeiramente, pontuar a experiência sociolinguística da criança nos quatro domínios do quotidiano, seguindo os critérios apresentados na Tabela 5, e, de seguida, somar os pontos atribuídos e dividir o total pela pontuação máxima possível (i.e., 40).

TABELA 5 – Medidas quantitativas para a pontuação dos domínios que compõem a variável RIQUEZA.INPUT.

<i>Domínios (Questões)</i>	<i>Categorias</i>	<i>Medidas quantitativas</i>	<i>Pontuação máxima possível</i>
(i) <i>Media</i> (n.º 18)	Nunca	0	25 pontos <sup>a</sup>
	1 dia por semana	1	
	2 dias por semana	2	
	3 ou 4 dias por semana	3	
	5 ou 6 dias por semana	4	
	Todos os dias	5	
(ii) <i>Atividades<sup>b</sup></i> (n.º 22.1)	Nenhuma	0	7 pontos
	Uma atividade	1	
	Duas atividades	2	
	Três atividades	3	
	Quatro atividades	4	
	Cinco atividades	5	
	Seis atividades	6	
	Mais de seis atividades	7	
(iii) <i>Amigos</i> (n.º 19 e n.º 21)	Não tem amigos ou Apenas X	0	4 pontos
	Mais X do que Y	1	
	Y e X em quantidades iguais	2	
	Mais Y do que X	3	
	Apenas Y	4	

(iv) Ensino formal (n.º 22.1, n.º 24 e n.º 25)	A criança recebe pouca ou nenhuma instrução formal em Y.	0	4 pontos
	A criança frequenta aulas de Y, de carácter extracurricular, uma vez por semana ou até 1h30m por semana.	1	
	A criança frequenta aulas de Y, de carácter extracurricular, mais do que uma vez por semana ou mais do que 1h30m por semana.	2	
	A criança frequenta uma escola bilingue em que as aulas são dadas maioritariamente na língua do país de acolhimento.	3	
	A criança frequenta uma escola bilingue em que as aulas são dadas maioritariamente em Y.	4	

Nota. <sup>a</sup> Máxima pontuação possível tendo em conta o número de atividade listadas (= 5); <sup>b</sup> As atividades relacionadas com o ensino formal da LH não são contabilizadas neste parâmetro, mas no (iv); X = língua do país de acolhimento; Y = língua de herança.

O resultado desta operação pode variar entre o mínimo de 0 e o máximo de 4, sendo que, quanto mais próximo de 4 estiver o valor, maior será a diversidade e a complexidade do *input* que a criança recebe no quotidiano.

## 5. Considerações finais

O presente instrumento de recolha de dados sobre a experiência linguística de crianças bilingues pretende colmatar a inexistência de um instrumento deste género em língua portuguesa, desenvolvido para o estudo de crianças lusodescendentes, mas com o potencial de ser adaptado ao estudo de falantes de herança de outras variedades do português.

No presente artigo focámos a sua potencialidade de medição do contacto que as crianças bilingues, oriundas de famílias emigrantes (portuguesas), têm com a sua LH. Neste sentido, foi detalhado o procedimento de quantificação das quatro variáveis que têm sido apontadas como sendo

as mais influentes no desenvolvimento da língua de origem. Contudo, o questionário permite, também, extrair informação sobre o grau de contacto com a língua maioritária através da análise de variáveis como o início de exposição à língua maioritária, o número de elementos do agregado familiar que usam a língua maioritária na comunicação doméstica e a escolarização, exclusiva ou não, nessa língua. Uma vez que, em regra, a língua maioritária é a principal língua de socialização da criança fora da família e, como discutido na introdução, a variação na quantidade de *input* desta língua é menor do que na da LH, a influência de variáveis ambientais sobre o desenvolvimento da língua maioritária tem um peso consideravelmente menor do que na língua minoritária (Gathercole & Thomas 2009), sobretudo na faixa etária coberta pelo questionário apresentado (6 a 10 anos). Como tal, não é recomendado para estudos focados no desenvolvimento bilingue em faixas etárias mais novas (até aos 5 anos), que pretendem investigar a influência dos fatores individuais sobre o desenvolvimento de ambas as línguas da criança bilingue.

## REFERÊNCIAS

- Benmamoun, E., Montrul, S. & Polinsky, M. 2013. Heritage languages and their speakers: opportunities and challenges for linguistics. *Theoretical Linguistics*. 39(3-4): 129-181.
- Bialystok, E., Luk, G., Peets, K. F. & Yang, S. 2010. Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*. 13: 525-531.
- Bohman, T. M., Bedore, L. M., Peña, E. D., Mendez-Perez, A. & Gillam, R. B. 2010. What you hear and what you say: Language performance in Spanish-English bilinguals. *International Journal of Bilingual Education and Bilingualism*. 13(3): 325-344.
- Bridges, K. & Hoff, E. 2014. Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics*. 35: 225-241.
- Bylund, E. & Díaz, M. 2012. The effects of heritage language instruction on first language proficiency: a psycholinguistic perspective. *International Journal of Bilingual Education and Bilingualism*. 15(5): 593-609.
- Chondrogianni, V. & Schwartz, R. G. 2020. Case marking and word order in Greek

- heritage children. *Journal of child language*. 47(4): 766-795.
- Christensen, L. B., Johnson, R. B. & Turner, L. A. 2015. *Research methods, design, and analysis*. Harlow, England: Pearson.
- Cobo-Lewis, A., Pearson, B., Eilers, R. & Umbel, V. 2002. Effects of bilingualism and bilingual education on oral and written Spanish skills: a multifactor study of standardized test outcomes. In: D. K. Oller & R. Eilers (Eds.). *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters, 98-117.
- Codó, E. 2008. Interviews and Questionnaires. In: L. Wei & M. G. Moyer (Eds.). *The Blackwell guide to research methods in bilingualism and multilingualism*. Malden, MA: Blackwell Publishing, 158-176.
- Cohen, C., Bauer, E. & Minniear, J. 2021. Exploring how language exposure shapes oral narrative skills in French-English emergent bilingual first graders. *Linguistics and Education*. 63: 100905. <https://doi.org/10.1016/j.linged.2021.100905>
- Correia, L. em preparação. *A influência de fatores extralinguísticos sobre o desenvolvimento lexical de crianças bilíngues. Um estudo sobre falantes de herança do Português Europeu residentes em diferentes contextos migratórios*. Tese de doutoramento. Universidade do Minho.
- Correia, L. & Flores, C. 2017. The Role of Input Factors in the Lexical Development of European Portuguese as a Heritage Language in Portuguese-German Bilingual Speakers. *Languages*. 2: 30. <https://doi.org/10.3390/languages2040030>
- Correia, L., Flores, C. & Melo-Pfeifer, S. (submetido). O Ensino (do) Português na Alemanha: um foco nos falantes de herança. In: F.C. Olmo, S. Melo-Pfeifer & S. Souza (Eds.). *O que quer, o que pode esta língua: contextos, estatutos e práticas de ensino do Português Língua Não Materna numa visão crítica*. Porto: Universidade do Porto.
- Cote, L. R. & Bornstein, M. H. 2014. Productive vocabulary among three groups of bilingual American children: comparison and prediction. *First Language*. 34: 467-485.
- De Houwer, A. 2017. Minority Language Parenting in Europe and Children's Well-Being. In: N. J. Cabrera & B. Leyendecker (Eds.). *Handbook on Positive Development of Minority Children and Youth*. Cham, Suíça: Springer, 231-246.
- Dörnyei, Z. & Taguchi, T. 2010. *Questionnaires in Second Language Research: Construction, Administration and Processing* (2.<sup>a</sup> ed.). New York: Routledge.
- Flores, C. & Barbosa, P. 2014. When reduced input leads to delayed acquisition: a study on the acquisition of clitic placement by Portuguese heritage speakers. *International Journal of Bilingualism*. 18(3): 304-325.

- Flores, C., Santos, A. L., Jesus, A. & Marques, R. 2017. Age and input effects in the acquisition of mood in Heritage Portuguese. *Journal of Child Language*. 44(4): 795-828.
- Garraffa, M., Beveridge, M. & Sorace, A. 2015. Linguistic and Cognitive Skills in Sardinian-Italian Bilingual Children. *Frontiers in Psychology*. 6: 1898. <https://doi.org/10.3389/fpsyg.2015.01898>
- Gathercole, V. C. M. 2002. Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction. In: D. K. Oller & R. Eilers (Eds.). *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters, 207-219.
- Gathercole, V. C. M. 2014. Bilingualism matters: One size does not fit all. *International Journal of Behavioral Development*. 38: 359-366.
- Gathercole, V. C. M. & Thomas, E. M. 2007. Factors contributing to language transmission in bilingual families: the core study-adult interviews. In: V. C. M. Gathercole (Ed.). *Language Transmission in Bilingual Families in Wales*. Cardiff: Welsh Language Board, 59-181.
- Gathercole, V. C. M. & Thomas, E. M. 2009. Bilingual first-language development: dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*. 12: 213-237.
- Gollan, T. H., Starr, J. & Ferreira, V. S. 2015. More than use it or lose it: The number-of-speakers effect on heritage language proficiency. *Psychonomic Bulletin & Review*. 22(1): 147-155.
- Gutiérrez-Clellen, V. F. & Kreiter, J. 2003. Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*. 24(2): 267-288.
- Ishizawa, H. 2004. Minority language use among grandchildren in multigenerational households. *Sociological Perspectives*. 47(4): 465-483.
- Kaščelan, D., Prevost, P., Serratrice, L., Tuller, L., Unsworth, S. & De Cat, C. 2020. A Review of questionnaires quantifying bilingual experience in children: do we document the same constructs? Ms. <https://doi.org/10.31219/osf.io/nry2s>
- Makarova, V., Terekhova, N. & Mousavi, A. 2019. Children's language exposure and parental language attitudes in Russian-as-a-heritage-language acquisition by bilingual and multilingual children in Canada. *International Journal of Bilingualism*. 23(2): 457-485.
- Meisel, J. 2008. Child second language acquisition or successive first language acquisition? In: B. Haznedar & E. Gavrusseva (Eds.). *Current trends in child second language acquisition: a generative perspective*. Amsterdam: John Benjamins Publishing, 55-80.

- Montrul, S. 2016. *The acquisition of heritage languages*. Cambridge: Cambridge University Press.
- Paradis, J. 2011. Individual differences in child English second language acquisition: comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*. 1(3): 213-237.
- Pearson, B. Z. 2007. Social factors in childhood bilingualism in the United States. *Applied Psycholinguistics*. 28(3): 399-410.
- Place, S. & Hoff, E. 2016. Effects and noneffects of input in bilingual environments on dual language skills in 2 ½-year-olds. *Bilingualism: Language and Cognition*. 19(5): 1023-1041.
- Polinsky, M. 2015. Heritage languages and their speakers: State of the field, challenges, perspectives for future work, and methodologies. *Zeitschrift für Fremdsprachwissenschaft*. 26: 7-27.
- Rezzonico S., Goldberg A., Milburn T., Belletti A. & Girolametto L. 2017. English Verb Accuracy of Bilingual Cantonese-English Preschoolers. *Language, Speech, and Hearing Services in Schools*. 48(3): 153-167.
- Rodina, Y. & Westergaard, M. 2017. Grammatical gender in bilingual Norwegian-Russian acquisition: the role of input and transparency. *Bilingualism: Language and Cognition*. 20(1): 197-214.
- Rothman, J. 2009. Understanding the Nature and Outcomes of Early Bilingualism: Romance Languages as Heritage Languages. *International Journal of Bilingualism*. 13(2): 155-63.
- Schmid, M. S. & Dusseldorp, E. 2010. Quantitative analyses in a multivariate study of language attrition: the impact of extralinguistic factors. *Second Language Research*. 26(1): 125-160.
- Sorenson Duncan, T. & Paradis, J. 2020. How does maternal education influence the linguistic environment supporting bilingual language development in child second language learners of English? *International Journal of Bilingualism*. 24(1): 46-61.  
<https://doi.org/10.1177/1367006918768366>
- Tuller, L. 2015. Clinical use of parental questionnaires in multilingual contexts. In: S. Armon-Lotem, J. D. Jong & N. Meir (Eds). *Assessing multilingual children: disentangling bilingualism from language impairment*. Bristol: Multilingual Matters, 301-30.
- Unsworth, S. 2012. Quantity-oriented and quality-oriented exposure variables in simultaneous bilingual acquisition. In: N. de Jong, K. Juffermans, M. Keijzer & L.

- Rasier (Eds.). *Papers of the Anéla 2012 Applied Linguistics Conference*. Delft, The Netherlands: Eburon, 13-22.
- Unsworth, S. 2013. Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: the case of Dutch gender. *Bilingualism: Language and Cognition*. 16(1): 86-110. <https://doi.org/10.1017/S1366728912000284>
- Unsworth, S. 2016a. Quantity and Quality of Language Input in Bilingual Language Development. In: E. Nicoladis & S. Montanari (Eds.). *Bilingualism Across the Lifespan*. Berlin, Boston: De Gruyter, 103-122.
- Unsworth, S. 2016b. *Bilingual Language Experience Calculator. Manual: Background and instructions for use*. Retirado da Internet a 26-03-2021. <https://www.iris-database.org/iris/app/home/detail?id=york%3a928327&ref=search>
- Unsworth, S. 2019. Quantifying Language Experience in Heritage Language Development. In: M. Schmid & B. Köpke. (Eds.). *The Oxford handbook of first language attrition*. Oxford: Oxford University Press, 434-445.
- Zhang, D. 2008. *Between two generations: Language maintenance and acculturation among Chinese immigrant families*. Texas, USA: LFB Scholarly Publishing LLC.

Recensões



Elena Castroviejo, Louise McNally, Galit Weidman  
Sasson (Eds.). *The semantics of gradability, vagueness,  
and scale structure*. Springer. 2018. 299 p. ISBN: 978-3-  
319-77790-0

Inês Cantante  
cantante.ines@gmail.com  
(Centro de Linguística da Universidade do Porto)

As editoras deste livro são autoras reconhecidas nas áreas da semântica e da pragmática, em particular no que respeita ao estudo de predicados graduáveis e da interação entre os fatores semântico-pragmáticos no processamento da linguagem e os processos cognitivos que a envolvem. Ao focar os temas da graduabilidade, estrutura escalar e vagueza, o volume compõe uma obra que utiliza um vasto conjunto de dados experimentais para apresentar conclusões acerca de temas tão vastos como a multidimensionalidade dos adjetivos, a interpretação de predicados vagos ou a subjetividade de expressões graduáveis.

O primeiro capítulo constitui uma breve introdução, pelas editoras, da motivação e necessidade de uma obra como esta, confrontando as várias perspetivas que abordam os temas em foco, terminando com uma breve explicação da organização da obra e um excelente resumo, por capítulo, que, mantendo uma estrutura sucinta, é bastante preciso, sendo, por essas razões, facilmente acompanhado pelo leitor.

O segundo capítulo, de Paul Égré e Jérémy Zehr, denominado “Are gaps preferred to Gluts? A closer look at borderline contradictions”, analisa casos de *borderline contradictions* (cf. Ripley, 2011), i.e., casos em que “one feels equally attracted toward applying and toward denying the predicate” (p. 26), e que podem ocorrer em construções conjuntivas (“x is P and not P”), ou disjuntivas (“x is neither P nor not P”) (pp. 26-27). Os resultados revelaram uma tendência para o uso de formas disjuntivas, configurando esta uma “preference for incompleteness over inconsistency” (p. 40), que privilegia, dessa forma, a informatividade da escolha. Talvez a proposta

mais inovadora neste estudo seja a de que a atribuição do significado é feita a nível local e não global, havendo assim uma redução do nível da frase para o nível das orações que a constituem.

O terceiro capítulo, “Multidimensionality, Subjectivity and Scales: Experimental Evidence”, de Stephanie Solt, aborda a forma como a multidimensionalidade dos adjetivos se relaciona com questões de subjetividade (e escalaridade), assumindo-se que esta constitui uma fonte de ordenação para os níveis de subjetividade, já que “different individuals may weight these component dimensions differently” (p.61). No decorrer da sua investigação, a autora propõe que a ordenação dos valores pode ter origem num de dois fatores: multidimensionalidade do adjetivo ou julgamento pessoal do falante, com caráter avaliador. De acordo com estas duas fontes de subjetividade, os adjetivos podem, então, subdividir-se em, não dois, mas três subtipos: objetivos, subjetivos ou mistos.

O capítulo seguinte, “Online Processing of ‘Real’ and ‘Fake’: the cost of being too strong”, utiliza dados do Alemão, continuando ligado à temática da multidimensionalidade, ao analisar nomes modificados pelos adjetivos “real” (*verdadeiro*) e “fake” (*falso*). Os autores, P. Schumacher, P. Brandt e H. Weiland-Breckle, defendem a ideia de que adjetivos como ‘fake’, que introduzem uma contradição (entre adjetivo e nome), obrigam a um maior tempo de processamento, por parte dos falantes, já que exigem uma espécie de ‘reparo’, ou seja, re-arranjo conceptual. Este capítulo torna-se interessante pelo facto de, ao testar o processamento *online* dos falantes, apresentar uma proposta que se relaciona com a de Égré & Zehr (Cap. 2): os falantes evitam inconsistências.

No quinto capítulo, “Gradable Nouns as Concepts without Prototypes”, Hanna de Vries expõe uma perspetiva muito interessante, que combina uma proposta de Kamp & Partee (1995), a Teoria dos Protótipos (i.e., certos conceitos poderão estar associados a características prototípicas e “the prototype of a concept C is defined as the abstract entity which is most C-like” (p. 116)), com a graduabilidade de adjetivos (cf. Kennedy & McNally 2005). De Vries testa a hipótese de que adjetivos com valor máximo, como *limpo*, têm protótipos, enquanto adjetivos absolutos sem valor máximo, como *sujo*, não têm, propondo uma distinção entre esta ‘prototypicality’ (p. 114) e a noção de acessibilidade (termos mais presentes na memória dos falantes são

mais facilmente processados). Apesar de pressupor alguns conhecimentos prévios, este capítulo é muito completo, pois são combinados os conceitos de vagueza, graduabilidade e escalas abertas e/ou fechadas.

O capítulo seguinte, “Education as a source of vagueness in criteria and degree”, de S. Verheyen e G. Storms, analisa, de uma perspetiva sociolinguística, o processo de categorização aplicado a predicados vagos. O objetivo é observar como é que as diferentes competências linguísticas dos falantes, decorrentes de um maior ou menor grau de escolaridade, se relacionam com a interpretação de conceitos vagos, com base em duas origens para a sua indeterminação: vagueza de grau ou vagueza de critérios. Os resultados obtidos permitem afirmar que o nível de educação dos participantes parece, pelo menos no âmbito muito específico deste estudo, ter efeitos na categorização dos termos. Apesar de se tratar de um estudo de fácil compreensão, é muito específico, pelo que os resultados poderão estar muito dependentes dos estímulos dados aos participantes e dos participantes em si, constituindo, ainda assim, uma boa investigação inicial.

No capítulo “Intensification, gradability and social perception: the case of totally”, A. Beltrama pretende avaliar se o significado social associado a uma expressão, profundamente dependente do contexto social, pode, também, estar associado a fatores linguísticos. Para isso, o autor selecionou o modificador (intensificador) ‘*totally*’ (*totalmente*), pelo facto de ter a particularidade de conjugar características sociais – ao marcar uma certa proximidade social entre os falantes – e semânticas, já que é utilizado como forma de mostrar que o topo máximo da escala foi atingido (cf. Kennedy & McNally 2005) ou, de acordo com outras propostas (cf., e.o, Cantante 2018; 2020), como forma de reforço epistémico, que, segundo Beltrama, revela “individual certainty of the speaker towards the truth of the proposition” (p. 173). De facto, os resultados permitem concluir que este último caso, ‘*speaker-oriented totally*,’ facilita a atribuição de um significado social saliente à expressão, o que comprova que as propriedades semânticas têm influência na construção de um significado social. Este estudo tem particular interesse pelo facto de constituir um ponto de ligação entre abordagens exclusivamente semânticas e abordagens sociais.

Em “Perceived informativity and referential effects of contrast in adjectivally modified NPs”, da autoria de H. Aparicio, C. Kennedy e M.

Xiang, é analisada a distinção entre adjetivos absolutos e relativos, com a qual o leitor, neste ponto do livro, deverá estar já ambientado. Focando nomes modificados por adjetivos, o objetivo dos autores passa por compreender a forma como os falantes de uma língua atribuem referência às entidades designadas por esses nomes. Mais concretamente, é avaliado o efeito que a contrastividade poderá ter na atribuição de um referente ao nome modificado, normalmente associado à informatividade do modificador (cf. Sedivy *et al.* 1999). Os resultados parecem evidenciar uma estreita relação entre diversos fatores: a informatividade do adjetivo, fatores pragmáticos e fatores léxico-semânticos, diretamente relacionados com a estrutura escalar dos adjetivos. A investigação permitiu, ainda, comprovar que o processamento semântico é incremental, construído em estreita relação com o contexto (quando o contexto é mais restritivo, a interpretação fica facilitada).

O capítulo seguinte, “Modified Fractions, Granularity and Scale Structure”, por Chris Cummins, analisa expressões de quantidade (numéricas) para verificar se os números envolvidos e a informação dada pela frase (ou seja, o contexto linguístico) têm influência nas interpretações dos falantes. Para isso, Cummins testou a forma como os participantes interpretavam frases com frações, com o objetivo de avaliar a proposta de Krifka (2002) de que a granularidade das escalas poderá ter repercussões na interpretação dos ouvintes: escalas com granularidade mais fina dão informações mais precisas, mas deverão ter maior custo cognitivo para o falante. Os resultados mostram que, mais importante do que a granularidade, no momento da decisão, deverá ser o conjunto de possibilidades/alternativas apresentadas ao falante pelo contexto pragmático. A ideia final deste estudo é a de que a noção de granularidade, herdada de Krifka (2002), tem de ser refinada para dar conta destas questões complexas. Apesar de ser um bom ponto de partida, a avaliação de frações poderá constituir uma limitação desta investigação, dado que os participantes podem não ter, todos, o mesmo domínio sobre a variável em análise.

O penúltimo capítulo do livro, “Decomposition and Processing of Negative Adjectival Comparatives”, analisa a ideia de que os adjetivos negativos (como *short*) são mais complexos do que as correspondentes formas positivas, porque deverão ser constituídos por duas partes: a forma positiva do adjetivo (por exemplo, *tall*) e um morfema de negação (para o

inglês, 'little') (ou seja, *short* deverá ser um complexo do tipo *little tall*) (cf. Büring 2007; Clark & Chase 1972). Os autores, D. Tucker, B. Tomaszewickz e A. Wellwood, pretendem evidenciar que esta decomposição, a existir, deverá refletir-se num maior tempo de processamento por parte dos falantes. Além de realizarem testes em duas línguas diferentes, Inglês e Polaco, os autores testam a mesma questão utilizando uma linguagem matemática ( $A > B$ ), para verificar se apenas fatores linguísticos estão em jogo no momento do processamento. Os resultados abonam a favor da proposta da decomposição, mostrando, além disso, que os falantes traduzem as notações matemáticas para a sua própria língua.

Por fim, o livro termina com um estudo, de Rick Nouwen e Kajub Dotlacil, "Cumulative Comparison: Experimental Evidence for Degree Cumulation". Havendo casos em que o comportamento dos graus parece aproximar-se do comportamento das entidades, nomeadamente casos de frases comparativas com *than-clauses*, os autores pretendem avaliar se, tal como as entidades, também os graus podem desencadear mecanismos de leituras distributivas ou cumulativas. Os resultados mostram que frases com estruturas comparativas (*than-clauses*) parecem, efetivamente, introduzir relações cumulativas entre graus, o que comprova a existência de pluralidades de graus.

De um modo geral, este livro apresenta uma perspetiva muito abrangente e atual acerca do estudo da graduabilidade de um ponto de vista experimental. Ao agregar um conjunto de estudos muito diversos entre si, mas com um fio condutor comum, em que os resultados são comprovados empiricamente, o leitor pode, através do acompanhamento atento de cada capítulo, chegar a conclusões acerca das perspetivas apresentadas. O facto de ser constituído por capítulos curtos que, em cada caso, apresentam as perspetivas teóricas essenciais para a sua fundamentação (apesar de, em certos casos, haver necessidade de alguns conhecimentos prévios), tornam a leitura deste livro bastante agradável.

## REFERÊNCIAS

- Büring, D. 2007. Cross-polar nomalies. In: T. Friedman & M. Gibson (Eds.). *Proceedings of Semantics and Linguistic Theory 17*. Connecticut: Linguistic Society of America, 37-52.
- Cantante, I. 2018. *Sobre a semântica dos adjetivos adverbiais modais*. Dissertação de Mestrado. Faculdade de Letras da Universidade do Porto.
- Cantante, I. 2020. Sobre a escalaridade dos adjetivos adverbiais modais. *Linguística*. 15: 41-70.
- Clark, H. H. & Chase, W. G. 1972. On the process of comparing sentences against pictures. *Cognitive Psychology*. 3(3): 472-517.
- Kamp, H. & Partee, B. 1995. Prototype theory and compositionality. *Cognition*. 57(2): 129-191.
- Kennedy, C. & McNally, L. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language*. 81(2): 345-381.
- Krifka, M. 2002. Be brief and vague! And how bidirectional optimality theory allows for verbosity and precision. In: D. Restle & D. Zaefferer (Eds.). *Sounds and systems: Studies in structure and change, a Festschrift for Theo Vennemann*. Berlin: Mouton de Gruyter, 439-458.
- Ripley, D. 2011. Contradictions at the borders. In: R. Nouwen, R. van Rooij, U. Sauerland & H.-C. Schmitz (Eds.). *Vagueness in communication*. Lecture Notes in Computer Science Berlin: Springer, Vol. 6517, 169-188.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G. & Carlson, G. N. 1999. Achieving incremental semantic interpretation through contextual representation. *Cognition*. 71(2): 109-147.

Philipp Cimiano; Christian Chiarcos; John P. McCrae; Jorge Gracia (2020). *Linguistic Linked Data. Representation, Generation and Applications*. Springer International Publishing. ISBN 978-3-030-30225-2

Purificação Silvano  
msilvano@letras.up.pt

*Centro de Linguística da Universidade do Porto, Faculdade de Letras da  
Universidade do Porto (Portugal)*

## 1. Background

During the last decades, there has been a proliferation of language resources, which are of utmost relevance to research in many areas from Linguistics to Natural Language Processing (NLP). However, frequently these resources lack structural and conceptual interoperability (Chiarcos et al. 2020), which are key conditions warranted in a Semantic Web where multilingual information from different resources should be more easily accessed, identified, and interpreted. Linguistic Linked Open Data (LLOD) enables such a reality.

The LLOD Infrastructure, envisioned by a working group of Open Language Foundation, the Open Linguistics Working Group (OWLIG) (Chiarcos et al. 2011), combines two prior existing concepts, open data and linked data, and applies them to language resources. As other Web resources, linguistic resources must comply with the following Linked Open Data's directrices regarding publication and representation: (a) data openly licensed; (b) entities referred by URIs (Uniform Resource Identifier); (c) URIs resolvable over HTTP (Hypertext Transfer Protocol); (d) data represented by web standards, such as HTML (Hypertext Markup Language), RDF (Resource Description Framework) or JSON-LD (Javascript Object Notation); and (e) resources linked to each other (Chiarcos et al. 2013). These Linked Data Principles (Bizer et al. 2009) ensure homogeneity of the linguistic resources, and thus, a more global access to them, which, in turn, can lead to more

cooperation across disciplines (Chiarcos *et al.* 2020). In fact, in this manner, LLOD provides several benefits such as representation, by means of linked graphs, which allow a more adaptable representation format for linguistic data; interoperability; expressivity, due to vocabularies like Web Ontology Language (OWL) and Lemon; and dynamicity, as a result of the perpetual enhancement of web data (Chiarcos *et al.* 2013; <https://linguistic-lod.org>).

Throughout the years, the LLOD infrastructure has been enriched with contributions from several projects. *Creating knowledge out of Interlinked Data* (LOD2) was one of the first (2010-2014) to contribute to the development of LLOD infrastructure. More recently, *Ready-to-use Multilingual Linked Language Data for Knowledge Services across Sectors* (Prêt-à-LLOD), a European H2020 project, has been working towards further expansion of LLOD infrastructure, and its sustainability. Towards these goals, Prêt-à-LLOD is developing methods to discover, transform and link linguistic data to be published as LLOD, and is demonstrating the usability and efficacy of those methods in real-world problems of language technology industries (e.g. Semantic Web Company, Oxford University Press) by means of use case pilots (Declerck *et al.* 2020). Prêt-à-LLOD has also been collaborating closely with *European Lexicographic Infrastructure* (ELEXIS), which is creating a dictionary matrix based on LLOD. The COST action *European Network for Web-Centred Linguistic Data Science* (NexusLinguarum) aims to promote the construction of a global ecosystem of multilingual and semantically interoperable linguistic data by developing linked data technologies, combined with NLP techniques and multilingual resources (<https://nexuslinguarum.eu/the-action>), which will be of great value to the LLOD Infrastructure.

The LLOD Infrastructure has grown so significantly that, first, in 2014, the category “linguistics” was integrated in the Linked Open Data cloud, and second, in 2018, the LLOD cloud was incorporated as a domain-specific addendum (<https://linguistic-lod.org>). In fact, the LLOD cloud’s increase has been higher than the LOD cloud (Chiarcos *et al.* 2020). At the moment of the writing, the language resources available in the LLOD cloud are 220, distributed in the following categories: corpora, lexicons and dictionaries, terminologies, thesauri and knowledge bases, linguistic resource metadata, linguistic data categories, typological databases and others.

The appeal of LLOD is substantiated by the increasing community of researchers working not only towards the creation or conversion of digital linguistic resources into linked data (LD), but also towards the development of standards and models to do so. Moreover, LLOD has stirred up the interest of researchers who are eager to learn more about this paradigm. It is in this regards that *Linguistic Linked Data. Representation, Generation and Applications*, by Philipp Cimiano, Christian Chiarcos, John P. McCrae and Jorge Gracia, who have authored seminal works in LLOD, is a must-read book.

## 2. The book

This book offers a thorough description of LLOD to students, teachers and researchers from Linguistics and Computational Sciences who work with language resources and are keen on learning about guiding principles and techniques to create, reconvert and publish those resources as linked data, and about possible applications of LLD resources. The book summarizes a wide array of contents about different aspects of LLD with clear explanations, illustrative examples, and references for additional information.

The book is very well organized in five parts: *Preliminaries*, *Modelling*, *Generation and Exploitation*, *Use Cases* and *Conclusions*. The first part is key to understand the rest of the book, as stated by the authors, because it presents the basic concepts underlying LLD. The second and third parts explain how to model language resources into LD and how to represent and to discover linking between datasets. The fourth part is dedicated to the description of different applications of LLD. The fifth part encompasses the conclusions. As one can infer from this overview, the book is quite overarching covering the necessary theoretical principles to comprehend the subject, but also mentioning use cases to demonstrate the benefits of adhering to the LLOD model.

Zooming in the structure of these five parts, each part comprises a number of chapters, and each chapter is divided into sections preceded by an *Abstract* and followed by *Summary and Further Reading*, and *References*. The abstract is very helpful, allowing the reader to gather more information

about the chapter's content before reading it. The summary gives the reader a systematization of the main points discussed in the chapter, which is always useful to consolidate what the reader learnt, while the further reading part specifies the references to broaden the reader's knowledge about the issues explored in the chapter.

The first three chapters integrate part I – *Preliminaries*. Chapter 1 demonstrates how LD is the perfect scenario to attain the *FAIR Principles*, that are, *Findability*, *Accessibility*, *Interoperability* and *Reusability*. In fact, the already mentioned Linked Open Data's directrices guarantee that language resources are findable, accessible, interoperable, and reusable. Chapter 2 focuses on RDF, the quintessentially format for LLD, describing its semantics and its most well-known serializations (e.g. Turtle (Terse RDF Triple Language), JSON-LD). The contribution of OWL to represent ontological knowledge, and the query language SPARQL (Protocol and RDF Query Language) are also explained. It goes without saying that the foundational nature of this chapter means that if the reader's intention is to model language resources resorting to RDF or Turtle, work with OWL or carry out queries using SPARQL, further study and practice are needed. The last chapter of the first part provides the motivation for LLD and outlines its beginning, the story behind LLOD cloud, and the substantial work of the LLOD community, namely of OWLG.

Chapters 4 to 8 compose part II – *Modelling*. The first chapter of this part gives a detailed account of the Ontolex-lemon, a generic lexicon model, which was developed by the very active Ontology-Lexicon Community Group, and which enables the representation of lexical information as linked data. Its popularity is proven by the large collection of dictionaries that use the lemon model (cf. DBpedia, Apertium, among others). Chapter 5 instructs the reader on how to annotate data according to RDF principles. It starts with a summing up of annotation languages used by NLP and Digital Humanities, such as the CoNLL format and TEI (Text Encoding Initiative), showing how these can be adapted to create LLD. The following section proceeds to discussing two compatible RDF-based formalisms for referencing data on the web: Web Annotation, and a simpler alternative, that is NIF (NLP Interchange Format). Chapter 6 complements the previous chapter by elaborating on the necessary vocabularies to represent linguistic

annotations in an interoperable manner, namely those grounded on OWL and RDF models, such as CoNLL-RFD and POWLA. The technologies and techniques to query the resulting annotated corpora are also revised. Since it is important to encode not only linguistic data, but also the information about the language resources that contain those linguistic data, chapter 7 refers the models for the representation of language resources metadata, so that they can be easily found and reused, two essential values of LLOD. It is the case of DC-Terms (DCMI Metadata Terms vocabulary) for general metadata (title, license, description creator...) and Meta-Share.owl ontology for information pertaining to the language resources (resource type, modality, number of languages...). The last chapter of the second part focuses on strategies to achieve conceptual interoperability, being one of those the laborious undertaking of harmonizing linguistic categories through the creation of general terminology repositories like ISOcat, and more specific ones, such as OLiA (Ontologies of Linguistic Annotation).

The next group of chapters (9-11) constitute part III – *Generation and Exploitation*. Chapter 9 takes the reader through the process of publishing LLD, itemizing and illustrating each of the six steps: (i) specification – analysis and description of data sources; (ii) modelling – creation or selection of vocabularies to represent in RDF; (iii) generation – production of RDF datasets from the data sources; (iv) linking – connection of RDF datasets; (v) publication – release of the datasets; (vi) exploitation – development of applications that use the datasets. Being at the core of LLD paradigm the linking concept, the techniques to represent the links between datasets in the same language and across languages, and to discover those links are, obviously, quite pertinent. For this reason, chapter 10 is fully devoted to link representation and discovery. Chapter 11 argues in favor of using the models compliant with LLD formalisms, like NIF, for creating NLP pipelines, and it also exemplifies how this workflow can be operationalized.

Finally, chapters 12-14 form part IV – *Use Cases*. The linking of multilingual wordnets is addressed in chapter 12, while the following chapter highlights the prominence of LOD in Digital Humanities. In both cases, the reader becomes acquainted with the broad range of applications that integrate LLOD technologies, in particular in the field of Digital Humanities. In chapter 14, the authors underline the hurdles of finding language

resources in the repositories, and of standardizing metadata pertaining to those repositories, and present Linghub, a portal that gathers metadata from different repositories, as a solution to overcome those obstacles.

Chapter 15 (Part IV) closes this incursion into the world of LLOD with a succinct recapitulation of the topics tackled in the book and with a glimpse into the required future steps towards the expansion of the LLOD infrastructure.

As it can be assumed from this brief synopsis, the book under review supplies the reader with an extensive survey about the topic of LLOD, divulging the models and best practices so that more researchers adhere to this paradigm, and together continue to invest in the development of the LLOD infrastructure.

#### REFERENCES

- Bizer, C., Heath, T. & Berners-Lee, T. 2009. Linked data – The story so far. *International Journal on Semantic Web and Information Systems*. 5: 1-22.
- Chiarcos, C., Hellmann, S. & Nordhoff, S. 2011. Towards a Linguistic Linked Open Data cloud: The Open Linguistics Working Group. *TAL (Traitement Automatique des Langues)*. 52(3): 245-275. Retrieved January 12, 2021, from the World Wide Web: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=9AAA230DDF3A6AA313FE939EF95C30AC?doi=10.1.1.377.2076&rep=rep1&type=pdf>
- Chiarcos, C., McCrae, J., Cimiano, P. & Fellbaum, C. 2013. Towards open data for linguistics: Linguistic linked data. In: Alessandro Oltramari, Piek Vossen, Lu Qin & Eduard Hovy (Eds.). *New Trends of Research in Ontologies and Lexical Resources: Ideas, Projects, Systems*. Springer, Berlin, Heidelberg, 7-25. Retrieved January 12, 2021, from the World Wide Web: <https://link.springer.com/content/pdf/10.1007%2F978-3-642-31782-8.pdf>
- Chiarcos, C., Klimek, B., Fäth, C., Declerck, T. & McCrae, J. P. 2020. On the Linguistic Linked Open Data Infrastructure. In: Georg Rehm, Kalina Bontcheva, Khalid Choukri, Jan Hajič, Stelios Piperidis & Andrejs Vasiļjevs (Eds.). *Proceedings of the 1<sup>st</sup> International Workshop on Language Technology Platforms (IWLTP 2020)*. Marseille, France: European Language Resources Association, 8-15. Retrieved January 12, 2021, from the World Wide Web: <https://www.aclweb.org/anthology/2020.iwltp-1.2/>

Declerck, T., McCrae, J. P., Hartung, M., Gracia, J., Chiarcos, C., Montiel-Ponsoda, E., Cimiano, P., Revenko, A., Sauri, R., Lee, D., Racioppa, S., Nasir, J., Orlikowski, M., Lanau-Coronas, M., Fäth, C., Rico, M., Elahi, M., Khvalchik, M., Gonzalez, M. & Cooney, K. 2020. Recent developments for the Linguistic Linked Open Data Structure. In: Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odiijk & Stelios Piperidis (Eds.). *Proceedings of the 12<sup>th</sup> Conference on Language Resources and Evaluation (LREC 2020)*. Marseille, France: European Language Resources Association, 5660-5667. Retrieved January 12, 2021, from the World Wide Web: <https://aclanthology.org/2020.lrec-1.695.pdf>

