WALTER TSUYOSHI SANO

A TRADUÇÃO DA ESCRITA
sistema de transposição ortográfica
do árabe padrão para o português brasileiro

V E R S ã O C O R R I G I D A

O exemplar original encontra-se disponível
no Centro de Apoio à Pesquisa Histórica (CAPH) da FFLCH.

Dissertação apresentada à Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo para obtenção do título de Mestre em Língua, Literatura e Cultura Árabe

Orientadora: Safa Alferd Abou Chahla Jubran

São Paulo

2011

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas
da Universidade de São Paulo

Sano, Walter Tsuyoshi

A tradução da escrita: sistema de transposição
ortográfica do árabe padrão para o português brasileiro /
Walter Tsuyoshi Sano ; orientadora Safa Alferd Abou
Chahla Jubran. — São Paulo, 2011.
197 f. ; il.

Dissertação (Mestrado) — Faculdade de Filosofia,
Letras e Ciências Humanas da Universidade de São Paulo.
Departamento de Letras Orientais. Área de concentração:
Língua, Literatura e Cultura Árabe.

1. Escrita (linguística). 2. Língua árabe. 3. Ortografia.
4. Português do Brasil. 5. Tradução. 6. Transliteração.
I. Título. II. Jubran, Safa Alferd Abou Chahla.

Dedicatória

A todas as pessoas que comigo
instantes e momentos compartilharam,
ideias e ideais partilharam.

Àqueles que, em um e outro momento,
apoiaram minhas iniciativas
e toleraram minhas idiossincrasias.

A quem sabe e entende seguir eu uma senda
ao lado de iguais, enquanto me ergo
sobre ombros de gigantes.

Agradecimentos pessoais

AOS MEUS PAIS, Walter Sano e Yuriko Sano, ambos professores, pela criação que recebi e pelo apoio incontestado que recebo na opção pela vida acadêmica.

À prof.^a dr.^a Safa Alferd Abou Chahla Jubran, por todo apoio e toda confiança, e pela atenção, paciência e compreensão ao longo do processo de elaboração e realização desta dissertação.

Aos professores e colegas das disciplinas cursadas —da graduação como da pós-graduação—, por fazerem da academia ambiente propício à nossa maturação intelectual¹.

Aos professores Mamede Mustafa Jarouche e Paulo Chagas de Souza, por terem composto a banca do exame de qualificação.

¹Agradecimentos especiais ao professor Marcos Lopes, por me apresentar o $\text{\LaTeX}$, sem o qual esta dissertação muito perderia tipograficamente; ao professor Clóvis Luiz Alonso Júnior, pela conferência da expressão em latim cunhada para uso na introdução (página 27); à colega Indaiá de Santana Bassani, pelo livro de NIMER (2005), usado em 4.2.5 (páginas 109–110); à colega Paula da Costa Caffaro, pelas gramáticas do árabe; ao colega Sérgio Ferreira Mendes, pela informação pertinente ao uso do ponto médio ⟨·⟩ em catalão, e ao informante Victor, pela confirmação do uso corrente da distinção entre ⟨LL⟩ e ⟨LL⟩.

Agradecimentos gerais

Às bibliotecárias² e aos demais funcionários da Biblioteca Florestan Fernandes, pelo cuidado dispensado à preservação e manutenção do acervo material e digital do conhecimento.

À Faculdade de Filosofia, Letras e Ciências Humanas (FFLCH-USP), pela oportunidade de realização do curso de mestrado em Língua, Literatura e Cultura Árabe.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de mestrado.

²Recentemente ingressou no quadro de funcionários um bibliotecário. Ainda assim, optamos por concordar o plural com a maioria feminina.

عالم بلا عمل مثل الغيم بلا مطر

«Saber sem produzir
é chover sem molhar.»

ditado árabe

Resumo

SANO WT (2011) **A tradução da escrita** — sistema de transposição ortográfica do árabe padrão para o português brasileiro. 197 pp. Dissertação (Mestrado) — Faculdade de Filosofia, Letras e Ciências Humanas; Universidade de São Paulo, São Paulo.

NESTA DISSERTAÇÃO tomamos por objeto de estudo a romanização, em português, de palavras árabes (e.g, termos, topônimos e antropônimos). Estudamos por que alguns nomes são grafados de maneiras diferentes —em um ou outro jornal, por exemplo—, analisamos esquemas existentes de romanização e propomos um sistema que seja não apenas conciso e regular, mas também abrangente e minucioso, podendo ser usado tanto por leigos quanto por acadêmicos. Os seis capítulos encontram-se divididos em duas partes. Na primeira parte, abordamos a relação entre transposição ortográfica (i.e, transcrição, transliteração romanização e latinização) e tradução; apresentamos a língua árabe e sua escrita; ressaltamos elementos pertinentes da ortografia do português brasileiro. Na segunda parte, cotejamos oito esquemas de romanização árabe-português, apontando suas (des)semelhanças; apresentamos nosso sistema de transposição ortográfica, em seus quatro módulos (leigo simplificado, leigo comum, acadêmico amplo e acadêmico estrito), incluindo exemplos de aplicação em nomes, expressões e um breve texto; concluímos antecipando os desdobramentos da proposta.

Palavras-chave: escrita; língua árabe; ortografia; português do Brasil; tradução; transliteração.

Abstract

SANO WT (2011) **The translation of writing** — a system for Standard Arabic to Brazilian Portuguese orthographic transposition. 197 pp. Dissertação (Mestrado) — Faculdade de Filosofia, Letras e Ciências Humanas; Universidade de São Paulo, São Paulo.

THIS MASTER'S THESIS presents our research concerning the romanization of Arabic words (e.g. technical terms, toponyms and anthroponyms) into Portuguese. We have first studied why the way some nouns and proper names are written varies from source to source, then analysed some of the existing romanization schemes and, finally, proposed a system both concise and regular as well as comprehensive and detailed, one which is suitable for laymen and scholars alike. There are six chapters, distributed in two parts. In the first part, we see four orthographic transposition methods (transcription, transliteration, romanization and latinization) and how they relate to translation; the Arabic language and its writing script; and relevant aspects of the Brazilian Portuguese orthography. In the second part, we compare eight Arabic-Portuguese romanization schemes, highlighting their (dis)similarities; introduce all four modules (layman's simplified, layman's common, scholar's broad and scholar's narrow) of our orthographic transposition system, including samples comprising names, expressions and a small text; and finish off with an assessment on how said proposal may unfold.

Keywords: Arabic language; Brazilian Portuguese; orthography; translation; transliteration; writing.

خلاصة

SANO WT (2011)

ترجمة الكتابة: نظام إملائي للنقل الحرفي للعربية في اللغة البرتغالية البرازيلية.
197 pp. Dissertação (Mestrado) — Faculdade de Filosofia, Letras e Ciências Humanas; Universidade de São Paulo, São Paulo.

إن الهدف من دراستنا في هذه الأطروحة هو كتابة الكلمات العربية (على سبيل المثال، المصطلحات وأسماء الأماكن وأسماء العلم) بالحروف اللاتينية، للإستخدام في اللغة البرتغالية. نقوم بالبحث عن أسباب تدوين بعض الأسماء بطرق مختلفة في صحيفة وأخرى مثلاً. نقوم أيضاً بتحليل أنظمة موجودة وباقتراح نظام كتابة بالحروف اللاتينية لا يتميز بالموجزية وبالانتظام فحسب بل بالشمولية وبالذقة أيضاً مما يمكن استخدامه من قبل الناس العاديين والأكاديميين. يتم تقسيم الفصول الستة في جزأين. في الجزء الأول، نناقش العلاقة بين النقل الإملائي (كالنسخ، والنقل الحرفي والكتابة بالحروف اللاتينية) وبين الترجمة ونقدم اللغة العربية ونظام خطها، بينما نسلط الضوء على عناصر ذات صلة في هجاء اللغة البرتغالية البرازيلية. أما في الجزء الثاني فنقارن بين ثمان مخططات لكتابة العربية بالحروف اللاتينية، مشيرين إلى (عدم) التشابه بينها، ثم نقدم اقتراحنا لنظام هجائي يعرض في أربع وحدات (عادي مبسط، وعادي مشترك، متخصص واسع ومتخصص محصور) بما في ذلك من أمثلة تطبيقية لأسماء وعبارات ونصوص قصيرة كما وأننا نستنتج مقدماً عواقب هذا الاقتراح.

الكلمات الرئيسية: الكتابة؛ اللغة العربية؛ الإملاء؛ البرتغالية في البرازيل؛ الترجمة؛ النقل الحرفي.

SANO WT (2011) **Tarjama^t al-kitāba^t**: niḍām ^qimlā^qiyy li-'n-naql al-ḥarfiyy li-'l-^carabiyya^t fi 'l-luġa^t al-burtuġaliyya^t al-barāziliyya^t. 197 pp. Dissertação (Mestrado) — Faculdade de Filosofia, Letras e Ciências Humanas; Universidade de São Paulo, São Paulo.

² ^Qinn al-hadaf min darāsati-nā fi haḍihi al-^qaṭrūḥat huwa kitāba^t al-kali_
³ māt al-^carabiyya^t (^calā sabīl al-miṭāl, al-muṣṭalḥāt wa-^qasmā^q al-^qamākin
⁴ wa-^qasmā^q al-^calam) bi-'l-ḥurūf al-lātiniyya^t, li-'l-^qistikḍm fi 'l-luġat al-
⁵ -burtuġaliyya^t. Naḵūmu bi-'l-baḥṭ ^can ^qasbāb tadwīn ba^cḍ al-^qasmā^q bi-
⁶ ṭurukmuḵtalifa^t fi ṣaḥīfa^t wa-^qukrā maṭal^{an}. naḵūmu ^qayḍ^{an} bi-taḥlīl ^qan_
⁷ ḍima^t mawjūda^t wa-bi-^qaḵtirāḥ niḍām kitābat bi-'l-ḥurūf al-lātiniyya^t lā
⁸ yatamayyizu bi-'l-mawjūziyya^t wa-bi-'l-^qantiḍām faḥasb bal bi-'t-tu_
⁹ mūliyya^t wa-bi-'d-daḳat ^qayḍ^{an} mimmā yumakkin ^qistikḍāma-hu min
¹⁰ ḳibal an-nās al-^cādiyyīn wa-al-^qakādīmiyyīn. Yatimmu taḵsīm al-fuṣūl
¹¹ as-sitta^t fi jiz^qayn. Fī 'l-jiz^q al-^qawwal, nunāḳix al-^calāka^t bayna an-naql
¹² al-^qimlā^qiyy (ka-'n-nasḳ, wa-an-naql al-ḥarfiyy wa-'l-kitābat bi-'l-ḥurūf
¹³ al-lātiniyyat) wa-bayna at-tarjama^t wa-nuḵaddim al-luġa^t al-^carabiyya^t
¹⁴ wa-niḍām ḳaṭṭi-hā, baynamā nusalliṭ al-^qaḍwā^q ^calā ^canāṣir ḍāta ṣilla^t
¹⁵ fī hijā^q al-luġa^t al-burtuġaliyya^t al-barāziliyya^t. ^Qammā fī 'l-jiz^q at-tānī
¹⁶ fanuḵārin bayna ṭamāni muḳaṭṭaṭāt li-kitābat al-^carabiyyat bi-'l-ḥurūf
¹⁷ al-lātiniyya^t, muxīrīna ^qilā (^cadami) at-taxābuh baynahā, ṭumma nuḳa_
¹⁸ ddim ^qiḵtirāḥi-nā li-naḍām hijā^qiyy yu^craḍ fī ^qarba^c waḥadāt (^cādiyy
¹⁹ mubassat, wa-^cādiyy muxtarak, mutakaṣṣiṣ wāsi^c wa-mutakaṣṣiṣ maḥ_
²⁰ ṣūr) bi-mā fī ḍalika min ^qamṭila^t taṭbīḳiyya^t li-^qasmā^q wa-^cinārāt wa-
²¹ -nuṣūṣ ḳaṣīrat kamā wa-^qinn-nā nasantij muḳaddam^{an} ^cawāḳib hāḍā al-
²² -^qaḵtirāḥ.

Al-kalimāt ar-ra^qiṣiyya^t: al-kitāba^t; al-luġa^t al-^carabiyya^t; al-^qimlā^q; al-
 -burtuġaliyya^t fī 'l-barāzīl; at-tarjaha^t; an-naql al-ḥarfay.

Sumário breve

Introdução 27–33

Parte primeira — da tradução e da escrita

- 1 Transposição ortográfica e tradução 35–58
- 2 A língua árabe e sua escrita 59–77
- 3 O português brasileiro e o alfabeto latino 78–93

Parte segunda — da tradução da escrita

- 4 A romanização do árabe 96–121
- 5 Proposta de transposição ortográfica 122–159
- 6 Considerações finais 160–165

Elementos pós-textuais

Posfácio 166–172

Apêndice 173–178

Anexos 178–191

Referências 192–195

Índice 196

Sumário geral

Introdução 27

I A tradução e a escrita

- 1 Transposição ortográfica e tradução 35–58
 - 1.1 Questões práticas 36
 - 1.1.1 O cotidiano 36
 - 1.1.2 O projeto 36
 - 1.1.3 Exemplos 37
 - 1.1.3.1 Uma língua, duas grafias 37
 - 1.1.3.2 Um equívoco, dois dígrafos 38
 - 1.1.3.3 A alface e o califa 38
 - 1.1.3.4 A azeitona e o militante 40
 - 1.1.4 Procurando informações 41
 - 1.1.4.1 Um contra-exemplo 41
 - 1.1.4.2 Trinta e três chefes de estado 41
 - 1.1.4.3 Um exercício de pesquisa 43

1.2	Questões teóricas	45
1.2.1	Decisões de tradução	45
1.2.2	Escopo e relevância	46
1.2.3	Formas de tradução	47
1.2.4	O leitor	48
1.2.5	As estratégias	49
1.2.6	Outras estratégias	51
1.2.7	Metáforas e palavras isoladas	52
1.2.8	Avaliação	52
1.2.9	Releitura	53
1.3	Na academia	53
1.3.1	A falsafa	53
1.3.2	Nomes próprios	55
1.3.3	Aristóteles	56
1.3.4	Diferentes referentes	57
1.4	Questões práticas e teóricas na academia	58
2	A língua árabe e sua escrita	59–77
2.1	A língua	60
2.1.1	Variantes	60
2.1.2	Morfologia	60
2.1.3	Sintaxe	64
2.2	A escrita	64
2.2.1	Designação	64
2.2.2	Origens	65
2.2.2.1	da escrita	65
2.2.2.2	das consoantes	65
2.2.2.3	das vogais	66
2.2.3	Os usuários	67
2.2.4	Línguas não-semíticas	68
2.2.5	Características básicas	70
2.2.5.1	Letras solares e letras lunares	70

2.2.5.2	Vogais longas, semivogais e suporte da hamza ^t	70
2.2.5.3	Marcas suplementares	72
2.2.6	Cursividade	74
2.2.7	Ortografia	75
2.2.7.1	e morfologia	75
2.2.7.2	e suas regras	75
2.2.7.3	e o alif	77
3	O português brasileiro e o alfabeto latino	78–93
3.1	A língua — fonemas	79
3.1.1	Consoantes	79
3.1.2	Semivogais	80
3.1.3	Vogais	81
3.2	A escrita — correspondências grafema–fonema	82
3.2.1	Nulas	82
3.2.2	Biunívocas	83
3.2.3	Unívocas	83
3.2.4	Múltiplas	85
3.2.5	Nulas, biunívocas, unívocas e múltiplas	87
3.3	Seleção de grafemas	87
3.3.1	Regras e ordem	89
3.3.1.1	Regras	89
3.3.1.2	Ordem	90
3.3.2	Etapas	90
3.3.2.1	Primeira etapa: fonemas vocálicos	90
3.3.2.2	<i>Intermezzo</i> : fonemas consonantais	90
3.3.2.3	Segunda etapa: correspondências biunívocas	91
3.3.2.4	Terceira etapa: conflito com vogal	91
3.3.2.5	Quarta etapa: conflito com consoante	92
3.3.3	Os grafemas selecionados	93

II A tradução da escrita

- 4 A romanização do árabe 96–121
 - 4.1 nos jornais 97
 - 4.1.1 Diferentes formas de romanizar 97
 - 4.1.2 Manuais de redação e estilo 98
 - 4.1.3 Manual *online* da Folha de São Paulo
— nomes árabes 99
 - 4.1.4 Comentários 100
 - 4.1.5 Análise 101
 - 4.1.6 O chefe de Estado da Líbia (outra vez) 102
 - 4.2 na academia 103
 - 4.2.1 Esquemas e sistemas de romanização 103
 - 4.2.2 O esquema de Cunha 104
 - 4.2.3 O esquema de Nasr 104
 - 4.2.4 O sistema de Machado 104
 - 4.2.5 O sistema de Nimer 109
 - 4.2.6 O sistema de Sabbagh 110
 - 4.2.7 O sistema da UNESP 111
 - 4.2.8 O sistema leigo de Jubran 113
 - 4.2.9 O sistema acadêmico de Jubran 116
 - 4.3 em análise 117
 - 4.3.1 Correspondências consensuais
— fonemas equivalentes 117
 - 4.3.2 Correspondências não-consensuais
— fonemas equivalentes 117
 - 4.3.3 Correspondências não-consensuais
— fonemas familiares 119
 - 4.3.4 Correspondências não-consensuais
— fonemas não-familiares 120
 - 4.3.5 Correspondências não-consensuais
— outros fonemas 120

- 5 Proposta de transposição ortográfica 122–159
 - 5.1 A proposta 123
 - 5.1.1 Os princípios 123
 - 5.2 A romanização acadêmica ampla 123
 - 5.2.1 As vogais 124
 - 5.2.2 As consoantes — romanização consensual de fonemas equivalentes 124
 - 5.2.3 As consoantes — romanização não-consensual de fonemas equivalentes 124
 - 5.2.4 As consoantes — romanização não-consensual de fonemas familiares 125
 - 5.2.5 As consoantes — romanização não-consensual de fonemas não-familiares 127
 - 5.2.6 As consoantes — romanização não-consensual de ⟨ξ⟩ /ɣ/ e ⟨ϵ⟩ /ʎ/ 130
 - 5.2.6.1 As não-letras 132
 - 5.2.6.2 As letras 134
 - 5.2.7 Correspondências grafema–fonema da RAA 135
 - 5.2.8 Alguns pontos pertinentes 138
 - 5.2.8.1 à ortografia 138
 - 5.2.8.2 à morfossintaxe 138
 - 5.2.8.3 ao |al-| 139
 - 5.3 A romanização acadêmica estrita 140
 - 5.3.1 Linha de base 140
 - 5.3.2 Sobrescrito 140
 - 5.3.3 Subescrito 141
 - 5.3.4 Caixa-baixa e caixa-alta 141
 - 5.3.5 Outros 141
 - 5.4 A romanização leiga comum 141
 - 5.4.1 Grafemas compostos 142
 - 5.4.2 O ponto médio 143
 - 5.5 A romanização leiga simplificada 144
 - 5.6 Resumo das propostas 146

- 5.7 Exemplos de aplicação 146
 - 5.7.1 Geminação 149
 - 5.7.2 Qalif nulo 150
 - 5.7.3 Qalif sobrescrito 151
 - 5.7.4 Hamza^t 152
 - 5.7.5 Tā^q marbūṭa^t 153
 - 5.7.6 Assimilação 154
 - 5.7.7 Artigo definido e a hamza^t al-waṣl 155
 - 5.7.8 Artigo definido e suas elisões 156
 - 5.7.9 Sufixos pronominais 157
 - 5.7.10 Alguns nomes próprios 158
 - 5.7.11 O título 159
- 6 Considerações finais 160–165
 - 6.1 Letras 161
 - 6.2 Palavras 162
 - 6.3 Nomes 163
 - 6.4 Bibliotecas 163
 - 6.5 Desdobramentos da proposta 164

Elementos pós-textuais

Posfácio 166

Apêndice

- A Terminologia 173
 - A.1 Letras e sinais gráficos 173
 - A.2 Grafemas 174
 - A.3 Combinações 175
 - A.4 Grafema–fonema 175
 - A.5 Alfabetos 176
 - A.5.1 Abjad 176
 - A.5.2 Alfabeto latino 177
 - A.6 Transposição ortográfica 178

Anexos

- A MUAMMAR GADDAFI OU MU'AMMAR AL-QADHDHAFI? 179
- B Kaddafi ou Gaddafi: como se escreve
o nome do ditador da Líbia? 181
- C Entenda as variações de fonética e grafia
do nome do ditador da Líbia 183
- D How Do You Spell 'Gaddafi'? 185
- E FROM SATURDAY NIGHT LIVE 12/12/1981 187
- F How Many Different Ways Can You Spell 'Gaddafi'? 189

Referências bibliográficas 193

Índice de citações 196

Sumário de tabelas

Capítulo 1 — Transposição ortográfica e tradução

- 1 Duas grafias para a capital do Iraque. 38
- 2 Diferentes grafias e pronúncias para a capital da Coreia do Sul. 39
- 3 Consoante <خ> / fonema [x]. 40
- 4 Assimilação da consoante final do artigo definido. 41
- 5 Possibilidades de romanização para o artigo definido e os três fonemas seguintes do sobrenome <القذافي>. 44

Capítulo 2 — A língua árabe e sua escrita

- 6 O artigo definido e o genitivo. 61
- 7 Número, caso, definição e gênero em {professor bom}. 62
- 8 Raiz triconsonantal e formas derivadas referentes a {escrever}. 63
- 9 Conjugação verbal — perfectivo (aspecto perfeito) de <كَتَبَ> |kataba| {escrever}. 63
- 10 Dois grupos de grafemas derivados. 66

- 11 Quantidade de usuários dos sistemas de escrita
mais difundidos. 68
- 12 Letras solares e lunares, com os respectivos fonemas. 71
- 13 O artigo definido e as letras solares e lunares. 71
- 14 Vogais longas. 72
- 15 Os grafemas ⟨ا⟩, ⟨ي⟩ e ⟨و⟩ nas funções de vogal longa,
semivogal e suporte da hamza^t. 72
- 16 Vogais breves. 73
- 17 Posição das ḥaraka^t. 73
- 18 Marcação de caso em ⟨بَيْت⟩ |bayt| {casa} 73
- 19 Diferentes formas cursivas, de acordo
com a ligação a grafemas adjacentes. 75
- 20 Ortografia da partícula ⟨لِ⟩ |li| {para}. 76
- 21 Ortografia das partículas ⟨وَ⟩ |wa| {e} e ⟨بِ⟩ |bi| {com}. 76

Capítulo 3 — O português brasileiro e o alfabeto latino

- 22 Alofones comuns entre /H/ e /R/. 80
- 23 Pronúncia de R em ataque de sílaba não-inicial,
precedido de L. 81
- 24 Distribuição das vogais orais
do português brasileiro [...] 82
- 25 Vogais nasais do português brasileiro. 82
- 26 Correspondências biunívocas. 84
- 27 Correspondências unívocas grafema–fonema. 84
- 28 Correspondências unívocas fonema–grafema. 85
- 29 Correspondências múltiplas — grafemas. 86
- 30 Correspondências múltiplas — fonemas. 87
- 31 Fonemas consonantais e os grafemas a representá-los. 88
- 32 Os fonemas vocálicos e os grafemas correspondentes. 90
- 33 Os fonemas consonantais e os grafemas
(simples e/ou compostos) a representá-los. 91
- 34 Seleção dos correspondências biunívocas. 92

- 35 Seleção das correspondências em conflito com vogais. 93
- 36 Seleção das correspondências em conflito
com consoantes. 94
- 37 Grafemas selecionados. 94

Capítulo 4 — A romanização do árabe

- 38 Nomes próprios árabes romanizados por três jornais. 97
- 39 O artigo definido romanizado por três jornais. 98
- 40 O esquema de romanização de CUNHA (1987). 105
- 41 O esquema de romanização de NASR (2005). 106
- 42 O sistema de romanização de MACHADO (1990). 108
- 43 O sistema de romanização de NIMER (2005). 109
- 44 O sistema de romanização de SABBAGH (1988). 111
- 45 O sistema de romanização da UNESP. 114
- 46 O sistema de romanização leiga de JUBRAN (2004). 115
- 47 O sistema de romanização acadêmica de JUBRAN (2004). 116
- 48 Romanização consensual — fonemas equivalentes. 118
- 49 Romanização não-consensual de fonemas equivalentes. 118
- 50 Romanização não-consensual de fonemas familiares. 119
- 51 Romanização não-consensual de fonemas não-familiares. 120
- 52 Romanização não-consensual de ⟨ξ⟩ /ʕ/ e ⟨ϣ⟩ /ʔ/. 121

Capítulo 5 — Proposta de transposição ortográfica

- 53 Romanização das vogais. 124
- 54 Seleção de grafemas simples primitivos
— romanização consensual de fonemas equivalentes. 125
- 55 Seleção de grafemas simples primitivos
— romanização não-consensual
de fonemas equivalentes. 125
- 56 Romanização de fonemas familiares
— grafemas simples e compostos. 126

- 57 Seleção de grafemas simples primitivos
— romanização fonemas familiares. 126
- 58 Romanização de fonemas familiares
— grafemas derivados. 127
- 59 Uso do traço inferior. 128
- 60 Romanização de fonemas não-familiares
— grafemas primitivos e derivados. 129
- 61 Romanização de fonemas não-familiares
— grafemas primitivos e derivados. 131
- 62 Seleção de grafemas simples primitivos
— romanização de ⟨Ⲫ⟩ /ʕ/ e ⟨Ⲥ⟩ /ʔ/. 131
- 63 Isolamento tipográfico decorrente
do uso de ⟨ⲀⲀ⟩ ou ⟨ⲀⲀ⟩. 133
- 64 Aspas simples e duplas na romanização do hebraico. 133
- 65 Vantagem tipográfica do uso de ⟨ⲀⲀ⟩. 135
- 66 Grafemas latinos selecionados para as respectivas
correspondências grafema–fonema. 136
- 67 O sistema de romanização acadêmica ampla. 137
- 68 Romanização de fonemas familiares
— grafemas compostos. 143
- 69 Romanização de fonemas não-familiares
— grafemas compostos. 144
- 70 Romanização simplificada de fonemas familiares. 145
- 71 Romanização simplificada de fonemas não-familiares. 145
- 72 Ambiguidades geradas pelo uso de um mesmo grafema
para diferentes fonemas. 146
- 73 Letras solares e lunares (consoantes e semivogais). 147
- 74 Grafemas do abjad árabe
— nome, fonema e romanizações. 148
- 75 Exemplos de romanização — consoantes geminadas. 149
- 76 Exemplos de romanização
— correspondência nula de ⟨ⲀⲀ⟩. 150
- 77 Exemplos de romanização — maddat e ʿalif ḥanjariyyat. 151

- 78 Exemplos de romanização — suporte da hamza^t. 152
- 79 Exemplos de romanização — tā^q marbūṭa^t. 153
- 80 Exemplos de romanização — assimilação em ⟨ال⟩ |al-|. 154
- 81 Exemplos de romanização — hamza^t al-waṣl. 155
- 82 Exemplos de romanização — elisões em ⟨ال⟩ |al-|. 156
- 83 Exemplos de romanização — sufixos pronominais. 157
- 84 Exemplos de romanização — nomes próprios. 158
- 85 Exemplos de romanização — o título (RAA). 159
- 86 Exemplos de romanização — o título (RAE). 159

Sumário de figuras

Capítulo 5 — Proposta de transposição ortográfica

- 1 O registro das consoantes e das vogais breves;
e o registro, na RAE, da consoante geminada. 149
- 2 O ʔalif nulo — valor meramente ortográfico. 150
- 3 Madda^t e ʔalif sobrescrito. 151
- 4 A hamza^t sem suporte, apoiada no ʔalif,
no yā^q e no wāw. 152
- 5 Registro da tā^q marbūṭa^t *versus* hā^q final 153
- 6 Registro da assimilação de /l/, na RAA e na RAE. 154
- 7 Hamza^t waṣl. 155
- 8 Supressão ortográfica de grafemas do artigo definido.. 156
- 9 Sufixos pronominais no acusativo e no genitivo. 157
- 10 O nome da capital do Iêmen. 158
- 11 O nome do chefe de Estado da Líbia. 158

Anexo E — FROM SATURDAY NIGHT LIVE 12/12/1981

- 12 Programa *Saturday Night Live* de 1981-12-12. 188

Lista de sinais de pontuação

“ ... ”	aspas curvas	uso metafórico
«...»	aspas angulares ¹	citação
(...)	parênteses	breve acréscimo de informação
[...]	colchetes	transcrição fonética; legenda de abreviaturas
/.../	barras oblíquas	transcrição fonológica
...	barras verticais	transposição ortográfica
⟨...⟩	colchetes angulares	referência a palavra ² , letra ou símbolo
{...}	chaves	tradução ³
—...—	travessões	inserção de comentário ⁴

¹Também chamadas de *guillemets*.

²Função normalmente reservada às aspas simples (‘...’). Os colchetes angulares são também usados para delimitar palavras árabes.

³Função normalmente reservada aos colchetes, especialmente em referências bibliográficas, para apresentar traduções livres dos títulos.

⁴Em final de período, usamos apenas o travessão de abertura, seguido de espaço em branco.

Lista de variantes tipográficas

com serifa	texto normal	
sem serifa	títulos, notas de margem e citações em bloco	
romana⁵	texto normal	
<i>oblíqua</i>	<i>foreign words</i> {palavras estrangeiras}	
<i>itálica</i>	ênfase	
<u>sublinhada</u>	palavra ou expressão a ser <u>comentada</u>	comentada em nota de margem.
VERSALETE	sobrenomes, siglas, abreviaturas e letras	

⁵Também chamada de regular.

Introdução

Ex litteris scientia.

«Das letras, o conhecimento.»

do autor

Nosso interesse por sistemas de escrita acompanha-nos desde longa data. Dos primeiros *kana* dos silabários japoneses —aprendidos ainda na infância— às configurações de mão do *sign writing* (um dos meios usados para grafar línguas de sinais) —apreendidos já na graduação—, nunca deixou de nos fascinar a variedade de soluções encontradas pela engenhosidade humana para efetuar o registro escrito das mais variadas línguas.

Para quem queira ampliar seus conhecimentos nessa área, o curso de Letras proporciona uma oportunidade ímpar. Aqui, o aluno de graduação tem à sua disposição variadas oportunidades para aprender ou apenas ter contato com diferentes sistemas de escrita. As mais notáveis encontram-se no DLO —vale ressaltar que esse departamento congrega habilitações exclusivamente de línguas que empregam sistemas

DLO: Departamento de Letras Orientais

de escrita diferentes do alfabeto latino—, nas disciplinas das cadeiras de árabe, armênio, chinês, japonês, hebraico, russo e, mais recentemente, coreano; outras, nas disciplinas de sânscrito e grego antigo (DLCV), grego moderno —língua neo-helênica— (DLM) e línguas de sinais (DL), além da disciplina Sistemas de Escrita (DL).

DLCV: Departamento de Letras Clássicas e Vernáculas

DLM: Departamento de Letras Modernas

DL: Departamento de Linguística

Tendo cursado optativas abrangendo a maioria das línguas mencionadas, percebemos que pouco ou nada se pesquisava acerca de sistemas de escrita. Sendo os estudos de Paleografia (ciência afim à Filologia) o mais próximo desse tipo de estudo, não deixa de ser curioso notar que pesquisas tratando literalmente de letras sejam tão restritas na Faculdade de Letras. Isso se torna ainda mais evidente se considerarmos que, com a diversidade de habilitações de que dispomos e da variedade do material disponível em nossa biblioteca, em relação às línguas que se grafam com outros sistemas de escrita, não são poucas as pesquisas publicadas nem poucos os livros catalogados (e lidos) no alfabeto latino. Com esta pesquisa, procuramos começar a preencher essa lacuna.

Curiosamente, ao explicarmos nosso objeto de estudo (a romanização do árabe), não raro as pessoas perguntavam-nos por que não pesquisar a romanização do japonês — provavelmente sob a premissa de ser natural que um descendente estude a língua. Nossa resposta sempre tocou em dois pontos principais: já existem (duas ou três) formas oficiais de romanização; as questões que se podem levantar existem em quantidade reduzida. No caso do árabe, em contraste, não há uma romanização oficial, que seja internacionalmente aceita e utilizada; e as diferenças linguísticas e ortográficas em relação ao português são bastante mais diversificadas.

Isso pode ser percebido facilmente ao abrir um jornal ou uma revista e constatar que palavras japonesas raramente assumem grafias discordantes entre um veículo e outro. Por sua vez, à exceção dos que circulam há mais tempo, topônimos e antropônimos árabes são regularmente grafados de maneiras diferentes neste ou naquele jornal, nesta ou naquela revista; e, no caso de rádio e televisão, pronunciados de maneiras as mais variadas, conforme o grau de conhecimento —ou da adesão

a uma ou outra grafia— por parte do apresentador, entrevistador ou entrevistado. Para os que conhecem —mesmo que minimamente— a língua de origem, as grafias adotadas no português brasileiro ou não registram ou registram erroneamente certos elementos importantes (distintivos) dos termos e dos nomes árabes. Do desafio de abordar um tema tão cotidiano, mas tão pouco estudado, surgiu a presente pesquisa.

Em seu início, nossa pesquisa tinha por objetivo investigar como a mídia impressa (jornais e revistas) procede à romanização de termos, expressões e nomes próprios de origem árabe que não estejam dicionarizados —que não sejam nomes de países e suas capitais, por exemplo— e elaborar uma proposta que pudesse padronizar tal procedimento. No entanto, em pouco tempo percebemos que essa empreitada seria muito pouco frutífera, uma vez que demandaria haver, nas redações, uma pessoa que conhecesse a língua árabe — além de depender do interesse dos meios de comunicação, que eventualmente deixam de seguir as normas de seus próprios manuais de estilo.

Surgiu daí a possibilidade de, mantendo a intenção inicial (padronizar a romanização), elaborar uma proposta para o meio acadêmico. Diversos fatores contribuíram para essa mudança, acenando com um prognóstico favorável. Primeiramente, apresentaríamos um trabalho destinado a quem tem contato mais direto com o árabe — linguistas, professores de língua e cultura, pesquisadores das mais variadas áreas (e.g, literatura, história, geopolítica, relações internacionais), entre outros tantos. Além disso, passaríamos a dialogar com um público que faz de seu cotidiano a atenção a pormenores e o cuidado com minúcias. O *modus operandi* da imprensa passou, então, a um segundo plano, restringindo-se uma referência ao que encontramos no dia a dia e ao que se tornaria o cerne de nossa pesquisa.

A etapa seguinte seria estudar o que já se havia feito no tocante à romanização do árabe. Para recortar melhor nosso objeto de pesquisa, concentramo-nos em esquemas voltados ao português brasileiro; compilamos o trabalho de seis autores e comparamo-los, a fim de identificar

o que era relevante (ou não) registrar na passagem de um a outro sistema de escrita.

Os esquemas apresentaram similaridades —observamos concordância em cerca de metade das correspondências entre o grafema do sistema de escrita original e o utilizado na romanização—, mas algumas diferenças. As mais aparentes referiram-se à maneira como se representam os fonemas que não existem no português, mas são familiares —por serem encontrados em línguas como o inglês e o espanhol—, e os que praticamente só ocorrem em línguas semíticas, em especial os fonemas faringalizados (enfáticos).

Basicamente, podem ser adotados dígrafos e/ou diacríticos, cada qual possuindo vantagens e desvantagens. O uso de dígrafos facilita a digitação em teclados, mas pode gerar ambiguidades, por serem comuns os encontros consonantais em árabe — por exemplo, ⟨fatha⟩ deve ser lido [fa.θa] ou [fat.ha]? Os diacríticos evitam esse problema, mas causam outros, como a falta de consistência em seu emprego — por exemplo, se o ponto indica faringalização em ⟨ṭ⟩ (oclusiva alveolar desvozeada faringalizada), então ⟨ġ⟩ deveria representar uma oclusiva velar vozeada faringalizada, não uma fricativa velar vozeada não-faringalizada.

Escolher qual procedimento adotar não foi difícil. Uma vez que visávamos a atender um público especialista —que, podíamos supor, estaria interessado menos em simplificações e facilidades do que em ver grafadas em caracteres latinos distinções relevantes no árabe— escolhemos os diacríticos para indicar as correspondências grafema–fonema que escapavam ao uso de letras simples (sem diacríticos).

No entanto, tão logo começamos a experimentar com essa forma preliminar de romanização, notamos que ela poderia servir de base para outros esquemas. A partir do que viria a se tornar a romanização acadêmica ampla (RAA), derivamos um esquema de transliteração, a romanização acadêmica estrita (RAE), que permite registrar elementos e fenômenos outrora presentes apenas na escrita original. E, uma vez estabelecida a regularidade de representação dos diacríticos, não foi difícil aplicar esse mesmo princípio ao uso de dígrafos, no caso da romanização

leiga comum (RLC), e à remoção completa de marcas distintivas (diacríticos ou dígrafos), no caso da romanização leiga simplificada (RLS). A proposta havia se desenvolvido e contava agora com quatro esquemas distintos de romanização, cada um servindo a um propósito.

A RAA permite registrar distinções normalmente ignoradas nos esquemas de romanização estudados — por exemplo, *hā^q* final *versus* *tā^q marbūṭa^t* e *nūn* final *vs.* *tanwīn* (nunação). Ela foi projetada para oferecer um considerável grau de correspondência com a escrita original, permitindo que seja usada, por exemplo, em dicionários etimológicos e em trabalhos acadêmicos, bem como na catalogação de obras, em bibliotecas.

A RAE pode ser considerada uma forma de transliteração. Ela oferece um grau ainda mais elevado de correspondência, permitindo identificar elementos ortográficos até agora só registrados na escrita árabe. Por exemplo, pode-se explicitar qual letra (árabe) é usada como suporte da hamza.

A RLC é voltada para uso em contextos de menor rigor na produção da romanização. Em vez de diacríticos, empregam-se dígrafos, mas mantendo os mesmos princípios adotados na atribuição de diacríticos. De certo modo, RAA e RLC são variantes de um mesmo conceito, a única diferença sendo o uso de diacríticos ou dígrafos, respectivamente. A RAA permite estabelecer correspondências mais diretas entre a escrita árabe e sua romanização; a RLC apresenta-se como uma forma mais fácil e prática de ser produzida, uma vez que emprega caracteres prontamente disponíveis em qualquer teclado usado na digitação do português.

A RLS, como o próprio nome sugere, é uma versão simplificada da RLC. Não se usam diacríticos nem dígrafos; e neutralizam-se as diferenças entre grafemas de mesma letra básica. Por exemplo, <Ṭ>/<ṬḤ> e <Ṭ>/<ṬḪ> viram ambos <Ṭ>, i-e, diferentes grafemas (ou fonemas) árabes são representados por um mesmo grafema na romanização.

Resumindo, em nossa proposta as romanizações acadêmicas empregam diacríticos; as leigas, dígrafos. Basicamente, os usuários acadêmicos podem se valer da precisão e da acurácia de representação dos dia-

críticos, enquanto os leigos podem aproveitar a facilidade de produção (digitação) dos dígrafos. Em relação às desvantagens de um e de outro método, mencionadas anteriormente, atribuímos a cada diacrítico um valor único, tornando fácil identificar a relação dos grafemas em questão com suas contrapartes simples (sem diacrítico) — na prática, basta conhecer o valor do diacrítico para deduzir, no mesmo instante, a que fonema o grafema com diacrítico se refere; e apresentamos soluções bastante simples para contornar as eventuais ambiguidades geradas pelos dígrafos.

A fim de melhor abordar a questão da transposição ortográfica, discutir as bases de nossa proposta e apresentar nosso sistema, dividimos esta dissertação em duas partes, cada qual contendo três capítulos, a saber:

No **primeiro capítulo**, estudamos as relações entre *transposição ortográfica e tradução*. Procuramos apresentar as questões práticas e teóricas do estudo da transposição de palavras entre línguas que usem sistemas de escrita diversos. Abordamos diferentes casos de transposição ortográfica, verificamos como as escolhas exigidas por esse processo não muito diferem das envolvidas na tradução propriamente dita e salientamos os elementos considerados relevantes para adotarmos a presente linha de estudo.

No **segundo capítulo**, descrevemos *a língua árabe e sua escrita*. Informamos as características do árabe e a relação da língua com sua escrita. Esses pontos serão relevantes ao tratarmos das escolhas pertinentes às correspondências grafema–fonema de nossos esquemas de romanização e às formas de representação de fenômenos fonéticos e de elementos morfológicos.

No **terceiro capítulo**, falamos *d’o português brasileiro e o alfabeto latino*. Esmiuçamos de que maneiras um fonema pode ser grafado e que fonemas representa um dado grafema. E procuramos selecionar correspondências grafema–fonema que não gerem ambiguidades de produção e/ou interpretação. Uma tal biunivocidade será útil na elaboração

do sistema de romanização, que se utilizará das correspondências aqui selecionadas.

No **quarto capítulo**, analisamos três esquemas e cinco sistemas voltados para *a romanização do árabe*. Verificamos, primeiro, como a imprensa escrita brasileira aborda a questão da grafia de nomes estrangeiros — em especial os de origem árabe. Em seguida, analisamos como dicionários etimológicos ou bilíngues e uma proposta específica para o meio acadêmico procedem à transposição ortográfica do árabe padrão para o português brasileiro.

No **quinto capítulo**, apresentamos nossa *proposta de transposição ortográfica*. Estabelecemos os princípios para a elaboração do sistema de transposição ortográfica e apresentamos seus quatro módulos, a saber: romanização leiga simplificada; romanização leiga comum; romanização acadêmica ampla; romanização acadêmica estrita. Apresentamos também alguns exemplos de uso desses módulos em palavras e expressões árabes, além de um breve texto.

No **sexto capítulo**, discutimos os *desdobramentos da proposta*. Consideramos algumas das atuais limitações e das futuras possibilidades de desenvolvimento referentes ao sistema proposto. Das propostas verificamos o que precisa ser consolidado, especialmente na representação de determinados fenômenos. Dos testes e exemplos em que pudemos aplicar as romanizações —particularmente a RAA e a RAE—, levantamos o que pudemos perceber acerca de algumas questões, para as quais procuramos apontar soluções e possíveis alternativas. Do restante analisamos o que falta abordar, em termos de contextos de aplicação, e para quais áreas mais podemos expandir, em relação ao escopo de aplicação.

Esperamos, com este *sistema de transposição ortográfica do árabe padrão para o português brasileiro*, poder contribuir para que o registro em caracteres latinos de termos, expressões e nomes árabes possa ser tão preciso e acurado quanto possível.

Parte I

Da tradução e da escrita

CAPÍTULO 1

Transposição ortográfica e tradução

Dubium sapientiae initium.

«Na dúvida começa a sabedoria.»

René Descartes

PROCURAMOS APRESENTAR as questões práticas e teóricas do estudo da transposição de palavras entre línguas que usem sistemas de escrita diversos. Abordamos diferentes casos de transposição ortográfica, verificamos como as escolhas exigidas por esse processo não muito diferem das envolvidas na tradução propriamente dita e salientamos os elementos considerados relevantes para adotarmos a presente linha de estudo.

1.1 Questões práticas

A transposição de nomes estrangeiros está presente em nosso dia a dia. Ao ler um jornal, assistir ao noticiário ou dizer um nome, estamos lidando com decisões que precisamos tomar ou que foram por outrem tomadas.

1.1.1 O cotidiano

A língua árabe há muito faz parte do cotidiano brasileiro. Sua presença estende-se desde o legado mouro na península ibérica até as imigrações de fins do século XIX e início do século XX, que trouxeram ao Brasil música, dança e culinária dos países de origem dos imigrantes.

As palavras de origem árabe constantes no léxico em uso no Brasil pouco ou nenhum estranhamento causam ao brasileiro; o mesmo pode-se dizer da maioria dos topônimos e antropônimos. No entanto, nos últimos anos tem crescido a quantidade de referências a lugares e pessoas antes pouco ou nunca mencionados. Acompanham esse fenômeno a novidade de grafias que fogem à ortografia do português brasileiro, bem como uma certa inconsistência ou falta de uniformidade no modo como esses nomes são grafados. Isso faz com que, não raro, as pessoas perguntem-se por que o nome daquela cidade ou pessoa é escrito de uma maneira em um jornal e de outra em uma revista; e que tomem por incorreta a pronúncia do âncora de um telejornal ou a de um entrevistado. Entre o certo e o incerto, fica a pergunta sobre por que existem diferentes grafias e se seria possível estabelecer-lhes uma uniformização.

topônimos:
nomes
de países,
cidades e
acidentes
geográficos.
antropônimos:
nomes
de pessoas.

1.1.2 O projeto

A proposta inicial deste estudo consistia na análise do modo como a mídia escrita brasileira grava topônimos e antropônimos originalmente escritos em árabe. Quando lemos notícias que envolvem nomes oriundos desse sistema de escrita diverso do utilizado em português, um rá-

pido cotejo entre diferentes meios de comunicação —jornais e revistas semanais, principalmente— permite constatar que a grafia apresenta variações de um veículo a outro.

Por exemplo, o nome do Chefe de Estado da Líbia aparece como ⟨Muamar Kadafi⟩ n'O Estado de São Paulo e ⟨Muammar Gaddafi⟩ na Folha de São Paulo. Isso não se observa em nomes já consolidados em português brasileiro, como no caso de nomes de países e suas capitais. No entanto, entre português brasileiro (PB) e português europeu (PE) algumas diferenças podem ser observadas (*vide* 1.1.3.1, abaixo).

E há situações em que a adaptação de topônimos e antropônimos não encontra consenso em relação a nomes grafados no alfabeto latino e conhecidos de longa data. Por exemplo, encontra-se o nome da cidade estadunidense de New York grafado de pelo menos três formas: sem alterações, como no original, ⟨New York⟩, i.e, usa-se o endônimo; com todos os elementos ortográficos adaptados, ⟨Nova Iorque⟩, i.e, utiliza-se um exônimo; ou mesmo ⟨Nova York⟩, em uma forma a mesclar tradução ((⟨New⟩ → ⟨Nova⟩) e manutenção ((⟨York⟩ = ⟨York⟩) de seus elementos constituintes.

endônimo:
nome usado
pelos falantes
locais.

exônimo:
nome usado
por estrangeiros.

1.1.3 Exemplos

Observemos alguns exemplos de como podem diferir alguns exônimos dos respectivos endônimos; e como estão distantes certas formas mais recentes das já consolidadas na língua.

1.1.3.1 Uma língua, duas grafias

Para escrever o nome da capital do Iraque, ⟨بغداد⟩ /bayda:d/, no Brasil grafa-se ⟨Bagdá⟩; em Portugal, ⟨Bagdade⟩ (tabela 1, página 38). Recorrendo à grafia original, em árabe, percebemos haver dois elementos relevantes: Uma vogal longa /a:/, que carrega o acento da palavra; uma consoante final /d/. Podemos supor que a última sílaba /da:d/ percebe-se de maneiras sutilmente diferentes em Portugal e no Brasil. Em ou-

tras palavras, os brasileiros parecem aproximar /da:d/ de palavras como <dá>; os portugueses, de <idade>.

Tabela 1: Duas grafias para a capital do Iraque.

árabe	tr·fonol	português
بغداد	baɣda:d	Bagdá (PB) Bagdade (PE)

[Tr]anscrição [fonol]ógica.

1.1.3.2 Um equívoco, dois dígrafos

Por vezes, ocorre primazia da escrita sobre a fala. Quem assistiu aos Jogos Olímpicos de 1988 deve lembrar-se de ter visto o nome da capital da Coreia do Sul grafado, em inglês, desta maneira: <Seoul>, pronunciado /soʊl/. Em português, por outro lado, grafamos desta: <Seul>, lido /seul/. Ao compararmos com o original, <서울> /sɐ.ul/, notamos uma divergência, em português, em relação à pronúncia da primeira vogal /ɐ/ — <ㅊ>, na escrita original —, que em português estaria mais próxima de /ɔ/¹ do que de /e/. Cabe mencionar que a relação entre ortografia e pronúncia, em português, aproxima-se — ou deriva — da grafia adotada em francês, <Séoul> /se.ul/, que denota ter sido identificado <OU> como dígrafo; em inglês, por outro lado, o dígrafo está em <EO> (tabela 2).

1.1.3.3 A alface e o califa

No étimo de <alface> há dois elementos morfológicos: o artigo definido <al->; o nome propriamente dito, <-face>. MACHADO (1990) remete-nos ao árabe <al-khass>, que em uma leitura ingênua pronunciar-se-ia /al.kas/, ignorando-se o ھ — algo similar poderia acontecer com <Århus>²

¹Entre os fonemas /ɐ/ e /ɔ/ existe apenas diferença de arredondamento.

²Nome da segunda maior cidade da Dinamarca; também grafado <Aarhus>, não implicando mudança de pronúncia ou interferência na alfabetação.

Tabela 2: Diferentes grafias e pronúncias para a capital da Coreia do Sul.

idioma	grafia	leitura
coreano	서울	sɐ.ul
inglês	<u>S</u> eoul	souɫ
francês	S <u>e</u> oul	se.ul
português	Seul	se.ul

O sublinhado indica o dígrafo.

ou ⟨Dvořák⟩³, cujos diacríticos seriam desconsiderados. Essa leitura levantaria uma questão acerca da relação entre a suposta plosiva velar /k/ do étimo e a fricativa labiodental [f] da palavra corrente.

Ao consultar a grafia em árabe ⟨الخص⟩, encontramos a consoante ⟨خ⟩, que corresponde a uma fricativa velar [x]. Essa consoante aparece em outra palavra: califa⁴. Consultando a grafia em árabe ⟨خليفة⟩, encontramos a primeira consoante ⟨خ⟩ —mesma de ⟨-face⟩—, desta vez associada ao ⟨c⟩. Temos, então, uma mesma consoante —e um mesmo fonema— árabe consolidada diferentemente em duas palavras incorporadas ao português: alface; califa.

Esta mesma palavra, em situações diversas, pode aparecer grafada de outras formas. Um exemplo seria o prédio inicialmente chamado de ⟨Burj Dubai⟩, rebatizado com o novo nome de ⟨Burj Khalifa⟩ ⟨برج خليفة⟩.

Outro exemplo pode ser encontrado no jogo *online* Guerra Khan⁵, em que existe a possibilidade d'o personagem do jogador obter o título de *Halif*. A propósito, e ainda a respeito da mesma consoante, esse jogo oferece outros títulos⁶, entre os quais encontra-se o de *Sheih*, que em

³Sobrenome tcheco; normalmente refere-se ao compositor Antonín Dvořák.

⁴Ao contrário do que ocorre em muitas das palavras portuguesas de origem árabe, em ⟨califa⟩ não se encontra o artigo definido ⟨al-⟩.

⁵<http://www.guerrakhan.com>

⁶ Os oito títulos, em ordem decrescente de importância, são: Halif; Magrib halif;

português encontra-se dicionarizado como ⟨xeque⟩, do árabe ⟨شيخ⟩.

Encontramos, para uma mesma consoante ⟨خ⟩ —ou fonema /x/—, registros nas mais variadas grafias (tabela 3): c; f; h; KH; QU. Em todas as instâncias, trata-se de uma consoante desvozeada.

Tabela 3: Consoante ⟨خ⟩ / fonema [x].

árabe	tr·fonol	português	outros
خس	xas	(al)face	(al-)khass
خليفة	xali:fa	califa	khalifa, halif
شيخ	ʃajx	xeque	sheih

[Tr]anscrição [fonol]ógica.

O sublinhado indica grafemas e dígrafos correspondentes a /x/.

1.1.3.4 A azeitona e o militante

Além da transposição de consoantes ou fonemas, outros elementos são transpostos de maneiras variadas. O registro —ou não— de alguns fenômenos fonéticos pode ser observado ao compararmos, por exemplo, ⟨azeitona⟩ com ⟨Al-Zarqawi⟩⁷. O étimo de ⟨azeitona⟩ compõe-se, à semelhança do de ⟨alface⟩ (1.1.3.3, p. 38), de dois elementos, quais sejam: o artigo definido ⟨al-⟩; o nome ⟨zeitona⟩, de ⟨زيتون⟩ |zaytūn|. Note-se que o l de ⟨al⟩ não consta na grafia em português; isso aproxima da pronúncia original a do português. No caso de ⟨Al-Zarqawi⟩, a pronúncia a essa grafia associada desconsidera a assimilação (tabela 4).

No árabe padrão, grafa-se o nome precedido pelo artigo por inteiro ⟨الزيتون⟩, mas na pronúncia opera-se a assimilação regressiva do fonema alveolar /l/. Tal assimilação ocorre, no árabe padrão, na combinação do

King; Vesire; Sheih; Sanjack bei; Ulemi; Beduin. Lista recuperada em 2011-07-05, disponível em <http://help.xs-software.com/game/browser/khanwars/language/br/topic/41/main//game.html>

⁷Sobrenome de ⟨أبو مصعب الزرقاوي⟩ |ʔabū Muṣʿab Az-Zarqāwī|, membro da القاعدة |Al-Ḳaʿida| morto em 2006-06-07.

artigo ⟨al⟩ com palavras iniciadas por fonemas coronais, exceto /z/; na gramática tradicional árabe, esses fonemas são chamados de ⟨solares⟩, em oposição aos ⟨lunares⟩, que não se assimilam (*vide* 2.2.5.1, p. 70).

Tabela 4: Assimilação da consoante final do artigo definido.

árabe	assimilação	grafia	leitura
الزيتون	[al.zaj.tu:n] → [az.zaj.tu:n]	azeitona	/a.zej.to.na/
الزرقاوي	[al.zar.qa.wi:] → [az.zar.qa.wi:]	Al-Zarqawi	/al.zar.ka.wi/

1.1.4 Procurando informações

Voltando ao ponto inicial, vejamos algumas das implicações de haver diferentes formas para grafar um mesmo nome.

1.1.4.1 Um contra-exemplo

Quem quiser procurar por informações sobre o 40.^o presidente estadunidense não precisará mais do que pesquisar por ⟨Ronald Reagan⟩. Talvez tenha de filtrar resultados que remetam à pregressa carreira de ator ou ao porta-aviões nuclear CVN-76 (batizado de «Ronald Reagan»), mas não encontrará problemas pertinentes à grafia do nome. Além disso, terá à disposição resultados em diferentes línguas.

1.1.4.2 Trinta e três chefes de estado

O mesmo exercício, aplicado a um nome que se pode escrever de variadas maneiras, pode se mostrar uma tarefa bastante trabalhosa. No caso do atual Chefe de Estado da Líbia, por exemplo, os veículos de comunicação —de uma mesma língua e até de um mesmo país— nem sempre adotam a mesma grafia; tal é o caso, mencionado em 1.1.2 (p. 37), d'O Estado de São Paulo (⟨Muamar Kadafi⟩) e da Folha de São Paulo (⟨Muammar Gaddafi⟩). Além disso, pesquisar em outras línguas aumenta

ainda mais as variantes a considerar; para citar um serviço bastante conhecido e utilizado —ainda que não seja uma fonte oficial—, na Wikipedia⁸ encontramos grafias diferentes, dependendo da língua em que se encontrar o texto. Apenas para citar alguns exemplos, temos: Muammar al-Gaddafi (português); Muamar el Gadafi (espanhol); Mu'ammar Ghaddafi (italiano); Mouammar Kadhafi (francês); Muammar Gaddafi (inglês)⁹.

Na mesma linha, um *website*¹⁰ elenca 32 grafias distintas para esse mesmo nome, quais sejam:

1. Muammar Qaddafi	17. Moamar Gaddafi
2. Mo'ammar Gadhafi	18. Mu'ammar Qadhdhafi
3. Muammar Kaddafi	19. Muammar Khaddafi
4. Muammar Qadhafi	20. Muammar al-Khaddafi
5. Moammar El Kadhafi	21. Mu'amar al-Kadafi
6. Muammar Gadafi	22. Muammar Ghaddafy
7. Mu'ammar al-Qadafi	23. Muammar Ghadafi
8. Moamer El Kazzafi	24. Muammar Ghaddafi
9. Moamar al-Gaddafi	25. Muamar Kaddafi
10. Mu'ammar Al Qathafi	26. Muammar Quathafi
11. Muammar Al Qathafi	27. Mohammed Q'udafi
12. Mo'ammar el-Gadhafi	28. Muammar Ghaddafi
13. Moamar El Kadhafi	29. Muamar Al-Kaddafi
14. Muammar al-Qadhafi	30. Moammar Khadafy
15. Mu'ammar al-Qadhdhafi	31. Moammar Qudhafi
16. Mu'ammar Qadafi	32. Mu'ammar al-Qaddafi

⁸<http://www.wikipedia.com>
Último acesso em 2011-07-05.

⁹Desde nossa primeira consulta, foram alteradas as grafias em espanhol e inglês (anteriormente ⟨Muammar al-Gaddafi⟩) e em italiano (anteriormente ⟨Muammar Ghaddafi⟩)

¹⁰<http://www.geocities.com/Athens/8744/spelling.htm>
Último acesso em 2011-07-05.

E vale ressaltar que o próprio *website* adota uma 33.^a variante, ⟨Mu’ammār al-Qadhafī⟩, pois «ela é considerada a melhor transliteração, a partir do árabe, do primeiro e último nomes do líder líbio»¹¹.

Pode-se notar, por essa lista, que as diferenças são de variadas ordens. Apenas em relação ao artigo definido (no sobrenome), tem-se: caixa-alta ou caixa-baixa; emprego ou não do hífen; e mesmo registro ou não do artigo. Na grafia das consoantes, encontram-se: registro ou não da consoante ⟨خ⟩ /x/; registro ou não da geminação (consoantes dobradas); variadas grafias para as consoantes ⟨ق⟩ /q/ e ⟨ذ⟩ /ð/. Quanto às vogais, observam-se alternâncias entre: ⟨u⟩ e ⟨o⟩; ⟨a⟩ e ⟨e⟩, em posição átona; ⟨A⟩, ⟨U⟩ e ⟨E⟩, posposto ao ⟨ق⟩ /q/; ⟨I⟩ e ⟨Y⟩.

Addendum: Durante o processo de arremate desta dissertação, ao conferirmos as fontes referentes ao tema desta subsubseção, localizamos algumas menções a duas outras listas: uma de 72 variantes, que teria sido elaborada pela Biblioteca do Congresso (*Library of Congress*, EUA); uma de 112 variantes, compilada pela rede ABC NEWS (EUA). Esse material pode ser encontrado nos anexos A–F.

1.1.4.3 Um exercício de pesquisa

Apenas para ilustrar alguns dos obstáculos impostos por tamanha variedade de formas de romanizar um mesmo nome, imaginemos como seria procurar por um hipotético livro escrito por ele, ⟨معمّر القذافي⟩.

Se a procura for feita com base no sobrenome —e tendo em mente as variantes apresentadas em 1.1.4.2 (p. 42)—, a consulta chega a um impasse antes mesmo da primeira letra.

Há variadas possibilidades referentes ao artigo definido ⟨al⟩ (tabela 5). Ele pode: não ser registrado; ser grafado com ⟨A⟩ ou ⟨E⟩; estar em caixa-

¹¹No original: «it is considered to be the best transliteration from the Arabic of the Libyan Leader’s first and last names.»

-alta ou caixa-baixa; ser ligado por hífen —ou não— ao nome que determina. Além disso, a alfabetação pode ser feita desconsiderando-se o artigo —como se faz com ⟨de⟩, ⟨da⟩, ⟨dos⟩, ⟨von⟩, ⟨van⟩ e afins— ou considerando-o parte integrante do nome — como o ⟨Mac⟩/⟨Mc⟩ escocês (e.g, MacLeod) e o ⟨O'⟩ irlandês (e.g, O'Brien).

E não são menos numerosas as possibilidades que se seguem (tab. 5). Quem conhecer o fonema /q/ poderá começar pela letra Q; quem for leitor d'O Estado de São Paulo, pela letra letra κ; quem for leitor da Folha de São Paulo, pela letra G; e, podemos supor, quem apenas tiver ouvido o nome, mas não tiver tido contato com uma das formas escritas, pela letra c. Somente na primeira letra temos quatro opções: Q; κ; G; c.

A seguir, novas escolhas devem ser feitas. A segunda letra, seguinte ao Q, pode ser A ou U; ou UA; ou um apóstrofo. A seguinte ao κ pode ser A ou então H — na verdade, HA. Ao G, ou A ou H; se H, depois podem vir A ou E. Na sequência, a consoante após a vogal pode ser grafada D, DD, DH ou DHDH; ou TH; ou ZZ.

Tabela 5: Possibilidades de romanização para o artigo definido e os três fonemas seguintes do sobrenome ⟨الْقَذَافِي⟩.

grafema árabe	ال	ق	ـ	ذ
fonema	/al/	/q/	/a/	/ð:/
romanizações	Ø	Q	A	D
	al-	K	E	DD
	Al-	KH	U	DH
	Al	G	'	DHDH
	el-	C	—	TH
	El	—	—	ZZ

1.2 Questões teóricas

Vejamos algumas questões acerca da teoria subjacente à prática da tradução.

1.2.1 Decisões de tradução

Via de regra, o trabalho do tradutor consiste em transportar um texto de uma língua para outra, transmitir significados da língua de partida para a de chegada. E, a fim de traduzir um conceito mais específico para leitores que pouco o conheçam, os tradutores podem recorrer a variados procedimentos.

Para citar apenas um exemplo, uma palavra de referente tipicamente brasileiro como <sertão> não encontra equivalente direto em inglês. Em um texto traduzido para esse idioma, pode-se esperar encontrar diferentes procedimentos, que incluem: empregar sempre uma mesma palavra (e.g, *wasteland*); alternar diferentes palavras de um mesmo campo semântico (e.g, *wasteland*, *hinterland*, *badlands*), no intuito de que o leitor apreenda o todo a partir de alguns elementos comuns aos equivalentes escolhidos; adotar um equivalente conhecido na anglofonia (e.g, *outback*); manter o original, adaptado ou não (e.g, *sertao* ou *sertão*) e talvez incluir uma nota explicativa.

O conjunto de procedimentos empregados nessa tarefa varia de acordo com o material sendo trabalhado, os propósitos da produção e o tipo de leitor, atentando para a interpretação, a transmissão de significados e a passagem de conceitos de uma língua a outra. Porém, ao tratarmos de assuntos que envolvam elementos estrangeiros, não raro nos deparamos com palavras que ou não se grafam como usualmente o fazemos ou sequer se escrevem com os caracteres que utilizamos.

Em nosso trabalho, consideramos estar a especificidade da tradução presente também na transposição ortográfica. Se o trabalho com nomes próprios e sua grafia nem sempre pode ser considerado tradução *stricto sensu*, muitas das ferramentas usadas em textos podem (e são) aplicadas

à transposição de nomes. Desse modo, passar um nome próprio de uma língua para outra pode ser considerado tradução, especialmente se ele estiver originalmente grafado em sistema de escrita diverso.

Como embasamento teórico dessa afirmação, em CHESTERMAN & WAGNER (2002) encontramos os diversos aspectos pertinentes à prática e à teoria da tradução. Faremos uma releitura sob o viés da transposição ortográfica, salientando as semelhanças de procedimento e solução existentes entre a tradução tradicional¹² e a de nomes próprios.

1.2.2 Escopo e relevância

As decisões que um tradutor precisa tomar (ao lidar com um nome próprio a ser grafado) dependem de saber para qual de dois públicos distintos —leigo ou especialista— o texto é dirigido. Dois conceitos que orientam essas decisões são a teoria de escopo e a de relevância.

Andrew Chesterman resume a teoria de escopo dizendo que «a forma segue a função»¹³ e a de relevância em um simples «seja pertinente ao propósito»¹⁴. Isso de certa forma contrasta com a ideia de que «o dever do tradutor é traduzir com exatidão [...] e na totalidade»¹⁵, uma vez que o propósito contempla justamente essa necessidade de traduzir “na totalidade”.

Nos meios de comunicação de massa, o «efeito (benefício) máximo do contexto»¹⁶ pode ser obtido bastando que o leitor identifique o referente; o «esforço (custo) mínimo de processamento»¹⁷ é determinado pela baixa necessidade de exigir do leitor conhecimentos da fonética,

¹²Aproveitando a aliteração, não podemos deixar passar a oportunidade de citar o conhecido ditado italiano «traduttore, traditore», que alude às mudanças, modificações e adaptações inerentes à atividade de tradução.

¹³No original: «form follows function» (p. 40).

¹⁴No original: «be relevant to the purpose» (p. 43).

¹⁵No original: «a translator's duty is to translate accurately [...] and in full». (p. 44).

¹⁶No original: «maximum contextual effect (benefit)».

¹⁷No original: «minimum processing effort (cost)».

da fonologia ou da ortografia da língua de partida. Por sua vez, em um ambiente acadêmico podem-se trabalhar elementos mais pormenorizados, provendo mais informações (e.g. etimologia, morfologia, afixos, etc.) e demandando conhecimentos mais profundos do funcionamento da língua em questão.

1.2.3 Formas de tradução

Da classificação de Shuttleworth e Cowie, apresentada por Chesterman (p. 52), podemos destacar — e adaptar à nossa leitura — as formas de tradução: **cultural**, presente em latinizações como <Confúcio> ou <Confucius> (em inglês), para <孔夫子> (em *pinyin*, <Kǒng Fūzǐ>); **horizontal**, cuja definição poderia ser adaptada para abordar, em vez da proximidade cultural de duas línguas, a semelhança do código grafemático (e.g. português e islandês); **vertical**, visível na opção por formas como <shampoo> (por <xampu>) e <layout> (por <leiaute>); **integral**, em que se preservam elementos fonêmicos ou grafemáticos do original (transcrição ou transliteração plena e pura); **interlinear**, quando se apresentam conjuntamente as formas adaptada e original (geralmente entre parênteses); **intralingual**, quando se lida com uma ortografia não-padrão e uma padrão (e.g. as letras de Adoniran Barbosa ou Luiz Gonzaga).

pinyin:
forma oficial
de romanização
do chinês.

Dos tipos intermediários apresentados por Wagner (p. 53), três destacam-se — ainda segundo nossa releitura. A **tradução reduzida** observa-se nos jornais, nas revistas e nos demais meios voltados ao público leigo, uma vez que a «informação básica»¹⁸ pouco ou nada extrapola os limites da necessidade que o leitor tem de associar o nome ao referente correto. Na mesma linha, a **tradução naturalizada** remete, assim como a tradução cultural (supracitada), à latinização, que visa a fazer com que um nome estrangeiro «seja percebido como um [nome] pertencente à cultura de chegada»¹⁹. Por sua vez, a **tradução direta** (ou absoluta) corresponderia a uma transcrição ou transliteração, em que se preservam

¹⁸No original: «basic information».

¹⁹No original: «feels like an original [name] in the target culture».

todas as relações com os elementos, de um ou de outro tipo (fonêmico ou grafêmico), da língua de partida.

1.2.4 O leitor

No que diz respeito aos tipos de leitor que se espera terão contato com o material produzido (p. 54 e 55), Chesterman apresenta exemplos bastante significativos. Os **leitores heterogêneos**, para os quais «um estilo mais popular pode ser [mais] adequado»²⁰, correspondem diretamente aos leitores de jornais e revistas; os **homogêneos**, que se utilizam de «linguajar mais específico e termos técnicos»²¹, seriam os produtores e receptores de textos acadêmicos. Os leitores que sejam falantes nativos ou não-nativos podem ser entendidos, no âmbito da transposição ortográfica, como aqueles conhecedores ou não da língua de partida, respectivamente.

A **distância cultural** de um texto traduzido apresenta-se, no tocante à ortografia, nos sistemas de escrita de partida e de chegada, quer sejam eles próximos (e.g, alfabetos latino e cirílico) ou distantes (e.g, alfabeto grego e silabário cherokee). Quanto aos **destinatários** e **receptores**, a distinção pode ser aplicada eminentemente a transposições grafemáticas que visem a um grupo homogêneo, composto de leitores que podem ser: **excluídos**, que pouco ou nada conseguem decodificar do que esteja grafado, mesmo quando estiver sendo usado um sistema de escrita conhecido; **observadores**, pessoas que têm conhecimento dessa língua, mas para quem as minúcias não têm valor acadêmico; **participativos**, ou seja, pesquisadores que trabalhem na mesma área ou com a mesma língua de partida.

Se o público alvo for constituído de um grupo heterogêneo, como no caso da mídia de massa, torna-se ainda mais «eticamente recomendável traduzir de modo a ser o menor possível o grupo de leitores ex-

²⁰No original: «a more popular style might be appropriate».

²¹No original: «more in-group language, technical terms».

cluídos»²². Isso serve de base (ou justificativa) para considerarmos que a transposição feita pelos jornais —malgrado haver inconsistências entre um e outro— seja adequada —e mesmo eticamente pertinente— ao leitor a que se dirigem os textos. Por sua vez, esses mesmos sistemas mostram-se inadequados para leitores homogêneos que sejam receptores participativos. Para estes leitores é possível —se não necessário— que o código ofereça a maior quantidade possível de informações —e aqui apoia-se nossa proposta. Por exemplo, se um antropônimo não registra uma ou outra consoante, para um público leigo isso pouca ou nenhuma diferença faz, contanto que fique claro qual o referente; para um público especializado, é provável que essa identificação (a do referente) seja igualmente possível, mas a transposição ortográfica talvez pareça pobre, incompleta ou deficiente.

1.2.5 As estratégias

A abordagem pelo viés da **solução de problemas**, transferindo do produto para o processo a análise, aproxima-se do que se deve fazer ao elaborar os sistemas de romanização. No caso específico da transposição ortográfica, o problema concerne à incompatibilidade entre o produto final (o topônimo ou antropônimo apresentado no sistema de escrita da língua de chegada) e os meios empregados (o sistema de transposição grafemática). As **estratégias** ou **ferramentas conceituais** consistem no conjunto de regras e orientações que conduzem a passagem, de um a outro sistema de escrita, de um nome próprio entre duas línguas.

Das três categorias de estratégia discutidas (pp. 57 e 58), nem todas se aplicam à transposição ortográfica de nomes. As **de bloqueio** parecem ser pouco necessárias, enquanto as **de pesquisa** são mais ou menos pertinentes, variando conforme o caso e a conveniência. As **textuais**, por sua vez, situam-se bastante próximas, chegando mesmo a incluir, nas estratégias textuais globais —que consideram o texto como um todo,

²²No original: «ethically appropriate to translate in such a way that the group of excluded readers is as small as possible» (p. 55).

seus propósitos e destinatários— um tipo de adaptação em que não escape resquício algum da cultura da língua de partida, «quicá nem mesmo nomes próprios»²³.

Quanto à **estrangeirização**, um bom exemplo pode ser encontrado em JAROUCHE (2002). Alguns nomes encontram-se em uma grafia algo estranha aos nossos olhos, como ⟨Šahrāzād⟩ e ⟨Ja[°]far⟩ (por ⟨Sherazade⟩ e ⟨Jafar⟩); isso contribui para a sensação de estarmos lendo uma obra de origem estrangeira. A essa estratégia contrapõe-se a **latinização**, que visa a tornar familiar uma forma estrangeira, como se observa em nomes como ⟨Confúcio⟩ (*vide* 1.2.3, p. 47).

O binômio adaptação–estrangeirização é subdividido por Wagner em quatro categorias (p. 59). Da **simplificação** aproximam-se as formas empregadas na mídia; ainda que não se trate de termos técnicos, as minúcias fonéticas, fonêmicas e ortográficas da escrita original são reduzidas, ao ser apresentadas ao leitor leigo. A **explicação** costuma ficar reservada aos manuais de redação e estilo, particularmente nas seções dedicadas à grafia dos nomes de determinadas línguas. A **condensação** pode ser observada na supressão de certos elementos, especialmente os inexistentes em português (e.g, tons com valor distintivo, quantidade vocálica, morfemas de caso, etc.). A **tradução direta** corresponde ao que se faz em uma transcrição ou transliteração, quando determinados elementos (fonéticos ou grafemáticos) do original são transpostos integralmente.

E os conceitos de troca e mudança²⁴ talvez sejam mais representativos do que acontece na transposição ortográfica do que na tradução tradicional. Essa abordagem, que considera a tradução um tipo de recodificação, encaixa-se bastante adequadamente em alguns procedimentos como a troca de SH por x (e.g, ⟨shampoo⟩ → ⟨xampu⟩) ou a de w por u (e.g, ⟨Kuwait⟩ → ⟨Kuuait⟩ → ⟨Kuit⟩).

²³No original: «perhaps not even proper names» (p. 58).

²⁴No original, *shift* e *change*

1.2.6 Outras estratégias

Por fim, Chesterman apresenta estratégias de ordem sintática, semântica e pragmática. Das **estratégias sintáticas**, a **tradução literal** é a que mais perto se encontra da correspondência grafema-grafema da transliteração, considerando-se que o resultado apresenta-se «ao máximo próxima da forma encontrada na língua de partida»²⁵. O **empréstimo** ou **decalque** depende menos de decisões pertinentes à transposição grafemática do que da linha de tradução do texto em si — por exemplo, um jornal pode escrever «Al-Qaeda, A Base». A **transposição**, no sentido apresentado por Chesterman, aparentemente não se aplica a muitas instâncias do trabalho com nomes próprios, salvo nos casos em que um *w* ou *y* sejam substituídos por um *u* ou *i* e sofram alguma metátese (vide 1.2.5, p. 50). A **mudança na estrutura sintagmática**, a **mudança de coesão**, a **troca de nível** e a **mudança de esquema** têm correlação com mudanças (ou supressões) no registro de elementos como qualidade ou quantidade vocálica ([o] → [u]; [a:] → [a]), acento tônico (oxítono para paroxítono), afixos (artigo definido ⟨al⟩ grafado junto, separado ou unido por hífen à palavra) ou consoantes dobradas (mantidas como no original ou reduzidas a uma só). A **troca de unidade**, a **mudança na estrutura oracional** e na **estrutura do período** parecem não se aplicar ao nível grafemático.

Das **estratégias semânticas** (p. 61) e das **pragmáticas** (p. 62), poucos são os procedimentos que encontram correspondentes na transposição ortográfica. A **hiponímia–hiperonímia** pode ser observada quando se substitui um grafema associado ao arquifonema por grafemas relacionados a alofones (no árabe, A por E ou O). O **filtro cultural** atua da mesma forma que a latinização (no exemplo de ⟨Confúcio⟩, em 1.2.3, p. 47). A **mudança de informação** opera em mão dupla, acrescentando elementos ausentes no original ou suprimindo os que forem considerados pouco relevantes para o público alvo.

²⁵No original: «maximally close to the source-language form».

1.2.7 Metáforas e palavras isoladas

Nas soluções padrão para tradução de metáforas (p. 66), todos os modelos podem ser empregados, de uma ou de outra forma, para transpor uma palavra. A **reprodução da mesma imagem** e a **tradução literal** podem ser encontradas em passagens que apresentam a pronúncia “aproximada” de uma palavra estrangeira (e.g, Taunay pronuncia-se «tô-né»). O uso de uma imagem diferente ou de um símile é recorrente quando se trabalham nomes que possuam correspondentes na língua de chegada (e.g, ⟨Joãozinho e Maria⟩ por ⟨Hänsel und Gretel⟩; ⟨Pedro⟩ por ⟨Peter⟩). A **paráfrase** é análoga a figuras de linguagem como a perífrase e a antonomásia, que substituem o referente por um epíteto ou característica marcante. E a **supressão**, por sua vez, age sobre letras ou fonemas omitidos do produto final.

1.2.8 Avaliação

Chesterman apresenta algumas normas para avaliação da qualidade do produto final. A **norma de adequação** (ou aceitabilidade), que poderia cotejar as soluções utilizadas (para grafar esta ou aquela palavra) com as previamente existentes, parece-nos ser pertinente a um nível um pouco mais básico, qual seja, o da fonotática. Em outras palavras, «[se] estiver adequado, atenderá as expectativas do leitor acerca de [como uma dada palavra deve ser lida]»²⁶. A **norma de relação** lida com a maior ou menor necessidade d’a transposição ser fidedigna, oferecendo mais ou menos elementos do nome próprio original. A **norma de comunicação**, de mesmo modo, verifica quão decodificável apresenta-se o produto final. A **norma de responsabilidade**, por sua vez, possivelmente perpassa toda e qualquer instância da atividade de tradução, quer seja ela de textos ou de nomes próprios.

²⁶No original: «[if] it fits appropriately, it will meet the reader’s expectations about [how a given word should be pronounced]» (p. 92).

1.2.9 Releitura

Os conceitos trabalhados por CHESTERMAN & WAGNER (2002) são de alguma forma tão pertinentes à prática e ao estudo do tema deste trabalho, que podemos reler alguns trechos sob o viés da transposição ortográfica —com as substituições entre colchetes—, como se segue:

Podemos dizer, por exemplo, que [uma letra] *x* em [um sistema de escrita] *a* na maioria das vezes passa para [o sistema de escrita] *b* como [a letra] *y* — mas isso não significa que todos os tradutores sempre farão [a transposição] dessa mesma forma, em todos os casos.²⁷

Imagino que o cliente ficará satisfeito [...], o leitor entenderá mais [e] a [transposição] terá a [leitura desejada].²⁸

[...] os leitores [lerão] com mais facilidade [a letra] *a* do que [a letra] *b*, ou a estratégia *x* de [transposição] trará resultados melhores do que a estratégia *y* [...]²⁹

1.3 Na academia

A seguir, faremos uma análise de como diferentes formas romanizadas dos mesmos antropônimos podem denotar perspectivas distintas sobre os respectivos referentes.

1.3.1 A falsafa

Alguns exemplos de que formas diferentes podem implicar conteúdos diferentes pode ser encontrado em ATTIE (2002). Acerca do surgimento

²⁷No original: «We can say, for instance, that [letter] X in [writing system] A very frequently gets translated into [writing system] B as [letter] Y — but this does not mean that every translator will always [transpose] it that way, under all circumstances.» (p. 64).

²⁸No original: «I predict that the client will like it [...], the reader will understand better [and] the [transposition] will have the desired [output, i.e., the desired pronunciation].» (p. 65).

²⁹No original: «[...] readers will find [letter] A easier to [read] than [letter] B, or that [transposing] strategy X will have a better effect than strategy Y [...].» (p. 65).

de <falsafa>, conta-nos o autor que esse termo teria sido resultado da transposição do grego <φιλοσοφία>:

Vale lembrar que se, por um lado, na língua grega, os morfemas φιλία/σοφία (philia/sophia) se unem para dar, entre outras, a idéia de “amor à sabedoria”, por outro lado, em árabe — assim como nas transcrições que encontramos em outras línguas [...] — a idéia que liga os conceitos de amor e sabedoria se dá somente por uma analogia e um retorno ao termo grego. Os vocábulos usados para significar “amor” e “sabedoria”, na língua árabe, não possuem qualquer semelhança com os radicais gregos decorrendo, portanto, que no vocábulo فلسفة (falsafa) não há qualquer idéia que provenha dos radicais próprios da língua árabe. É uma pura transcrição da língua grega. (p.29)

Trata-se, pois, menos de uma simples tradução do que de uma assimilação. Neste processo, o termo passa a assumir valores próprios do meio em que é utilizado. Além de não se manterem todas as possíveis relações originais —entre os morfemas e entre a palavra e os contextos em que seja empregada—, novos usos e significados passam a ser a ele associados.

Por exemplo, ao falar dos usos correntes que fazemos em expressões do tipo «filosofia de vida», assinala o autor que «a intensidade com que o Ocidente e o Oriente assimilaram algumas tradições da filosofia da antiga Grécia não foi a mesma» (p. 32); e arremata dizendo que «a falsafa fôra, em muitos casos, mais estrangeira para os árabes do que a filosofia o foi para os ocidentais» (idem).

A palavra, assim cristalizada, fixa um significado específico; delimita um dentre diferentes significados originalmente possíveis. Em suma, refere-se a «[uma] filosofia original e renovada que não se confunde com particularidades filosóficas anteriores» (p. 33).

E a assimilação de conceitos não constitui matéria estranha à tradução — ou ao que se convencionou chamar de tradução. Observando a contraposição —ou o contraste— entre <falsafa> e <filosofia>, nota-se que ambos os termos designam, a rigor, uma mesma herança da cultura

grega clássica. Mas, no tocante à utilização e ao valor atribuído à disciplina já assimilada, eles não portam somente informações geopolíticas —quem os utiliza e onde—, mas também culturais — como o fazem.

1.3.2 Nomes próprios

Attie grafa nomes próprios sistematicamente de forma transliterada: Al-Kindī; Al-Fārābī; Ibn Sīnā (Avicena); Ibn Rušd (Averróis); Al-Ġazālī. No entanto, em citações de outras obras (de outrem), mesclam-se formas transliteradas e latinizadas. Por exemplo, na citação de Carra de Vaux (p. 34) leem-se os nomes «Al-Kindī, Al-Fārābī, **Avicena**, [...], **Averróis**» (grifos nossos), onde se podem notar as formas utilizadas para nomear Ibn Sīnā e Ibn Rušd. Não seria próprio ou adequado supor um desconhecimento de Carra de Vaux acerca da grafia transliterada desses dois nomes, tampouco indicar uma falta de uniformidade na grafia — algo que de fato ocorre, mas não por equívoco ou descuido, decerto. Seria mais adequado inferir ter sido esse um ato deliberado, uma intenção informativa pertinente ao *status* de que gozam esses dois falāsifa entre os ocidentais, sendo ambos expoentes da falsafa, em especial na Península Ibérica.

falāsifa:
filósofos
da falsafa

Algo semelhante pode ser observado também em outras instâncias. Os topônimos são grafados em sua forma corrente no português brasileiro (e.g, Bagdá, Egito, Marrocos, Síria, Líbano, Iraque); alguns nomes conhecidos são transliterados (e.g, <Muḥammad> por <Maomé> e <Islām> por <Islã>), enquanto outros são aportuguesados (e.g, <Alcorão>, em vez de <Al-qur'ān>). Além disso, há palavras derivadas como <arábico>, <arabismo> e <arabizar>.

É interessante ressaltar também que, independente d'o primitivo registrado ter sido aportuguesado ou não, tais derivações tendem a recuperar elementos por vezes bastante essecíficos da forma original das palavras derivadas. Basta comparar <islâmico> com <Islā (Islām)> e <corânico> com <(Al)Corão (Al-qur'ān)> para notar que o sufixo <-ico> (formador de adjetivos) toma emprestado à forma original da palavra

árabe a consoante final do lexema (morfema lexical), de onde surgem o ⟨-mico⟩ e o ⟨-nico⟩ das formas derivadas.

Há outros exemplos dessa aparente assimetria. Por exemplo, o autor faz opções interessantes quando afirma que «[no] Islām a teologia denominou-se كلام (kalām); a mística é o صوفية (ṣūfiya), isto é, o sufismo; e o entendimento pela demonstração lógica é a فلسفة (falsafa).» (p. 48). Segue usando ⟨kalām⟩ para tratar do equivalente árabe do que no ocidente entende-se como Teologia, e ⟨falsafa⟩ para o equivalente à Filosofia; mas adota o termo ⟨sufismo⟩, quando se refere à mística. Poder-se-ia desejar que houvesse simetria —ou “coerência”, talvez— no emprego dos termos; deveria ser tudo ou de um jeito (com termos da língua de partida) ou de outro (com termos da língua de chegada). Entretanto, deve-se salientar que um tal proceder careceria das sutilezas expressivas que encontramos ao analisar detidamente a escolha das palavras e as grafias adotadas; suprimiria ao menos uma das diferenças entre os pares ⟨kalām⟩/⟨teologia⟩, ⟨ṣūfiya⟩/⟨mística⟩ e ⟨falsafa⟩/⟨filosofia⟩: a existência, em português, de um termo próprio —⟨sufismo⟩— apenas para o segundo.

1.3.3 Aristóteles

No que diz respeito aos nomes próprios, tomem-se os exemplos de Aristóteles, Platão e Plotino. Quando nos referimos a esses três filósofos gregos por intermédio de seus nomes adaptados à língua árabe —respectivamente, Aristūṭālis, Aflaṭūn e Afluṭīn—, a informação nestas formas contida não se restringe a um mero registro de como os falāsifa a eles se referiam. Atendo-nos somente ao caso de Aristóteles, afirma-se que «a questão dos apócrifos, isto é, os textos atribuídos equivocadamente a Aristóteles fizeram com que, na verdade, Aristūṭālis não fosse o mesmo Aristóteles que conhecemos hoje» (p. 144). Esse fato redundou em «uma doutrina que o próprio Aristóteles não enunciou em nenhuma de suas obras» (idem), ou seja, as interferências (ou modificações) foram de tal ordem que «[o] Aristóteles árabe — Aristūṭālis — é, na verdade,

um Aristóteles neoplatonizado» (idem). No entanto, «o que poderia parecer um equívoco histórico da entrada dos elementos neoplatônicos sob o nome de Aristóteles, acabou se conformando num dos pilares centrais da construção do sistema de Al-Fārābī e Ibn Sīnā.» (p. 145).

Portanto, entendemos que a manutenção da forma transliterada ⟨Aristūṭālis⟩ reflete as adaptações particulares operadas não apenas sobre a imagem acústica do nome Aristóteles, mas também sobre os ensinamentos ao filósofo atribuídos. Podemos dizer que, de um certo ponto de vista, Aristóteles e Aristūṭālis apresentam-se como duas figuras históricas relativamente distintas, motivo pelo qual se torna válido —ou mesmo imprescindível— referir-mo-nos a um e a outro pelas respectivas variantes.

1.3.4 Diferentes referentes

Percebemos que uma maior ou menor proximidade com a forma original da palavra pode ser usada para denotar certas particularidades do referente. Pode-se preservar o sentido original, mitigando a influência das adaptações pelas quais passam os conceitos, ao ser transpostos de uma para outra língua, de uma para outra cultura; ou podem-se distinguir, do sentido original, as influências dessa mesma sorte de adaptações. A escolha entre adotar ou não uma grafia distinta da original pode ser usada justamente para marcar diferenças oriundas das interações por que tenha passado uma determinada palavra ou um dado nome, a fim de ressaltar ou indicar eventuais alterações dos conceitos ou valores originais.

Dessa forma, topônimos e, principalmente, antropônimos deixam de ser significantes de referentes unívocos. Eles passam a existir como as tantas palavras e expressões que tornam o trabalho do tradutor algo menos mecânico do que intelectual. E, mesmo quando não há essa sorte de distinções, cabe ao tradutor a tomada de decisões que devem pesar os valores codificados na língua de partida e a decodificação na língua de chegada.

1.4 Questões práticas e teóricas na academia

Neste primeiro capítulo procurámos apresentar a transposição ortográfica como algo presente tanto no cotidiano de qualquer pessoa quanto no trabalho de pesquisa acadêmica; e como objeto relevante para os estudos tradutológicos, uma vez que não muito difere dos procedimentos e das ferramentas em uso na tradução *stricto sensu*.

São basicamente três os principais contextos, no meio acadêmico, em que nosso trabalho pode contribuir:

- em trabalhos acadêmicos, que por variados motivos precisem registrar uma palavra ou um nome de origem árabe da forma mais próxima possível à original;
- em dicionários, tanto nos bilíngues quanto nos etimológicos ou nos que apresentem entradas em que se faça necessário registrar o étimo de uma palavra;
- em bibliotecas, especialmente nas fichas catalográficas (físicas ou digitais), no registro de autores, títulos e assuntos.

CAPÍTULO 2

A língua árabe e sua escrita

Verba volant scripta manent.

«As palavras voam,
os escritos ficam.»

ditado latino

INFORMAMOS as características do árabe e a relação da língua com sua escrita. Esses pontos serão relevantes ao tratarmos das escolhas pertinentes às correspondências grafema–fonema de nossos esquemas de romanização e às formas de representação de fenômenos fonéticos e de elementos morfológicos.

2.1 A língua

O árabe faz parte do grupo das línguas semíticas, ramo da família das línguas afro-asiáticas. Atualmente é falado na Península Arábica e no norte da África, do Mar Vermelho ao Oceano Atlântico, sendo língua materna de mais de 180 milhões de pessoas (CRYSTAL, 1997, p. 318).

2.1.1 Variantes

A língua árabe apresenta-se em uma forma clássica e em uma coloquial — sendo esta composta de diversos dialetos, que variam conforme o país ou a região. O árabe clássico (ou literário) é a língua sagrada do Islã, havendo uma variante, o árabe padrão moderno, usada como *lingua franca* entre falantes de árabe (CRYSTAL, 1997, p. 318).

lingua franca:
língua comum
a falantes de
diferentes línguas
ou dialetos,
mutuamente
inteligíveis
ou não.

2.1.2 Morfologia

O árabe padrão moderno apresenta:

- dois gêneros: masculino; feminino. Em termos de concordância, há de se atentar também para a animacidade, pois em árabe faz-se distinção entre humanos e não-humanos (animais e seres inanimados).
- três números: singular; dual; plural. O dual é sempre regular; o plural, por sua vez, pode ser do tipo regular (uma forma) ou fracto, com trinta diferentes formas (COWAN, 2007, pp. 193–195).
- três casos: nominativo; acusativo; genitivo. Todos são marcados por um sufixo vocálico próprio: o nominativo, por /u/; o acusativo, por /a/; o genitivo, por /i/.
- duas formas de definitude: definido; indefinido. Um nome definido normalmente está prefixado pelo artigo definido ⟨ال⟩ |al-|, salvo quando houver outro elemento que o defina (tabela 6)

Tabela 6: O artigo definido e o genitivo.

árabe	romanizado	significado
الباب	al-bāb	a porta
البيت	al-bayt	a casa
باب البيت	bāb al-bayt ⁱ	a porta da casa*

*Em árabe, o núcleo do sintagma nominal encontra-se definido pelo nome que está no genitivo.

Os adjetivos concordam com os substantivos em gênero, número, caso e definição (tab. 7) — quando o substantivo estiver definido, ao adjetivo prefixa-se o artigo definido ⟨ال⟩ /ʔal/.

Isso afeta, entre outras construções, «todos os plurais árabes que não se referem a seres racionais» (COWAN, 2007, p. 37), que, para efeito de concordância, «são gramaticalmente feminino singular, de forma tal que todos os adjetivos que qualificam plurais de objetos inanimados ou idéias abstratas e todos os pronomes que substituem um ou outro devem constar no feminino.»(idem).

Quanto à tipologia, CRYSTAL (1997) classifica o árabe como uma língua inflexional, sintética ou fusional, na qual «relações gramaticais são expressas por mudanças na estrutura interna das palavras. [...] Latim, grego e árabe são exemplos nítidos.»¹. Em outras palavras, a principal característica de sua morfologia é a base triconsonantal de suas palavras; com um sistema de infixos e afixos, flexionam-se ou derivam-se outras palavras (tab. 8).

As formas finitas dos verbos concordam com o sujeito em pessoa, gênero e número. Há formas masculinas e femininas para a segunda e terceira pessoas do singular, para a terceira do dual e para a segunda e terceira do plural; não há distinção de gênero para a primeira pessoa do singular, segunda do dual e primeira do plural (tab. 9).

¹No original: «Grammatical relationships are expressed by changing the internal structure of the words. [...] Latin, Greek, and Arabic are clear cases.» (p. 295).

Tabela 7: Número, caso, definição e gênero em {professor bom}.

núm.	caso	def.	gên.	árabe	romanizado
S	N	DEF	M	المُعَلِّمُ الْجَيِّدُ	al-mu ^c allim ^u al-jayyid ^u
			F	المُعَلِّمَةُ الْجَيِّدَةُ	al-mu ^c allima ^{tu} al-jayyida ^{tu}
		IN	M	مُعَلِّمٌ جَيِّدٌ	mu ^c allim ^{un} jayyid ^{un}
			F	مُعَلِّمَةٌ جَيِّدَةٌ	mu ^c allima ^{tun} jayyida ^{tun}
A	DEF	M	M	المُعَلِّمُ الْجَيِّدُ	al-mu ^c allim ^a al-jayyid ^a
			F	المُعَلِّمَةُ الْجَيِّدَةُ	al-mu ^c allima ^{ta} al-jayyida ^{ta}
		IN	M	مُعَلِّمًا جَيِّدًا	mu ^c allim ^{an} jayyid ^{an}
			F	مُعَلِّمَةً جَيِّدَةً	mu ^c allima ^{tan} jayyida ^{tan}
G	DEF	M	M	المُعَلِّمُ الْجَيِّدُ	al-mu ^c allim ⁱ al-jayyid ⁱ
			F	المُعَلِّمَةُ الْجَيِّدَةُ	al-mu ^c allima ^{ti} al-jayyida ^{ti}
		IN	M	مُعَلِّمٌ جَيِّدٌ	mu ^c allim ⁱⁿ jayyid ⁱⁿ
			F	مُعَلِّمَةٌ جَيِّدَةٌ	mu ^c allima ^{tin} jayyida ^{tin}
D	N	DEF	M	المُعَلِّمَانِ الْجَيِّدَانِ	al-mu ^c allim ^{tī} al-jayyidān ⁱ
			F	المُعَلِّمَتَانِ الْجَيِّدَتَانِ	al-mu ^c allimatān ⁱ al-jayyidatān ⁱ
		IN	M	مُعَلِّمَانِ جَيِّدَانِ	mu ^c allim ^{tī} jayyidān ⁱ
			F	مُعَلِّمَتَانِ جَيِّدَتَانِ	mu ^c allimatān ⁱ jayyidatān ⁱ
A/G	DEF	M	M	المُعَلِّمَيْنِ الْجَيِّدَيْنِ	al-mu ^c allimayn ⁱ al-jayyidayn ⁱ
			F	المُعَلِّمَتَيْنِ الْجَيِّدَتَيْنِ	al-mu ^c allimatayn ⁱ al-jayyidaynayn ⁱ
		IN	M	مُعَلِّمَيْنِ جَيِّدَيْنِ	mu ^c allimayn ⁱ jayyidayn ⁱ
			F	مُعَلِّمَتَيْنِ جَيِّدَتَيْنِ	mu ^c allimatayn ⁱ jayyidaynayn ⁱ
P	N	DEF	M	المُعَلِّمُونَ الْجَيِّدُونَ	al-mu ^c allimūn ^a al-jayyidūn ^a
			F	المُعَلِّمَاتُ الْجَيِّدَاتُ	al-mu ^c allimāt ^u al-jayyidāt ^u
		IN	M	مُعَلِّمُونَ جَيِّدُونَ	mu ^c allimūn ^a jayyidūn ^a
			F	مُعَلِّمَاتُ جَيِّدَاتُ	mu ^c allimāt ^{un} jayyidāt ^{un}
A/G	DEF	M	M	المُعَلِّمِينَ الْجَيِّدِينَ	al-mu ^c allimīn ^a al-jayyidīn ^a
			F	المُعَلِّمَاتُ الْجَيِّدَاتُ	al-mu ^c allimāt ⁱ al-jayyidāt ⁱ
		IN	M	مُعَلِّمِينَ جَيِّدِينَ	mu ^c allimīn ^a jayyidīn ^a
			F	مُعَلِّمَاتُ جَيِّدَاتُ	mu ^c allimāt ⁱⁿ jayyidāt ⁱⁿ

[Núm]ero: [s]ingular; [D]ual; [P]lural. [Caso]: [N]ominativo; [A]cusativo; [G]enitivo. [Def]inição: [DEF]inido; [IN]definido. [Gên]ero: [M]asculino; [F]eminino.

Tabela 8: Raiz triconsonantal e formas derivadas referentes a {escrever}.

árabe	romanização	significado
ك + ت + ب	K + T + B	(raiz triconsonantal)
كَتَبَ	kataba	[ele] escreveu
كُتِبَ	kutiba	foi escrito
كِتَابَ	kitāb	livro
كُتُبُ	kutub	livros
كَاتِبَ	kātib	escritor
مَكْتَبَ	maktab	escritório
مَكْتَبَةٌ	maktaba ^t	biblioteca

O negrito marca as consoantes da raiz.

Tabela 9: Conjugação verbal — perfectivo (aspecto perfeito) de <كَتَبَ> |kataba| {escrever}.

núm.	pess.	gên.	árabe	romanizado
S	1	M/F	كَتَبْتُ	katabtu
	2	M	كَتَبْتَ	katabta
	2	F	كَتَبْتِ	katabti
	3	M	كَتَبَ	kataba
	3	F	كَتَبَتْ	katabat
	3	F	كَتَبَتْ	katabat
D	2	M/F	كَتَبْتُمَا	katabtumā
	3	M	كَتَبَا	katabā
	3	F	كَتَبَتَا	katabatā
P	1	M/F	كَتَبْنَا	katabnā
	2	M	كَتَبْتُمْ	katabtum
	2	F	كَتَبْتُنَّ	katabtunna
	3	M	كَتَبُوا	katabū
	3	F	كَتَبْنَ	katabna

[Núm]ero: [s]ingular; [d]ual; [p]lural.

[Pess]oa: [1] primeira; [2] segunda; [3] terceira.

[Gên]ero: [m]asculino; [f]eminino.

2.1.3 Sintaxe

A ordem dos termos de uma oração tende a seguir a ordem vso, exceto quando se usam pronomes pessoais, que precedem o verbo (svo). Em construções nominais no tempo presente, ocorre cópula zero, i.e, «[n]as frases nominais simples, compostas por sujeito e predicativo do sujeito, as ligações correspondentes aos verbos ‘ser’ e ‘estar’ não são expressas.» (COWAN, 2007, p. 19).

vso: verbo, sujeito, objeto.

2.2 A escrita

Tema central de nossa pesquisa, veremos agora o abjad árabe e algumas de suas características.

2.2.1 Designação

A escrita árabe, diferentemente do alfabeto latino, concentra-se no registro das consoantes. Há variados termos para designar esse tipo de sistema de escrita, como consonantário (à semelhança de <silabário>) e alfabeto consonantal (COULMAS, 1991). CRYSTAL (1997) chama-o de «alfabeto “consonantal”» (com aspas); e BAUER (1996) explicita essa ressalva, salientando que «apesar de não haver caracteres básicos para as vogais curtas, a designação comum de que a escrita árabe seja ‘consonantal’ é incorreta, uma vez que se representam as vogais longas, mas não a geminação de consoantes.»²

Por isso, preferimos adotar o termo <abjad>, que DANIELS (1996c) explica como sendo formado «das primeiras letras do exemplo mais difundido, a escrita árabe, em sua ordem histórica»³ — <ا ب ج د>, formado pelas letras <ا>, <ب>, <ج> e <د>, que correspondem às quatro primeiras do

²No original: «[t]hough there are no basic characters for the short vowels, the common designation of the Arabic script as ‘consonantal’ is incorrect, since the long vowels are represented but consonant gemination is not.» (p. 561).

³«from the first letters of the most widespread exemple, the Arabic script, in their historic order» (p. 4).

abjad fenício. Com isso pretendemos empregar um termo que seja mais adequado às características e ao funcionamento da escrita árabe, além de evitar associações equivocadas, como a de que ela registraria única e exclusivamente consoantes, ou a de que se trataria de um alfabeto destituído de vogais.

2.2.2 Origens

2.2.2.1 da escrita

As origens da escrita árabe remetem, em última instância, ao aramaico. HIGOUNET (2003) supõe que «[...] ele seria derivado do alfabeto nabateu e, por ele, da escrita aramaica.» (p. 75); BAUER (1996) afirma que «a escrita norte-arábica desenvolveu-se como um ramo da escrita aramaica nabateia»⁴ e que as primeiras inscrições datam do século IV.

Apesar de haver relações entre língua árabe, seu sistema de escrita e religião islâmica, HIGOUNET (2003) ressalta que os três «são fenômenos distintos, cujos domínios nem sempre coincidem. [...] A própria escrita árabe surgiu antes do Islã. Mas foram os avanços do islamismo que evidentemente favoreceram sua difusão.» (p. 75).

2.2.2.2 das consoantes

O abjad árabe surgiu da modificação de uma outra escrita. COULMAS (1991) relata que «enquanto o alfabeto norte-semítico consiste de apenas 22 letras, o árabe possui um inventário consonantal mais rico, com 28 letras básicas. Algumas das letras adicionais foram criadas adicionando-se marcas diacríticas a outras.»⁵. HIGOUNET (2003) acrescenta que, «[p]ara distinguir certas consoantes diferentes figuradas [pela mesma forma bá-

⁴No original: «The North Arabic script developed as a branch of Nabatean Aramaic script» (p. 559).

⁵No original: «[...] while the North Semitic alphabet consists of twenty-two letters only, Arabic has a richer inventory of consonants, and comprises twenty-eight basic letters. Some of the additional letters were created by adding diacritical marks to others.» (p. 150).

sica], recorreu-se desde cedo a pontos diacríticos (ou seja, distintivos) postos acima ou abaixo das letras.» (p. 78). Esses pontos são chamados de *ʿajām* [ʿajām]. (tab. 10)

Tabela 10: Dois grupos de grafemas derivados.

forma básica	formas derivadas*		
	sem ʿajām	com ʿajām	
ب	—	ب /b/	ت /t/ ث /θ/
ح	ح /x/	ج /ʒ/	خ /ħ/

*Grafemas pareados com os respectivos fonemas.

Esse processo está em acordo com a afirmação de CRYSTAL (1997) de que «[e]m uma língua em que há muitos novos sons, deve-se tomar uma decisão acerca do que se deve fazer: inventar novas letras; combinar letras (formar dígrafos); ou empregar diacríticos (acentos).»⁶. E o mesmo fenômeno pode ser observado, atualmente, nas línguas não-semíticas que adotaram a escrita árabe e tiveram de imprimir-lhe algumas modificações (*vide* 2.2.4, página 69).

2.2.2.3 das vogais

O sistema fonológico árabe possui seis vogais, havendo distinção de quantidade: três vogais breves /a, i, u/; três longas /a:, i:, u:/. O registro das vogais breves é opcional, usando-se marcas colocadas acima ou abaixo das letras (*vide* 2.2.5.3, p. 72); trata-se de um desenvolvimento mais recente, do século VIII, e que é de pouco uso na escrita atual (HIGOUNET, 2003).

A respeito de ser opcional a indicação das vogais breves, DANIELS (1996a) afirma que:

⁶No original: «For a language where there are many new sounds, a decision has to be made about whether to invent new letters, combine letters into digraphs, or go in for diacritics (such as accents).» (p. 367).

há línguas para as quais um alfabeto não é o sistema de escrita ideal. os abjads semíticos prestam-se efetivamente à estrutura do hebraico, do aramaico e do árabe [...]. As marcas vocálicas suplementais desses abjads servem para resolver ambiguidades entre palavras com as mesmas consoantes e que não poderiam ser distinguidas pelo contexto.⁷

COULMAS (1991) acrescenta que as vogais são sempre marcadas «somente no Qur'ân e na poesia»⁸ e HIGOUNET (2003) completa, afirmando que o propósito seria o de orientar a leitura, indicando «a vocalização exata necessária à boa recitação do Corão» (p. 78).

Sobre as vogais longas, COULMAS (1991) acrescenta que «[a] indicação de vogais, na escrita árabe, desenvolveu-se na mesma linha das demais escritas semíticas, isto é, com *matres lectionis* e marcas diacríticas»⁹, explicando que «[o]s sinais acrescidos às consoantes, para indicar a qualidade vocálica de uma sílaba, são chamados de *matres lectionis* ou 'mães da leitura'»¹⁰. Um exemplo do hebraico seria a sequência dalet-vav-dalet (דוד), que pode representar tanto (dod) {tio} quanto (David); para desfazer essa ambiguidade, atualmente acrescenta-se um yud na grafia de (דייד) |David|, tornando-o distinto de (דוד) |dod|.

2.2.3 Os usuários

HIGOUNET (2003) afirma que «[a] escrita árabe é, com o hebraico quadrado, a única escrita consonântica ainda em uso [...] e é hoje, depois da escrita latina, uma das grandes escritas internacionais» (pp. 74-75).

⁷No original: «[...] there are languages for which an alphabet is not an ideal writing system. The Semitic abjads really do fit the structure of Hebrew, Aramaic, and Arabic very well [...]. The supplemental vowel markings on the abjads serve to remove ambiguity among words with the same consonants which cannot be resolved from the context.» (p. 27).

⁸No original: «only in the Qur'ân and in poetry» (p. 150).

⁹No original: «[v]owel indication in Arabic writing developed along the lines of other Semitic scripts: that is, with *matres lectionis* and diacritical marks» (p. 150).

¹⁰No original: «[c]onsonant signs that are used for this purpose of indicating the vowel quality of a syllable are called *matres lectionis* or 'mothers of reading'» (p. 147).

Nessa mesma linha encontra-se SACKS (2003), para quem «[o]s Três Grandes são os alfabetos romano, árabe e cirílico, cada qual servindo a diversas nações e línguas.»¹¹

De acordo com levantamento feito por SACKS (2003), o abjad árabe é o terceiro sistema de escrita mais utilizado no mundo, com 291 milhões de usuários. Está atrás apenas do alfabeto latino e da escrita chinesa; e à frente do abugida devanagari, do alfabeto cirílico, do abugida bengali e da escrita japonesa (tab. 11).

Tabela 11: Quantidade de usuários dos sistemas de escrita mais difundidos.

sistemas de escrita		usuários (em milhões)
tipo	denominação	
alfabeto	latino	1900
escrita	chinesa	900
abjad	árabe	291
abugida	devanagari	260
alfabeto	cirílico	252
abugida	bengali	125
escrita	japonesa	118

Dados de Sacks (2003)

2.2.4 Línguas não-semíticas

O abjad árabe pode ser usado para grafar, além do árabe padrão moderno, várias línguas não-semíticas. KAYE (1996) cita, entre as mais importantes,

as línguas berberes do norte da África do Norte, exceto o tuaregue [...]; as línguas iranianas persa, pachto, curdo e balúchi;

¹¹No original: «[t]he International Three are the Roman, Arabic, and Cyrillic alphabets, each serving multiple nations and languages.» (p. 2).

as indo-arianas urdu, sindi e caxemíri; a dravídica moplá; e as austronésias sulu, malagaxe e malaio.¹²

e SACKS (2003) completa essa lista com «[...] núbio no Sudão [e] uigur na China»¹³.

A esse respeito, HIGOUNET (2003) afirma que

[a] escrita árabe é, no mundo muçulmano de hoje, como a escrita chinesa em seu vasto domínio, um elemento de unidade, visto que é o veículo da língua clássica compreendida por todos os letrados, ao passo que a língua falada se fragmentou em múltiplos dialetos. (p. 79)

E HIGOUNET (2003) traça o percurso que teria sido feito, lembrando que

[s]endo a escrita do Qur'ân, ela espalhou-se da Península Arábica para todo o Oriente Próximo, centro-oeste e sudeste da Ásia, e muitas partes da África e algumas regiões periféricas da Europa. Ela foi adaptada para mais línguas pertencentes a mais famílias do que qualquer outra escrita semítica.¹⁴

o que, como seria de se esperar,

levou à formação de uma gama de variados tipos próprios de cada país e língua, à semelhança das particularidades do alfabeto romano em toda sua sorte de adaptações. [...] Das línguas para as quais o alfabeto árabe foi adaptado, algumas têm um inventário fonêmico mais rico do que o árabe. Caracteres novos ou derivados tiveram de ser cunhados para grafar essas línguas.¹⁵

¹²No original: «the Berber languages of North Africa, except Tuareg [...]; the Iranian languages Persian, Pashto, Kurdish, and Balochi; Indo-Aryan Urdu, Sindhi, and Kashmiri; Dravidian Moplá; and Austronesian Sulu, Malagasy, and Malay.» (p. 743).

¹³No original: «[...] Nubian in Sudan [and] Uighur in China» (p. 10).

¹⁴No original: «[a]s the script of the Qur'ân it spread out from the Arabian peninsula throughout the Near East, western central and south-eastern Asia, as well as to great parts of Africa and some peripheral regions of Europe. It was adapted to more languages belonging to more families than any other Semitic script.» (p. 150).

¹⁵No original: «led to the formation of a variety of different national and language-specific types much like the peculiarities of the Roman alphabet in its many specific applications. [...] Of the languages to which the Arabic alphabet was applied, some have a richer inventory of phonemes than Arabic. New or derived letter signs had to be invented for writing these languages.» (p.151).

Esse processo não muito difere do que ocorreu com o próprio abjad árabe (*vide* 2.2.2.2, p. 65).

Em tempos mais remotos, a adoção do abjad árabe ocorreu também «na Espanha reconquistada, com a literatura aljamiado de língua espanhola e, na Europa oriental, com dialetos eslavos» (HIGOUNET, 2003, p. 78). COMRIE (1996, p. 682) e KAYE (1996, p. 744) lembram que foi usado em turco até 1928.

2.2.5 Características básicas

Escreve-se o abjad árabe da direita para a esquerda, em linhas que se sucedem de cima a baixo. Além disso, CAGLIARI (2009) e COWAN (2007) lembram que não se faz, como no alfabeto latino, distinção de caixa-alta e caixa-baixa (maiúsculas e minúsculas). Para nosso trabalho, três outras características básicas do abjad árabe são-nos relevantes: a distinção entre letras solares e lunares; as funções dos grafemas das vogais longas; e o uso das marcas suplementares.

2.2.5.1 Letras solares e letras lunares

O abjad árabe constitui-se de 28 letras, que se dividem em dois grupos (tabela 12): <حروف شمسية> |ḥurūf xamsiyya^t| {letra[s] solar[es]}; <حروف قمرية> |ḥurūf ḵamariyya^t| {letra[s] lunar[es]}. As **letras solares** têm a característica de representar as consoantes coronais —exceto <ج> /ʒ/—, o que as distingue das **letras lunares**; quando iniciam uma palavra a que se prefixe o artigo definido <ال> |al-|, ocasionam a assimilação regressiva do /l/ (tab. 13).

SABBAGH (1988) explica a origem dos termos <solar, lunar> informando que as letras solares «fazem parte do grupo da letra “šīn”, que inicia a palavra شمس [š]ams (sol)», e que as lunares «fazem parte do grupo da letra “qāf”, que inicia a palavra قمر qamar (lua)».

Tabela 12: Letras solares e lunares, com os respectivos fonemas.

solares		lunares	
graf.	fon.	graf.	fon.
ت	/t/	ب	/b/
ث	/θ/	ج	/ʒ/
د	/d/	ح	/ħ/
ذ	/ð/	خ	/x/
ر	/r/	ع	/ʕ/
ز	/z/	غ	/ɣ/
س	/s/	ف	/f/
ش	/ʃ/	ق	/q/
ص	/sˤ/	ك	/k/
ض	/dˤ/	م	/m/
ط	/tˤ/	و	/w/
ظ	/ðˤ/	ي	/j/
ل	/l/	ه	/h/
ن	/n/	ء	/ʔ/

[Graf]ema; [fon]ema.

Tabela 13: O artigo definido e as letras solares e lunares.

árabe	letra	pronúncia
الشمس ←	ال + شمس	solar /al/ + /ʃams/ → [aʃ:ams]
البيت ←	ال + بيت	lunar /al/ + /bajt/ → [albajt]

⟨شمس⟩ /ʃams/ |xams| {sol}
 ⟨بيت⟩ /bajt/ |bayt| {casa}

2.2.5.2 Vogais longas, semivogais e suporte da hamza^t

De modo geral, há correspondência biunívoca entre as consoantes e seus respectivos grafemas. As vogais longas, por outro lado, possuem grafemas próprios (tab. 14), que podem assumir outras duas funções (tab. 15): grafar uma semivogal —somente ⟨و⟩ e ⟨ي⟩—; servir de apoio para ⟨ء⟩ — a ⟨همزة⟩ |hamza^t|, que representa o fonema /ʔ/.

Tabela 14: Vogais longas.

grafema	vogal	nome	
		árabe	romanizado
ا	/a:/	ألف	ʔalif
ي	/i:/	ياء	yāʔ
و	/u:/	واو	wāw

Tabela 15: Os grafemas ⟨ا⟩, ⟨ي⟩ e ⟨و⟩ nas funções de vogal longa, semivogal e suporte da hamza^t.

	ʔalif	yāʔ	wāw
grafema	ا	ي	و
vogal longa	/a:/	/i:/	/u:/
semivogal	—	/j/	/w/
suporte	أ إ	ئ ئ	ؤ ؤ

2.2.5.3 Marcas suplementares

A escrita árabe conta com diversas marcas suplementares, normalmente opcionais, chamadas conjuntamente de ⟨تشكيل⟩ |taxkīl|.

As breves são marcadas usando-se as ⟨حركات⟩ |ḥarakāt| (plural de ⟨حركة⟩ |ḥaraka^t|). Cada uma possui um nome (tab. 16) e pode ocupar uma

posição acima ou abaixo da consoante a que estiver associada (tab. 17). Essas marcas são facultativas, não sendo usadas exceto no al-ḵurʿān, em poesia e em material didático; ou para desfazer eventuais ambiguidades (*vide* 2.2.2.3, p. 66). Quando usadas no final de nomes, indicam o caso; havendo <تَنْوِين> |tanwīn| {nunação}, tem-se uma geminação ortográfica¹⁶, denotando indefinição (tab. 18).

Tabela 16: Vogais breves.

ḥaraka ^t	vogal	nome	
		árabe	romanizado
ـَ	/a/	فَتْحَة	fathā ^t
ـِ	/i/	كَسْرَة	kasra ^t
ـُ	/u/	ضَمَّة	ḍamma ^t

Tabela 17: Posição das ḥaraka^t.

		árabe	romanizado
بَ	←	ـَ + ب	b + a → ba
بِ	←	ـِ + ب	b + i → bi
بُ	←	ـُ + ب	b + u → bu

Outras marcas incluem <ـَـَ>, a <شَدَّة> |xadda^t|, que indica a geminação da consoante; e <ـَـْ>, o <سُكُون> |sukūn|, que sinaliza a ausência de vogal breve associada à consoante. As marcas usadas em conjunto com o <ل> ʿalif serão vistas em 2.2.7.3 (p. 77).

2.2.6 Cursividade

Em contraste com as formas quadradas usadas no hebraico, o abjad árabe adota uma forma cursiva, em que as letras, na sua maioria, são es-

¹⁶Dobra-se a marca, mas na pronúncia não se dobra a vogal; sufixa-se à vogal um /n/ — eis de onde vem o termo <nunação>.

Tabela 18: Marcação de caso em <بَيْت> |bayt| {casa}

caso	nome definido		nome indefinido	
	ḥaraka [†]	exemplos	tanwīn	exemplos
nominativo	ـُ	الْبَيْتُ ^q al-bayt ^u	ـٌ	بَيْتٌ bayt ^{un}
acusativo	ـَ	الْبَيْتَ ^q al-bayt ^a	ـِ	بَيْتًا bayt ^{an} *
genitivo	ـِ	الْبَيْتِ ^q al-bayt ⁱ	ـٍ	بَيْتٍ bayt ⁱⁿ

*O ^qalif tem valor meramente ortográfico.

critas ligadas umas às outras. Essa característica é partilhada com as escritas usadas no aramaico, no siríaco e no mandeu. Segundo HIGOUNET (2003), isso seria herança do nabateu, que possuía «ligaduras que ligam entre si a maior parte de suas letras, que podem tomar [...] formas independentes, iniciais, médias e finais diferentes.» (p. 78).

Isso faz com que seja necessário utilizar formas mais ou menos diferentes, conforme o grafema e sua ligação com grafemas adjacentes. Para tanto, há quatro formas: isolada; inicial (ligada à letra seguinte); medial (ligada às letras anterior e posterior); final (ligada à letra anterior). A maioria das letras apenas perde o *swash* nas formas inicial e medial, para que possam se ligar à letra seguinte, preservando apenas seu elemento identificador¹⁷, como ocorre com o <ب> bā^q e o <ح> ḥā^q. Um pouco poucas apresentam formas notadamente distintas, sendo a mais representativa o <و> hā^q. E, das 28 letras, 6 não se ligam à seguinte¹⁸, possuindo apenas formas isoladas e finais¹⁹, como o <د> |dāl| (tab. 19).

swash: um floreio, geralmente em forma de gancho, no final da letra.

¹⁷Trata-se da parte que, nas variantes, pouco ou nada se modifica.

¹⁸São elas: ^qalif <ا>; dāl <د>; ḏāl <ذ>; rā^q <ر>; zayn <ز>; wāw <و>.

¹⁹Nas tabelas que apresentam as variantes posicionais, é de praxe usar a forma isolada para representar também a inicial; e a final para a medial. Porém, no intuito de denotar algumas das características gráficas desses seis grafemas, neste trabalho adotamos a representação de BAUER (1996), que mostra somente as variantes inicial e medial que se ligam graficamente à letra seguinte.

Tabela 19: Diferentes formas cursivas, de acordo com a ligação a grafemas adjacentes.

formas				fonema	nome
final	medial	inicial	isolada		
ب	ـ	بـ	ب	b	bā ^q
ح	ـ	حـ	ح	ħ	ḥā ^q
هـ	ـ	هـ	هـ	h	hā ^q
د	ـ	ـ	د	d	dāl

2.2.7 Ortografia

2.2.7.1 e morfologia

A correspondência grafema–fonema do abjad árabe é quase totalmente biunívoca (*vide* 2.2.5.2, p. 70). Entretanto, ao contrário do que possa parecer, «a escrita árabe emprega alguns mecanismos para representar estruturas morfológicas mais profundas»²⁰ (BAUER, 1996).

Por exemplo, em relação ao gênero, «[normalmente], o feminino é formado a partir do masculino [...] com a adição do sufixo [-at].» (COWAN, 2007, pp. 12, 23). Esse sufixo é grafado com um grafema usado somente em posição final: <ة>, a <تاء مربوطة> |tā^q marbūṭa^t| {T amarrado}. Como o próprio nome sugere, trata-se de um <ت> |T| fechado; sua forma pode também ser entendida como um <ه> |H| final (هـ) acrescido dos dois pontos característicos do grafema <ت> tā^q.

2.2.7.2 e suas regras

Das regras de ortografia, uma é de especial interesse para nossa pesquisa: tanto o artigo definido quanto as partículas (e.g. preposições) compostas por apenas um grafema são grafados junto à palavra seguinte.

²⁰No original: «the Arabic script employs some devices to represent morphologically deeper structures.» (p. 561).

O artigo definido ⟨ال⟩ |al-| é quase sempre grafado em sua forma integral. Entretanto, quando com a preposição ⟨لِ⟩ |li| {para}, pode ocorrer supressão do ⟨ل⟩ ⁴alif ou, se o primeiro grafema seguinte ao artigo definido for o ⟨ل⟩ lām, de todo o artigo definido, ⟨ال⟩ |al-|, ortograficamente restando somente a indicação da assimilação (tab. 20).

E partículas como ⟨وَ⟩ |wa| {e} e ⟨بِ⟩ |bi| {com} —mas não ⟨فِي⟩ |fī| {em}, com dois grafemas— grafam-se junto à palavra seguinte. Caso ela esteja ligada ao artigo definido ⟨ال⟩ |al-|, a vogal será elidida, mas o grafema será mantido (tab. 21).

Tabela 20: Ortografia da partícula ⟨لِ⟩ |li| {para}.

árabe		romanizado		tradução
sintagma	acrécimo			
	رَجُل		rajul	homem
الرَّجُل ←	رَجُل + ال	artigo	ar-rajul	o homem
لِلرَّجُل ←	الرَّجُل + لِ	partícula	li-r-rajul ⁱ	para o homem
	لَيْل		layl	noite
الَّيْل ←	لَيْل + ال	artigo	al-layl	a noite
لِلَّيْلِ ←	الَّيْل + لِ	partícula	li-l-layl ⁱ	para a noite

Exemplos de COWAN (2007, p. 45).

Tabela 21: Ortografia das partículas ⟨وَ⟩ |wa| {e} e ⟨بِ⟩ |bi| {com}.

árabe		romanizado		tradução
sintagma	acrécimo			
	بَيْت		bayt	casa
الْبَيْت ←	بَيْت + ال	artigo	al-bayt	a casa
وَالْبَيْت ←	الْبَيْت + وَ	partícula	wa-'l-bayt	e a casa
بِالْبَيْت ←	الْبَيْت + بِ	partícula	bi-'l-bayt	com a casa

2.2.7.3 e o alif

O ⟨ا⟩ ^qalif, como visto anteriormente (2.2.5.2, p. 70), pode estar associado à vogal longa /a:/ ou servir de suporte ortográfico para a ⟨ء⟩ hamza^t (o grafema que representa a parada glotal /ʔ/). Na ortografia árabe, o ^qalif merece atenção especial em pelo menos mais cinco situações — as duas últimas referem-se ao ^qalif na posição final de algumas palavras.

A mais frequente ocorre com o artigo definido ⟨ال⟩ /ʔal/ |al-|. Na fala e na leitura fluentes, pode ocorrer elisão de /ʔa/, quando houver uma palavra que anteceda o artigo definido — além de alguns outros casos (COWAN, 2007). Essa elisão é marcada com ⟨َ⟩, a ⟨همزة الوصل⟩ |hamza^t al-waṣl| {hamza^t da junção}, formando ⟨أ⟩.

Além disso, em algumas palavras, pode ocorrer uma sequência de parada glotal e vogal longa /ʔa:/. O que poderia ser grafado como ⟨أأ⟩ dois ^qalif seguidos é substituído por uma ⟨مدة⟩ |madda^t| sobre apenas um ^qalif: ⟨آ⟩. Um exemplo bastante conhecido encontra-se em ⟨القرآن⟩ |al-ḡurʔān| {Alcorão}.

E, em uns poucos casos, a vogal longa /a:/ é grafada opcionalmente, sobrescrita: ⟨أَ⟩, o ⟨ألف خنجرية⟩ |^qalif ḡanjaryya^t|, como em ⟨الله⟩ |al-lāh| {Alá}.

Ele pode assumir a forma do ⟨ي⟩ yā^q, mas sem os pontos: ⟨ى⟩, o ⟨ألف مقصورة⟩ |^qalif maḡṣūra^t|. Assim grafado, quando houver a sequência de fonemas /ʔa:/, não se usa a madda^t, sendo a grafia ⟨أى⟩, como em ⟨رأى⟩ |raʔā| {[ele] viu}.

Ou pode aparecer após o final de algumas formas verbais no plural. Ele é usado, por exemplo, na terceira pessoa masculina do perfectivo: ⟨كُتِبُوا⟩ |katabū| {[eles] escreveram} (vide tab. 9). Nessa forma, o ^qalif tem valor meramente ortográfico, não estando diretamente associado a fonema algum.

CAPÍTULO 3

O português brasileiro e o alfabeto latino

Multum non multa.

«Qualidade, não quantidade.»

ditado latino

ESMIUÇAMOS de que maneiras um fonema pode ser grafado e que fonemas representa um dado grafema. E procuramos selecionar correspondências grafema–fonema que não gerem ambiguidades de produção e/ou interpretação. Uma tal biunivocidade será útil na elaboração do sistema de romanização, que se utilizará das correspondências aqui selecionadas.

3.1 A língua — fonemas

O português brasileiro (PB) possui um inventário de 33 fonemas, sendo 19 consoantes, 2 semivogais e 12 vogais. Para nossa proposta de transposição ortográfica (capítulo 5, páginas 122–159) interessam-nos mais as consoantes e semivogais do que as vogais, uma vez que o sistema vocálico do árabe padrão (AP) mostra-se relativamente mais simples (*vide* 2.2.2.3, página 66), tornando desnecessário pormenorizar todo o sistema vocálico do PB, para efeito de comparação.

3.1.1 Consoantes

As 19 consoantes do PB¹, no que tange ao modo como são realizadas em diferentes dialetos, podem ser divididas em três grupos: uniformes (sem variação de pronúncia); variáveis (com alofonia segundo o dialeto do falante); mistas (com variação de pronúncia em posições específicas).

São **uniformes**, em todos os dialetos do PB, as plosivas /p, b, k, g/, as fricativas /f, v/ e as nasais /m, n/. Via de regra, todas ocupam o ataque das sílabas em que aparecem².

Dentre as **variáveis**, as plosivas /t, d/ podem encontrar-se em distribuição complementar com [tʃ, dʒ], antes de /i/ ou /j/; a fricativa /H/³ apresenta alofones⁴, sendo alguns comuns a /R/ (tabela 22); as palatais /ɲ, ʎ/, segundo SILVA (2003), apresentam majoritariamente as realizações [j] (para a nasal) e [ʎ]⁵ (para a lateral), sendo possíveis as pronúncias [nʲ] e [j].

¹Subdivididas em seis plosivas /p, b, t, d, k, g/, sete fricativas /f, v, s, z, ʃ, ʒ, H/, três nasais /m, n, ɲ/ e três líquidas /R, l, ʎ/.

²Em palavras como <afta> /'af.ta/ e <apto> /'ap.to/, tende a ocorrer epêntese de [i], resultando nas realizações [a.fɨ.te] e [a.pɨ.tu].

³Para nossa análise, adotamos o arquifonema /H/ para transcrever a consoante que ocorre somente em início de sílaba (ataque simples), como em <rearranjo>.

⁴São eles: [h, x, r].

⁵Isso explica a homofonia existente entre <julho> e <Júlio>, ambos pronunciados [ʒu.ʎu].

E, no grupo das **mistas**, encontramos as fricativas /s, z, ʃ, ʒ/, que são uniformes no ataque das sílabas, podendo marcar variação dialetal quando em posição de coda; a rótica /R/⁶, que é uniforme em posição intervocálica e em ataques complexos, apresentando vários alofones em posição de coda⁷, sendo alguns comuns a /H/, (tabela 22); e a lateral /l/, que é uniforme em início de sílaba e em ataque complexo, mas que em posição de coda pode ser realizada [ɫ] ou, como na maioria dos dialetos do PB, [w]⁸.

Tabela 22: Alofones comuns entre /H/ e /R/.

arquifonema	grafia	realização	
		[h]	[x]
H (ataque)	rata	'ha.ta	'xa.ta
	marra	'ma.ha	'ma.xa
R (coda)	mar	'mah	'max
	carta	'kah.ta	'kax.ta

Cabe notar que a realização de /l/ como [w] parece não ter afetado o valor consonantal de ⟨l⟩ na coda de sílabas a que se segue um ⟨R⟩ —de forma similar ao que ocorre em ⟨Israel⟩ /is.Ha.'ɛl/—, visto que a pronúncia do ⟨R⟩, nesses contextos, permanece /H/ (tabela 23).

3.1.2 Semivogais

São duas as semivogais: /j/, representada pelo grafema ⟨i⟩; /w/, normalmente representada pelo grafema ⟨u⟩. Em alguns dialetos do PB, outros fonemas podem ser realizados como semivogais: a nasal palatal /ɲ/

⁶Para nossa análise, adotamos o arquifonema /R/ para transcrever o fonema que ocorre no ataque complexo (consoante + /r/), no ataque de sílabas não-iniciais e em posição de coda, como em ⟨prurir⟩.

⁷São eles: [ɾ, h, x, ɻ, ʎ, ɦ].

⁸Isso pode ser observado na homofonia de palavras como ⟨mau⟩ e ⟨mal⟩, ambas realizadas [maw].

Tabela 23: Pronúncia de R em ataque de sílaba não-inicial, precedido de L.

exemplo			contra-exemplo		
grafia	transcrição		grafia	transcrição	
	fonológica	fonética		fonológica	fonética
chilrear	ʃil.He.'aR	ʃiw.He.'aR	laurear	law.Re.'aR	law.Re.'aR
palrar	pal.'HaR	paw.'HaR	aura	'aw.Ra	'aw.Re

como [j] e a lateral palatal /ʎ/ como [j] (*vide* 3.1.1, p. 79); a lateral /l/ como [w], em posição de coda (*vide* 3.1.1, próxima página).

3.1.3 Vogais

As doze vogais dividem-se em sete orais⁹ e cinco nasais¹⁰.

As **orais** apresentam distribuição e variação em relação ao acento tônico (tabela 24), i-e, em posição tônica, todas as sete distinguem significado; pós-tônica medial e pré-tônica, apenas cinco constituem fonema; pós-tônica final, só três, em suas variantes reduzidas¹¹.

As **nasais** não apresentam essa variação, assim como não distinguem vogais média-altas e média-baixas (tab. 25). Não possuindo o árabe vogais nasais (*vide* 2.2.2.3, p. 66), não nos deteremos nesses fonemas.

⁹São elas: /i, e, ε/; /a/; /u, o, ɔ/.

¹⁰São elas: /ĩ, ĩ/; /ã/; /ũ, õ/.

¹¹Vogais reduzidas são alofones mais centralizados de algumas vogais ou de suas formas neutralizadas. Em outras palavras, as anteriores /i, e/ neutralizam-se em /ɪ/, que sofre redução para [ɪ]; a central /a/ sofre redução para [ə]; as posteriores /u, o/ neutralizam-se em /ʊ/, que sofre redução para [ʊ].

Tabela 24: Distribuição das vogais orais do português brasileiro, conforme sua tonicidade e posição relativa à sílaba tônica.

tonicidade	posição*	anteriores			centrais	posteriores		
tônica		/i/	/e/	/ɛ/	/a/	/u/	/o/	/ɔ/
átona	não-final	/i/	/e/		/a/	/u/	/o/	
	final	[ɪ]			[ɐ]	[ʊ]		

*Posição não-final: pré-tônica ou pós-tônica medial. Final: pós-tônica final.

Tabela 25: Vogais nasais do português brasileiro.

anteriores		centrais	posteriores	
/ĩ/	/ẽ/	/ã/	/ũ/	/õ/

3.2 A escrita — correspondências grafema–fonema

O PB emprega o alfabeto latino padrão de 26 letras¹², sendo três delas (κ, w e χ) usadas eminentemente para grafar estrangeirismos e seus derivados¹³. As consoantes do PB podem apresentar diferentes tipos de correspondência com os fonemas. Basicamente, trabalhamos com quatro categorias: nula; biunívoca; unívoca; múltipla.

padrão ou, ao menos, em sua forma mais difundida.

3.2.1 Nulas

Trata-se dos casos em que um elemento da fala —normalmente um fonema epentético, devido à sua própria natureza— não é registrado na forma escrita da palavra; e dos casos em que um grafema tem valor meramente ortográfico, não representando fonema algum.

No PB, por exemplo, o **fonema** epentético [j] pode ocorrer entre uma vogal e o /s/ em posição de coda, tornando homófonas as parônimas

¹²A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z.

¹³Por exemplo, ⟨Kafka⟩ e ⟨kafkiano⟩, ⟨Darwin⟩ e ⟨darwiniano⟩, ⟨Byron⟩ e ⟨byroniano⟩.

⟨mas⟩ e ⟨mais⟩, ambas pronunciadas [maj̃s]; o **grafema** que melhor representa esse tipo de correspondência é o **H inicial**, em palavras como ⟨**h**aver⟩, ⟨**h**erbicida⟩, ⟨**h**élio⟩, ⟨**h**ino⟩, ⟨**h**oje⟩, ⟨**h**omem⟩ e ⟨**h**umor⟩, além do utilizado em ⟨Bahia⟩ —mas ⟨baiano⟩— e ⟨Corinthians⟩ —mas ⟨corintiano⟩.

inicial, em oposição ao **H** formador dos dígrafos CH, LH e NH.

No AP, algo semelhante —ainda que não se trate exatamente do mesmo fenômeno— pode ser observado na marcação opcional das vogais breves (*vide* 2.2.2.3, p. 66) e em umas poucas instâncias do uso do ^aalif (*vide* 5.7.3, p. 151); e no uso do ^aalif no final de algumas formas verbais, com valor meramente ortográfico (*vide* 5.7.2, p. 150).

3.2.2 Biunívocas

Trata-se das correspondências exatas, de um para um, i-e, dado grafema —simples ou composto— representa única e exclusivamente um fonema, que é representado tão-somente por aquele grafema¹⁴.

No PB, encontram-se em alguns poucos casos. Os pares formados com **grafema simples** são ⟨P⟩–/p/, ⟨B⟩–/b/, ⟨T⟩–/t/, ⟨D⟩–/d/, ⟨F⟩–/f/¹⁵ e ⟨L⟩–/l/; os formados com **grafema composto** são ⟨NH⟩–/ɲ/ e ⟨LH⟩–/ʎ/ (tab. 34).

No AP, por outro lado, praticamente todos os grafemas e fonemas —salvo os associados às vogais longas e semivogais (*vide* 2.2.5.2, p. 70)— encontram-se nesta categoria.

3.2.3 Unívocas

Trata-se das correspondências em que um grafema corresponde a apenas um fonema, que pode ser representado também por outros grafe-

¹⁴A fim de não precisarmos escrever reiteradamente a mesma expressão, o termo ⟨grafema⟩ significará «grafema simples ou composto (dígrafo)» onde se puderem subentender ambas as possibilidades.

¹⁵O par ⟨v⟩–/v/ não pôde ser classificado como biunívoco devido ao uso, embora restrito, de ⟨w⟩–/v/. O fonema /v/ está classificado em 3.2.4 (p. 86).

Tabela 26: Correspondências biunívocas.

fonema		g·s	fonema		g·c
/p/	↔	P	/ɲ/	↔	NH
/b/	↔	B	/ʎ/	↔	LH
/t/	↔	T			
/d/	↔	D			
/f/	↔	F			
/l/	↔	L			

[G]rafema [s]imples; [g]rafema [c]omposto.

mas; e àquelas em que um fonema é representado por um só grafema, que pode representar ainda outros fonemas.

Os pares em que um **grafema simples** representa **um só fonema** são ⟨κ⟩–/k/, ⟨ç⟩–/s/, ⟨j⟩–/ʒ/, ⟨h⟩–/h/; os pares com **grafema composto** são ⟨sc, sç⟩–/s/¹⁶, ⟨ss⟩–/s/, ⟨xc⟩–/s/, ⟨ch⟩–/ʃ/, ⟨rr⟩–/R/ (tab. 27).

Tabela 27: Correspondências unívocas grafema–fonema.

fonema		g·s	fonema		g·c
k	←	K	s	←	SC/SÇ
s	←	ç	s	←	XC
ʒ	←	J	ʃ	←	CH
H	←	{H}	H	←	RR

[G]rafema [s]imples; [g]rafema [c]omposto.

As chaves { } indicam grafema usado em estrangeirismos.

¹⁶Dígrafos em distribuição complementar: sc, usado antes de ⟨e, i⟩; sç, antes de ⟨a, o, u⟩. Não confundir o grafema composto (dígrafo) sc com o encontro consonantal ⟨s+c⟩, usado em palavras como ⟨as.co⟩.

Os pares em que um **fonema** é representado por um só **grafema simples**¹⁷ são /f/–⟨x⟩ e /H/–⟨R⟩¹⁸; os pares com um **grafema composto** são /k/–⟨QU⟩ e /g/–⟨GU⟩ (tab. 28).

Tabela 28: Correspondências unívocas fonema–grafema.

fonema		g·s	fonema		g·c
f	→	x	k	→	QU
H	→	R	g	→	GU

[G]rafema [s]imples; [g]rafema [c]omposto.

Pode-se notar que alguns fonemas estabelecem relações mistas com seus respectivos grafemas e fonemas. Por exemplo, os fonemas /k, g/ são unívocos com seus respectivos grafemas e fonemas, que por sua vez estão em distribuição complementar entre si, i-e, /k, g/ são representados por ⟨c, g⟩ ou ⟨QU, GU⟩, dependendo da vogal seguinte: ⟨c, g⟩ antes de ⟨A, O, U⟩; ⟨QU, GU⟩ antes de ⟨E, I⟩. Além disso, os fonemas /f/ e /H/ estão em relação biunívoca com os dígrafos ⟨CH⟩ e ⟨RR⟩ e unívoca com os grafemas ⟨x⟩ e ⟨R⟩.

Um caso à parte, não trataremos do ⟨x⟩ como grafema do encontro consonantal /ks/, por se tratar de um dífono (um grafema, dois fonemas). Isso exclui também os encontros consonantais ⟨cc, cç⟩.

3.2.4 Múltiplas

Trata-se das correspondências em que um grafema representa mais de um fonema ou em que um fonema é representado por mais de um grafema.

¹⁷Diferentemente do que ocorre com as biunívocas, nas correspondências unívocas faz-se distinção entre ⟨grafema–fonema⟩ e ⟨fonema–grafema⟩, sendo o primeiro termo de cada par o que representa ou é representado univocamente.

¹⁸Somente em início de palavra.

Os **grafemas simples** são ⟨c⟩–/k, s/, ⟨g⟩–/g, ʒ/, ⟨r⟩–/H, R/, ⟨s⟩–/s, z/, ⟨z⟩–/z, s/, ⟨x⟩–/ʃ, s, z/. Os **grafemas compostos** são ⟨qu⟩–/k, kw/ e ⟨gu⟩–/g, gw/¹⁹ (tab. 29).

Tabela 29: Correspondências múltiplas — grafemas.

fonema		g·s	fonema		g·c
c, s	←	c	k, kw	←	qu
g, ʒ	←	g	g, gw	←	gu
H, R	←	r			
s, z	←	s			
z, s	←	z			
ʃ, z, s	←	x			

[G]rafema [s]imples; [g]rafema [c]omposto.

Os **fonemas representados por mais de um grafema simples** são /s/–⟨s, z, x, c/ç⟩, /z/–⟨z, s, x⟩, /v/–⟨v, w⟩, /j/–⟨i, y⟩, /w/–⟨u, w⟩. O **fonema representado por mais de um grafema composto** é /s/–⟨ss, sc/sç, xc⟩²⁰ (tab. 30).

Note-se haver sobreposição de alguns dos grafemas a representar /s, z/. Por motivos diversos, ⟨s, x, z⟩ podem representar um ou outro fonema: ⟨sapo, posa⟩; ⟨auxílio, exílio⟩; ⟨vez, vezes⟩.

¹⁹Na verdade, quando se representa também /w/ —especialmente com o Novo Acordo Ortográfico, que aboliu o uso do trema ⟨¨⟩, tornando obsoletas as formas ⟨qũ, gũ⟩—, os grafemas compostos passam a ser sequências de dois grafemas, um para a consoante e o seguinte para a semivogal.

²⁰Os dígrafos ⟨ss, sc/sç, xc⟩ só ocorrem em posição intervocálica, salvo raras exceções, como ⟨abscissa, obsceno⟩.

Tabela 30: Correspondências múltiplas — fonemas.

fonema		g·s	fonema		g·c
s	→	s, z, x, c/ç	s	→	ss, sc/sç, xc
z	→	z, s, x			
v	→	v, {w}			
j	→	j, {y}			
w	→	u, {w}			

[G]rafema [s]imples; [g]rafema [c]omposto.

As chaves { } indicam grafema usado em estrangeirismos.

3.2.5 Nulas, biunívocas, unívocas e múltiplas

Comparando os fonemas segundo os grafemas (simples e/ou compostos) que os representam (tab. 31), fica bastante nítida a variedade de grafias possíveis. Em termos de quantidade de variantes ortográficas, observamos: onze fonemas representados por apenas um grafema; seis fonemas, por dois grafemas; três, por três. Disso se distancia o /s/, representado por sete grafemas (quatro simples e três compostos).

3.3 Seleção de grafemas

Tendo analisado os fonemas e grafemas do português brasileiro, faremos uma seleção com base nos dados levantados em 3.1.1, 3.1.2 e 3.2 (pp. 79–87). Para nossa proposta, interessa-nos aproveitar as regularidades das correspondências grafema–fonema e evitar-lhes as exceções, para que a romanização (capítulo 5, pp. 122–159) seja tão simples quanto possível. Conseguindo selecionar somente correspondências biunívocas, teremos em mãos um inventário grafemático mais adequado²¹ para construir nosso sistema de transposição ortográfica.

²¹Mais adequado do que a ortografia do PB *per se* nos permitiria.

Tabela 31: Fonemas consonantais e os grafemas a representá-los.

plosivas e fricativas			nasais, líquidas e semivogais		
g·s	fon.	g·c	g·s	fon.	g·c
P	p		M	m	
B	b		N	n	
T	t			ɲ	NH
D	d				
{K}, C	k	QU	R	R	
G	g	GU	L	l	
				ʎ	LH
F	f				
{W}, V	v		{Y}, I	j	
C/Ç, X, Z, S	s	SS, SC/SÇ, XC	{W}, U	w	
X, S, Z	z				
X	ʃ	CH			
G, J	ʒ				
{H}, R	H	RR			

[G]rafema [s]imples; [fon]ema; [g]rafema [c]omposto.

As chaves { } indicam grafema usado em estrangeirismos.

3.3.1 Regras e ordem

A seleção será feita seguindo regras e obedecendo a uma ordem específica. As regras visam a auxiliar a seleção das correspondências mais adequadas e a solucionar eventuais impasses. A ordem presta-se a sequenciar a escolha dos grafemas, permitindo proceder às seleções, das mais simples para as mais complexas.

3.3.1.1 Regras

São seis as regras: de ataque; de isolamento; de simplicidade; de regularidade; de início; de conflito.

- **Regra de ataque** (regra geral). Tomem-se as correspondências grafema–fonema que ocorram em ataque de sílaba.
- **Regra de isolamento**. Tome-se a leitura esperada do grafema seguido de vogal (monossílabo cv).
- **Regra de simplicidade**. Tome-se preferencialmente o grafema simples, quando houver opções simples e compostas.
- **Regra de regularidade**. Tome-se a correspondência mais frequente e/ou regular, quando um grafema representar diferentes fonemas conforme a vogal seguinte. E atribua-se essa correspondência a todas as combinações cv.
- **Regra de início**. Tome-se o grafema que ocorra preferencialmente em:
 1. início de palavra (cv); ou
 2. posição intervocálica (vcv); ou
 3. ataque complexo (ccv).
- **Regra de conflito**. Havendo dois fonemas por um mesmo grafema representados, manter-se-á a correspondência do fonema que não puder associar-se a um grafema alternativo.

cv: sequência de [c]onsoante e [v]ogal.

3.3.1.2 Ordem

Primeiro selecionam-se as vogais. Depois, as consoantes, começando pelas biunívocas. Dentre as restantes, separam-se inicialmente as que estiverem em conflito com vogais; em seguida, as que estiverem em conflito entre si.

3.3.2 Etapas

Uma vez levantadas as correspondências grafema–fonema do PB (*vide* 3.2, pp. 82–87), seguiremos agora as quatro etapas da ordem proposta para a seleção de grafemas.

3.3.2.1 Primeira etapa: fonemas vocálicos

Antes de iniciarmos a análise e a seleção das correspondências envolvendo os fonemas consonantais, precisamos estabelecer os grafemas que representam os fonemas vocálicos. Não faremos modificação alguma às correspondências correntes no PB, exceto a de que elas devem excluir as semivogais /j, w/, que serão analisadas em 3.3.2.4 (p. 91). Portanto, as correspondências ficam sendo as mostradas na tab. 32.

Tabela 32: Os fonemas vocálicos e os grafemas correspondentes.

	anterior			central		posterior		
	fon.		graf.	fon.	graf.	fon.		graf.
alta	i	↔	I			u	↔	U
média	e, ε	↔	E			o, ɔ	↔	O
baixa				a	↔	A		

[Fon]ema; [graf]ema.

3.3.2.2 Intermezzo: fonemas consonantais

Vimos em 3.2 (pp. 82–87) que os fonemas consonantais do PB podem ser representados por grafemas simples e/ou compostos (tab. 33). Esses

dados — já em conformidade com a regra de ataque — serão usados nas próximas três etapas.

Tabela 33: Os fonemas consonantais e os grafemas (simples e/ou compostos) a representá-los.

plosivas			fricativas			nasais, líquidas e semivogais		
g·s	fon.	g·c	g·s	fon.	g·c	g·s	fon.	g·c
P	p		F	f		M	m	
B	b		{W}, V	v		N	n	
T	t		C, Ç, X, Z, S	s	SS, SC, SÇ, XC		ɲ	NH
D	d		X, S, Z	z		R	R	
{K}, C	k	QU	X	ʃ	CH	L	l	
G	g	GU	G, J	ʒ			ʎ	LH
			{H}, R	H	RR	{Y}, I	j	
						{W}, U	w	

[G]rafema [s]imples; [fon]ema; [g]rafema [c]omposto.

As chaves { } indicam grafema usado em estrangeirismos.

3.3.2.3 Segunda etapa: correspondências biunívocas

Dentre as correspondências que já sejam biunívocas (tab. 34), encontramos quatro plosivas /p, b, t, d/, uma fricativa /f/, as três nasais /m, n, ɲ/ e duas líquidas /l, ʎ/.

3.3.2.4 Terceira etapa: conflito com vogal

Uma vez separadas as correspondências inerentemente biunívocas, passamos às que podem ser representadas por um grafema que represente também uma vogal (tab. 35).

Os fonemas assim selecionados são as semivogais /j, w/, que em PB podem aparecer grafadas com ⟨i, u⟩ ou, no caso de nomes próprios e palavras estrangeiras, ⟨y, w⟩. No entanto, os grafemas ⟨i, u⟩ já foram selecionados — correspondem às vogais /i, u/ (vide tab. 32, p. 90).

Tabela 34: Seleção das correspondências biunívocas.

plosivas			fricativas			nasais, líquidas e semivogais		
fon.		graf.	fon.		graf.	fon.		graf.
p	↔	P	f	↔	F	m	↔	M
b	↔	B	v			n	↔	N
t	↔	T	s			ɲ	↔	NH
d	↔	D	z			R		
k			ʃ			l	↔	L
g			ʒ			ʎ	↔	LH
			H			j		
						w		

[Fon]emas; [graf]emas.

Aplicando-se a regra de conflito, mantemos ⟨I, U⟩ para as vogais e selecionamos ⟨Y⟩ para /j/ e ⟨W⟩ para /w/.

Indiretamente, isso resolve a seleção do grafema de /v/, que, analogamente ao que se observa com /j, w/, pode ser grafado com ⟨v, w⟩. Uma vez tendo sido selecionado ⟨w⟩ para /w/, resta apenas ⟨v⟩ para /v/.

3.3.2.5 Quarta etapa: conflito com consoante

Por fim, os oito fonemas restantes são representados por grafemas que apresentam conflitos com outras consoantes (entre si). Podemos separá-los em três grupos (tab. 36), segundo o tipo de conflito: /k, s, z, ʃ/, em ⟨C, S, Z, X⟩ (misto); /ʒ, g/, em ⟨G⟩; e /H, R/, em ⟨R⟩.

Os casos mais simples de resolver são o de /H, R/ e o de /g, ʒ/. Aplicando a regra de conflito, selecionamos ⟨R⟩ para /R/, deixando /H/ com ⟨H, RR⟩; pelas regras de isolamento e de início, eliminamos ⟨RR⟩, restando o ⟨H⟩. A regra de conflito também permite selecionar ⟨J⟩ para /ʒ/ e deixar ⟨G⟩ para /g/; as regras de simplicidade e de regularidade orientam-nos a preferir ⟨GU⟩ e a selecionar o ⟨G⟩, mantendo-lhe a correspondência com /g/ independentemente da vogal que o siga.

Tabela 35: Seleção das correspondências em conflito com vogais.

plosivas		fricativas		líquidas e semivogais	
fon.	graf.	fon.	graf.	fon.	graf.
k		v	($\longleftrightarrow$)	R	
g		s		j	$\longleftrightarrow$ {Y}
		z		w	$\longleftrightarrow$ {W}
		ʃ			
		ʒ			
		H			

[Fon]emas; [graf]emas.

As chaves { } indicam grafema usado em estrangeirismos.

Os parênteses () indicam seleção indireta.

Os conflitos de /k, s, z, ʃ/ são mais entremeados. Mas podemos resolvê-los em quatro passos. No primeiro, aplicamos as regras de conflito, de isolamento e de simplicidade para preterir <CH> e selecionar <x> para /ʃ/. No segundo, as regras de isolamento e de início permitem-nos selecionar z para /z/. No terceiro, as de simplicidade, de isolamento e de início reduzem as possibilidades ao <s>, para representar /s/. No quarto e último, resta-nos aplicar as regras de simplicidade e de regularidade para selecionar <κ> para /k/.

3.3.3 Os grafemas selecionados

Terminada a seleção de grafemas, constatamos ter atingido nosso objetivo de selecionar e estabelecer correspondências biunívocas para os fonemas do PB. Essa seleção —que, acreditamos, constitui-se dos grafemas a melhor representar seus respectivos fonemas— abrange tanto as vogais orais (tab. 32, p. 90) quanto as consoantes (tab. 37).

Tabela 36: Seleção das correspondências em conflito com consoantes.

grupo 1			grupo 2			grupo 3		
/H, R/			/ʒ, g/			/k, s, z, ʃ/		
fon.		graf.	fon.		graf.	fon.		graf.
H	↔	H	ʒ	↔	J	k	↔	K
R	↔	R	g	↔	G	s	↔	S
						z	↔	Z
						ʃ	↔	X

[Fon]emas; [graf]emas.

Tabela 37: Grafemas selecionados.

plosivas			fricativas			nasais, líquidas e semivogais		
fon.		graf.	fon.		graf.	fon.		graf.
p	↔	P	f	↔	F	m	↔	M
b	↔	B	v	↔	V	n	↔	N
t	↔	T	s	↔	S	ɲ	↔	NH
d	↔	D	z	↔	Z	R	↔	R
k	↔	K	ʃ	↔	X	l	↔	L
g	↔	G	ʒ	↔	J	ʎ	↔	LH
			H	↔	H	j	↔	Y
						w	↔	W

[Fon]emas; [graf]emas.

Parte II

Da tradução da escrita

CAPÍTULO 4

A romanização do árabe

Graecum est non legitur.

«Pra mim, isso é grego.»

ditado latino

VERIFICAMOS, primeiro, como a imprensa escrita brasileira aborda a questão da grafia de nomes estrangeiros — em especial os de origem árabe. Em seguida, analisamos como dicionários etimológicos ou bilingües e uma proposta específica para o meio acadêmico procedem à transposição ortográfica do árabe padrão para o português brasileiro.

4.1 nos jornais

Muitos são os sistemas de romanização do árabe. Eles costumam variar segundo a época em que foram criados, a língua de chegada e o propósito ou público-alvo a que se destinem. Como vimos em 1.1.2 (página 37) e 1.1.4.2 (páginas 41–43), dois jornais de um mesmo estado (São Paulo) podem grafar um mesmo nome próprio de formas distintas.

4.1.1 Diferentes formas de romanizar

Um breve piloto¹ fora realizado para verificar a viabilidade da pesquisa, em seu formato original. Ainda que o projeto tenha tomado rumo divergente, e apesar d’o levantamento do corpus inicialmente proposto ter-se mostrado pouco frutífero, os dados coletados permitem analisar alguns pontos relevantes.

Primeiramente, percebemos não haver identidade de grafia na escrita de antropônimos árabes (tabela 38). Cada um dos três jornais grafou de uma forma diferente os nomes apresentados. Isso pode ser atribuído a regras e orientações internas próprias, a agências de notícias (fontes) distintas ou a carência de conhecimentos que permitiriam aplicar efetivamente as regras e orientações ao material recebido das agências de notícias.

Tabela 38: Nomes próprios árabes romanizados por três jornais.

Folha	OESP	O Globo
Al Qaeda	Al-Qaeda	al-Qaeda
Mahmoud Abbas	Mahmud Abbas	—
Mummar Gaddafi	Muamar Kadafi	—

[Folha] de São Paulo; [OESP] O Estado de São Paulo.

¹Fontes: O Estado de São Paulo (2007-07-23 a 27); Folha de São Paulo (2007-07-23 a 27); O Globo (2007-07-23).

Outro aspecto interessante revelam-nos os exemplos extraídos. Pela grafia consistente —internamente a cada jornal—, torna-se claro que o artigo definido ⟨ال⟩ |al-| é de fácil reconhecimento (tab. 39). Particularidades dialetais aparentemente passam despercebidas —algo evidente em «Sharm el-Sheikh»—, mas o quadro geral demonstra ser o registro uniforme.

Tabela 39: O artigo definido romanizado por três jornais.

Folha	OESP	O Globo
Al Qaeda	Al-Qaeda	al-Qaeda
Abu Musab al Zarqawi	Al-Watan	—
Seif al Islam	Al-Asbah	—
—	Sharm el-Sheikh	—
—	Nuri al-Maliki	—
—	Qassim al-Moussawi	—

[Folha] de São Paulo; [OESP] O Estado de São Paulo.

4.1.2 Manuais de redação e estilo

Os manuais de redação e estilo editados pela mídia impressa —dos quais consultamos os impressos da Folha de São Paulo (Folha), d'O Estado de São Paulo e d'O Globo, e o *online* da Folha— diferem no tocante ao tratamento que deva ser dispensado às palavras cuja grafia original baseie-se em um sistema de escrita diferente do que utilizamos. Mas, em geral, raras são as ocasiões em que se atenta para fatores linguísticos que envolvam distinções estranhas ao português. Por exemplo: em mandarim ou cantonês, os tons; em coreano, as consoantes fracas e as plosivas aspiradas ou as enfáticas (termo distinto de ⟨faringalizada⟩); em árabe, as consoantes guturais e as faringalizadas (termo distinto de ⟨enfática⟩). Pouco ou nada se considera a respeito dessas particularidades fonêmicas, sendo todas elas invariavelmente reduzidas a uma mesma e comum vogal —no caso dos tons— ou consoante —no caso dos modos e pontos de articulação.

4.1.3 Manual *online* da Folha de São Paulo — nomes árabes

Para conhecermos melhor o que rege as decisões tomadas nos casos em questão, vejamos o manual *online* de redação e estilo da Folha de São Paulo², do qual constam as orientações para a grafia de nomes de diversas línguas. A seguir reproduzimos a íntegra do texto referente ao árabe; na sequência (4.1.5, p. 101), comentaremos as orientações apresentadas.

árabes — Embora as transliterações enviadas pelas agências de notícias sejam em geral corretas, elas costumam trazer dificuldades para o leitor brasileiro sem conhecimentos linguísticos. Por exemplo, o W, que representa o som U em posição semivocálica, pode ser lido por um brasileiro como um V, o que foge completamente à sonoridade original. Não é raro ouvir, mesmo na TV, a pronúncia *cuvait* para *Kuait*. Assim, a Folha faz algumas adaptações nas transliterações enviadas pelas agências.

Substitua os YY e WW por II e UU, respectivamente, exceto quando a combinação resultar numa vogal geminada. Se isso ocorrer, elimine uma delas: *Kuwait*=*Kuuait*=*Kuait*. Substitua o SH por CH: *Shatt al Arab*=*Chatt al Arab*.

Mantenha o H em qualquer situação que ele apareça. Embora esta letra possa causar alguma confusão na pronúncia para um brasileiro, ela marca na transliteração do árabe uma grande quantidade de nuances consonantais, cuja eliminação tornaria qualquer palavra irreconhecível para quem conheça a língua. Por exemplo, indica a diferença entre as letras *kha* (transliterada por KH) e a letra *kaf* (transliterada por um simples K).

No árabe há grande número de sons guturais que devem ser diferenciados. Isso exige fugir um pouco às regras da língua portuguesa. Por exemplo, use a letra Q sem o U que necessariamente a segue em português: *Qatar* e não *Catar*. Se de fato houver um U após o Q, observe se não se lhe segue um I ou U, caso em que será necessário o trema.

Cuidado com as letras S e G, que, em árabe, têm sempre o mesmo som. Isso exige que, em português, o S seja dobrado em posição intervocálica e o G seja seguido de U antes de E e I.

²http://www1.folha.uol.com.br/folha/circulo/manual_anexo_estrang_a.htm

Último acesso em 2011-07-25.

Não elimine as duplas consoantes, que, embora não signifiquem nada em português, explicam uma série de assimilações e dissimilações importantes para o conhecedor do árabe. De qualquer forma, a dupla consoante não chega a atrapalhar a leitura em português. A rigor, não existem as vogais E e O em árabe. Elas aparecem em alguns falares regionais e não deveriam ser escritas. Entretanto, as agências internacionais costumam trazer a pronúncia dos nomes de acordo com essas variações regionais. Embora o correto seja substituí-las pela vogal escrita (I, U ou A, dependendo do caso), isso é muito pouco prático num jornal. Mantenha, portanto, essas vogais: *lasser Arafat*. Cuidado com o artigo *al*. Em nomes próprios, nunca o elimine porque ele pode alterar o sentido do nome. Grafe-o em minúscula e sem hífen antes do nome que o segue: *Anuar al Sadat*. Se não for usado o primeiro nome —em títulos ou na segunda menção—, o *a* de *al* deve ser grafado em maiúscula. Se em português ocorrer artigo antes de termo árabe com *al*, suprima o artigo árabe para evitar redundância: *O líder da Fatah* e não *O líder da al Fatah*. Isso não vale para palavras portuguesas de origem árabe: o algodão, o álcool, o Alcorão. Lembre-se de que essas regras não valem para muitos topônimos de origem árabe, que têm seus nomes consagrados de outra forma em português: Argélia, Marrocos, Egito, Emirados Árabes Unidos. Lembre-se também de que árabes que vivem no Brasil ou em outros países ocidentais têm uma forma já latinizada de escrever seus nomes, respeite-a, mesmo que de encontro a essas regras.

* Use GU antes do E e I.

** [...]³

*** User (sic) SS em posição intervocálica.

4.1.4 Comentários

No primeiro parágrafo explicitam-se serem as agências de notícias (e.g. Reuters) a principal fonte da Folha de São Paulo e haver preocupação com a recepção desse tipo de informação pelo leitor brasileiro.

Nos quatro parágrafos seguintes e no início do último, para as adaptações mencionadas nas linhas 7–8 estabelecem-se alguns parâmetros:

³Observação suprimida por se tratar de comentário pertinente à língua russa.

deve-se substituir ⟨Y⟩ por ⟨I⟩, ⟨W⟩ por ⟨U⟩ (l. 9) e ⟨SH⟩ por ⟨CH⟩ (l. 11–12); manter o ⟨H⟩ onde quer que apareça (l. 13), especialmente quando formar dígrafos (l. 18–19). E não deixa de mencionar as consoantes guturais (l. 20) e as geminadas (l. 29–33).

O manual trata, ainda, das vogais ⟨E, O⟩ (l. 33–39). Considera o paradigma árabe, que trabalha com as vogais ⟨A, I, U⟩, mas faz ressalvas a dialetos.

O artigo ⟨ال⟩ |al-| (l. 39–48), quando associado a um nome próprio, não deve ser suprimido, pois «pode alterar o sentido do nome» (l. 40–41), mas deve sê-lo, caso se use o artigo definido do português (l. 44–46), exceto no caso de palavras que já fazem parte do léxico português (l. 46–48).

Pouco se explicam as escolhas feitas ao elaborar as orientações contidas no manual. Além disso, priorizam-se a conveniência da leitura (l. 4–6) e a praticidade da grafia (l. 38); excluem-se das orientações «muitos topônimos [...] consagrados de outra forma em português» (l. 48–51) e antropônimos de pessoas que já vivam «no Brasil ou em outros países ocidentais [e que] têm uma forma já latinizada de escrever seus nomes» (l. 51–54).

4.1.5 Análise

Algumas soluções apresentam-se como eminentemente *ad hoc*, i.e., prestam-se a um caso específico, desconsiderando-se situações análogas. Percebe-se isso quando o texto indica que se deve grafar o ⟨Q⟩ sem o ⟨U⟩ (l. 22–25) —contrariando a ortografia vigente no Brasil—, mas, contraditoriamente, que se deve grafar o ⟨G⟩ com o ⟨U⟩ (l. 26–28) —seguindo-a.

O mesmo vale para as orientações de grafar o ⟨s⟩ dobrado, em posição intervocálica (l. 26–28) —novamente em acordo com a ortografia vigente—, e de «não [eliminar] as duplas consoantes» (l. 29), uma vez que «explicam uma série de assimilações e dissimilações importantes para o conhecedor do árabe» (l. 30–31); mas não se considera como seria feita a distinção entre /s/ e /s:/, posto que, seguindo-se o manual,

ambos seriam grafados da mesma forma.

Em suma, há um misto de soluções que ora aproximam, (e.g, registro de consoantes dobradas) ora afastam (e.g, uso de ⟨ss⟩ intervocálico) a grafia final do original em árabe; e ora subvertem (e.g, ⟨Q⟩ sem ⟨U⟩ posposto), ora preservam (e.g, ⟨G⟩ com ⟨U⟩ posposto) princípios ortográficos do português.

4.1.6 O chefe de Estado da Líbia (outra vez)

No jornal Folha de São Paulo, o nome do já citado chefe de Estado da Líbia aparece grafado ⟨Muammar Gaddafi⟩ (vide 1.1.2, p. 37; 1.1.4.2, p. 41; tab. 38, p. 97). Com base no manual *online* da Folha (4.1.3, pp. 99–100), um breve e rápido cotejo dessa forma romanizada com a grafia em árabe ⟨مُعَمَّر الْقَذَافِي⟩ permite notar que:

1. As consoantes dobradas ⟨MM, DD⟩ foram grafadas como tais;
2. Não houve distinção de quantidade vocálica, como se poderia esperar — nenhuma menção às vogais longas é feita no manual.

mas também que:

1. A consoante ⟨ع⟩ /ʕ/ de /muʕam:ar/ não foi registrada, fazendo crer que somente o |al-| não pode ser eliminado, «porque ele pode alterar o sentido do nome» (l. 40–41), e que somente assimilações e dissimilações estão entre as distinções «importantes para o conhecedor do árabe» (l. 31–31); mas
2. O artigo não foi registrado, contrariando a orientação de que ele nunca deve ser eliminado de nomes próprios;
3. A consoante ⟨ق⟩ /q/ foi grafada com ⟨G⟩, provavelmente devido à grafia adotada por uma das agências de notícias, que supostamente enviam «transliterações [...] em geral corretas»⁴ (l. 2);

⁴Essa afirmação parece desconsiderar a relatividade do conceito de correção, no que diz respeito a línguas e seus sistemas — fonético-fonológico e ortográfico, principalmente.

4. Não foi marcada a distinção entre ⟨D⟩ e ⟨DH⟩, uma que «tornaria qualquer palavra irreconhecível para quem conheça a língua» (l. 16–17), não diferente do exemplo de ⟨κ⟩ e ⟨KH⟩ (l. 18–19).

Um trabalho de coleta e análise como esse, se estendido para mais veículos de comunicação, certamente produziria resultados semelhantes.

4.2 na academia

No meio acadêmico, a romanização de palavras, termos e nomes próprios árabes faz-se necessária em diversificadas instâncias. Ela está presente quer seja na pesquisa etimológica, quer seja na histórica ou geopolítica, passando pelo ensino de língua, literatura e cultura, bem como pela análise contrastiva entre a língua árabe e a portuguesa.

4.2.1 Esquemas e sistemas de romanização

Podem-se realizar romanizações, basicamente, por meio de dois recursos: esquema; sistema.

Os **esquemas de romanização** resumem-se a uma tabela que apresenta as correspondências entre o fonema da língua de partida e o grafema da língua de chegada; são bastante simples e nem sempre apresentam os grafemas da língua de partida; podem não trazer transcrição fonológica, preferindo-se uma descrição aproximada do fonemas em questão.

Um **sistema de romanização** abrange os usos da transposição ortográfica e quais elementos serão ou não representados, i-e, quem são os usuários (produtores e receptores) e em que contexto (tipo de publicação), bem como quais os níveis de exatidão desejados na representação, por exemplo, de fenômenos fonéticos ou fonológicos (e-g, assimilação) ou de regras ortográficas (e-g, grafia de clíticos).

A seguir, veremos dois esquemas e seis sistemas de romanização. Os esquemas são os de CUNHA (1987) e o de NASR (2005). Os sistemas são os

de MACHADO (1990), NIMER (2005), SABBAGH (1988) e o da UNESP, além dos dois sistemas (leigo e acadêmico) de JUBRAN (2004).

4.2.2 O esquema de Cunha

Em seu dicionário etimológico, CUNHA (1987) apresenta-nos uma tabela de «símbolos e valores fonéticos». No entanto, não se fazem distinções claras sobre quais símbolos são usadas na romanização de quais idiomas; nem exatamente qual seria o fonema sendo representado. Na tabela 40 procuramos organizar os símbolos que julgamos corresponder a alguns dos grafemas e fonemas árabes.

Dois símbolos não pudemos identificar:

- o ⟨z̤⟩, apesar de descrito como «j do ing[lês] *jail* (= ğ), porém mais enfático», pode se referir ao ⟨ظ⟩ /ðˤ/;
- os três símbolos ⟨‘ ‘ ’⟩, com a descrição de que «no árabe, indica forte aspiração», podem estar relacionados ao ⟨ξ⟩ /ʕ/, mas não se explica a distinção entre as variantes — o que também se observa em ⟨ṭ, t̤⟩ e ⟨ḍ, ḍ̤⟩.

4.2.3 O esquema de Nasr

Em seu dicionário árabe–português, NASR (2005) emprega dígrafos e diacríticos (tab. 41). Sua tabela de «Transliteração das palavras árabes» não indica como representar o ⟨ʾ⟩, apenas descrevendo-o como «[a] primeira letra do alfabeto árabe», e informando que «[é] considerada como vogal longa». Digno de nota é seu uso do grafema primitivo ⟨z⟩ — sem e com diacríticos — em três casos: ⟨ز⟩ /z/ |Z|; ⟨ذ⟩ /ð/ |Z|; ⟨ظ⟩ /ðˤ/ |Z|.

4.2.4 O sistema de Machado

Em seu dicionário etimológico, MACHADO (1990) emprega grafemas compostos (dígrafos) e derivados (com diacríticos) — a exemplo de Nasr (4.2.3, p. 104) — e faz distinção entre caixa-alta e caixa-baixa:

Tabela 40: O esquema de romanização de CUNHA (1987).

grafemas solares			grafemas lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ت	t	—	ء	ʔ	—	ا	a:	—
ث	θ	t̤, t̤	ب	b	—	ة	—	—
د	d	—	ج	ʒ	ž	إ	a	—
ذ	ð	ḏ, ḏ	ح	ħ	ḥ	ـ	i	—
ر	r	—	خ	x	—	أ	u	—
ز	z	—	ع	ʕ	—	أ	an	—
س	s	—	غ	ɣ	g̣	أ	in	—
ش	ʃ	š	ف	f	—	أ	un	—
ص	ṣ	ṣ	ق	q	ḳ			
ض	ḍ	ḍ	ك	k	—			
ط	ṭ	ṭ	م	m	—			
ظ	ð̣	—	ه	h	ḥ			
ل	l	—	و	u:/w	u			
ن	n	—	ي	i:/j	i			

[Graf]ema árabe; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semi-vogal.

Tabela 41: O esquema de romanização de NASR (2005).

grafemas solares			grafemas lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ت	t	T	ء	ʔ	A inicial	ا	a:	—
ث	θ	Th	ب	b	B	ة		—
د	d	D	ج	ʒ	J	ـَ	a	—
ذ	ð	Z	ح	ħ	H	ـِ	i	—
ر	r	R	خ	x	Kh	ـُ	u	—
ز	z	Z	ع	ʕ	ʕ	ـِـ	an	—
س	s	S	غ	ɣ	Gh	ـِـ	in	—
ش	ʃ	Ch	ف	f	F	ـِـ	un	—
ص	sʕ	Ṣ	ق	q	Q			
ض	dʕ	Ḍ	ك	k	K			
ط	tʕ	Ṭ	م	m	M			
ظ	ðʕ	Z	ه	h	H			
ل	l	L	و	u:/w	U			
ن	n	N	ي	i:/j	I			

[Graf]ema árabe; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semivogal.

grafemas simples: t; d; r; z; s; x; l; n; b; j; h; g; f; q; m; u; i; a.

grafemas compostos: th; dh; kh.

grafema derivado: ç.

outro: ⟨'⟩.

caixa-alta: D; T; Z; K; H.

E trabalha com mais do que apenas uma tabela de correspondências (tab. 42). Ainda que não seja propriamente um sistema, em «Nota sobre o método de transcrição do alfabeto arábico», o autor apresenta algumas observações, das quais destacamos:

4. — O tá marbutto de palavra isolada não será reproduzido independentemente, mas o *a* da sílaba que o precede será *â*: *madînâ*.

Quando em vocábulo seguido de outro iniciado por vogal, à qual aquele fonema se deve ligar foneticamente, será *-t-*: *madîna-t-an-nabbî*,

5. — Com o artigo definido arábico (*al-*) temos de admitir três casos:

a. — seguido de consoante lunar, será *al-*: *al-qaçr*;

b. — seguido de consoante solar, terá o *l* modificado em consoante igual à inicial do vocábulo que precede: *az-zait*;

c. — quando precedido de vocábulo terminado em vogal, que elimina o *a* do artigo, limita-se a ser *-l-*, o que obriga à representação daquela: *çoqu-l-kbîr*; *'abdu-l-naçr*.

A partir do segundo exemplo de (5.c), pode-se inferir que Machado não distingue ⟨ع⟩ /ʕ/ —que não consta na tabela de correspondências— de ⟨ء⟩ /ʔ/, pois usa-lhe o mesmo sinal ⟨'⟩ para grafar ⟨'abd-l-naçr⟩⁵ —provavelmente de ⟨عبد النصر⟩.

⁵Além disso, esse exemplo contraria a observação imediatamente anterior, de que «o artigo definido arábico (*al-*) [...], seguido de consoante solar, terá o *l* modificado em consoante igual à inicial do vocábulo que precede»; a romanização deveria ter sido ⟨'abd-n-naçr⟩.

Tabela 42: O sistema de romanização de MACHADO (1990).

grafemas solares			grafemas lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ت	t	t	ء	ʔ	'	ا	a:	a
ث	θ	th	ب	b	b	ة		—
د	d	d	ج	ʒ	j	ـ	a	—
ذ	ð	dh	ح	ħ	h	ـ	i	—
ر	r	r	خ	x	kh	ـ	u	—
ز	z	z	ع	ʕ	—	ـ	an	—
س	s	s	غ	ɣ	g	ـ	in	—
ش	ʃ	x	ف	f	f	ـ	un	—
ص	sʕ	ç	ق	q	q			
ض	dʕ	D	ك	k	K			
ط	tʕ	T	م	m	m			
ظ	ðʕ	Z	ه	h	H			
ل	l	l	و	u:/w	u			
ن	n	n	ي	i:/j	i			

[Graf]ema árabe; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semi-vogal.

4.2.5 O sistema de Nimer

Em seu dicionário etimológico, NIMER (2005) trabalha com um sistema de romanização bastante pormenorizado. Apesar de não fazer distinção entre ⟨ء⟩ e ⟨ل⟩ —corresponde ⟨ل⟩ a ⟨'⟩— é dos dois únicos autores a explicitamente apresentar a romanização das vogais breves —o outro sendo Sabbagh (4.2.6, p. 110) — e o único a distinguir o uso de ⟨و⟩ e ⟨ي⟩ no início ou no fim de uma sílaba (tab. 43).

Tabela 43: O sistema de romanização de NIMER (2005).

grafemas solares			grafemas lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ت	t	t	ء	ʔ	'	ا	a:	'
ث	θ	ṭ	ب	b	b	ة		-ah ^t
د	d	d	ج	ʒ	j	ـ	a	a
ذ	ð	ḏ	ح	ħ	ḥ	ـ	i	i
ر	r	r	خ	x	ḵ	ـ	u	u
ز	z	z	ع	ʕ	ʕ	ـ	an	—
س	s	s	غ	ɣ	ḡ	ـ	in	—
ش	ʃ	š	ف	f	f	ـ	un	—
ص	s ^ʕ	ṣ	ق	q	q			
ض	d ^ʕ	ḍ	ك	k	k			
ط	t ^ʕ	ṭ	م	m	m			
ظ	ð ^ʕ	ẓ	ه	h	h			
ل	l	l	و	u:/w	w/ú			
ن	n	n	ي	i:/j	í/í			

[Graf]ema árabe; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semi-vogal. Onde houver duas romanizações, a primeira refere-se à utilizada em ataque de sílaba; a segunda, em coda.

O sistema, que não deixa de contemplar a assimilação de ⟨ال⟩ |al-|, conta, ainda, com outras distinções:

- usa ⟨ǎ⟩, quando o ⟨ل⟩ final não representar /a:/;
- usa ⟨ǎ⟩, quando a palavra árabe terminar em ⟨ى⟩, o ⟨ألف مقصورة⟩ |ʔalif maḵṣūra^t|;

- usa ⟨i^é⟩ para representar o sufixo ⟨ي⟩;
- distingue ⟨ṣ⟩ ⟨ah^t⟩ de ⟨ṣl⟩ ⟨āṭ⟩.

Dois outros pontos de destaque são o uso de grafemas simples (primitivos e derivados), evitando-se os grafemas compostos; e o uso, na romanização, de ⟨'⟩ para ⟨ξ⟩ e de ⟨'⟩ para ⟨ء⟩.

4.2.6 O sistema de Sabbagh

Em seu dicionário árabe-português-árabe, SABBAGH (1988) emprega uma romanização que, na mesma linha da de Nimer (4.2.5, p. 109), procura adotar somente grafemas simples (tab. 44).

E o sistema indica explicitamente como devem ser realizadas: a romanização das vogais breves; a representação da assimilação; a indicação de diversos sinais (e.g. waṣla^t, madda^t, xadda^t); a interpretação de caracteres especiais (e.g. tā^q marbūṭa^t, ^qalif maḵṣūra^t); a localização do acento tônico.

Uma observação deve ser feita sobre a romanização de ⟨ġ⟩ /y/. Uma vez que o autor somente indicou —e usa— a forma em caixa-alta ⟨Ḡ⟩, resta a dúvida acerca de como grafar esse grafema em caixa-baixa⁶. Em outras palavras, não se sabe se o ponto deve ser registrado sobre a descendente ⟨g⟩, abaixo dela ⟨ḡ⟩ ou acima da letra ⟨ḡ̇⟩.

Essa questão, que pode parecer trivial, na verdade deriva da observação de que um diacrítico colocado sob a caixa-alta pode ser colocado sobre a caixa-baixa, como ocorre com a vírgula inferior:

vírgula inferior [...] Em virtude de sua descendente, o g palatal minúsculo letão (ḡ) é indicado por uma forma invertida e elevada desse diacrítico.

vírgula inferior invertida Este sinal diacrítico é uma variação da vírgula inferior utilizada na forma minúscula do g suave letão (ḡ). A forma maiúscula dessa letra é Ḡ. (O *g* é a única letra

⁶Caso similar pode ser visto em 5.2.5 (pp. 129–130), a respeito da forma em caixa-baixa de ⟨Ḍ⟩.

palatalizada do letão que tem uma descendente). (BRINGHURST, 2005)

Outro indício de não ser o posicionamento do diacrítico trivial, o par ⟨G, g⟩ —um G riscado— encontra uso em *skolt sami* (Lapônia), supostamente para registrar /ɣ/⁷.

Tabela 44: O sistema de romanização de SABBAGH (1988).

grafemas solares			grafemas lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ت	t	T	ء	ʔ	—	ا	a:	A
ث	θ	Ṭ	ب	b	B	ة		A (T)
د	d	D	ج	ʒ	J	ـ	a	—
ذ	ð	Ḍ	ح	ħ	Ḥ	ي	i	—
ر	r	R	خ	x	Ḳ	ـ	u	—
ز	z	Z	ع	ʕ	Ḡ	ـ	an	—
س	s	S	غ	ɣ	Ḡ	ـ	in	—
ش	ʃ	Š	ف	f	F	ـ	un	—
ص	sʕ	Ṣ	ق	q	Q			
ض	dʕ	Ḍ	ك	k	K			
ط	tʕ	Ṭ	م	m	M			
ظ	ðʕ	Ẓ	ه	h	H			
ل	l	L	و	u:/w	U			
ن	n	N	ي	i:/j	I			

[Graf]ema árabe; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semi-vogal.

4.2.7 O sistema da UNESP

O conjunto de Dicionários Temáticos Visuais Bilíngues⁸, do Grupo de Pesquisa (GP) em Lexicologia e Lexicografia Contrastiva da UNESP, apre-

⁷Por se tratar de um elemento usado por uma minoria étnica, pouco pudemos verificar além da existência de um endereço UNICODE para ⟨G⟩ (01E4) e outro para g (01E5).

⁸<http://www.dtvb.ibilce.unesp.br>
Último acesso em 2011-07-16.

senta, dentre os pares de línguas sendo trabalhados, o português-árabe⁹. Na descrição da microestrutura dessa seção do dicionário¹⁰, explica-se terem sido consultados os trabalhos de seis autores, cujos sistemas de transposição ora coincidiam, ora divergiam.

Com base nessa consulta, o GP dividiu as correspondências estabelecidas nas romanizações em três grupos. O **primeiro** congrega as correspondências que representam fonemas semelhantes a seus equivalentes em português e cuja transposição mostra-se «praticamente consensual entre os autores pesquisados»; o **segundo** contém as referentes a fonemas similares a alguns presentes em línguas «habitualmente mais conhecidas e difundidas, tais como inglês, francês, espanhol ou alemão»; o **terceiro** reúne as de fonemas que não possuem correspondentes nem em português, nem nas referidas línguas. A esses grupos chamaremos ⟨equivalentes⟩, ⟨familiares⟩ e ⟨não-familiares⟩, respectivamente.

Neste último grupo encontram-se as consoantes faringalizadas e o ⟨ǧ⟩ /ʕ/ {ʕayn}, representado de maneiras bastante diversas, conforme o autor, a saber:

c (dois autores);

ε (dois autores)¹¹;

' (um autor)¹²;

\ (um autor);

“A (um autor);

°a ou °A (tabela de transliteração das bibliotecas da USP e da UFRJ).

⁹<http://www.dtvb.ibilce.unesp.br/menu.php?dici=arab>

¹⁰http://www.dtvb.ibilce.unesp.br/introducao/microestrutura_do_dicionario_portugues_arabe.html

¹¹Note-se a semelhança de forma entre ⟨ε⟩ e ⟨ǧ⟩.

¹²Note-se o possível conflito com o registro de ⟨ʕ⟩ /ʕ/, cuja representação é normalmente feita também por ⟨ʕ⟩.

Por sua vez, o GP, em seu esquema de romanização (tab. 45), optou por adotar uma representação própria para o ^ʿayn:

[...] optamos por transliterar esta letra ع (‘aīn) para a minúsculo com trema = ä, tendo em vista que, embora essa letra seja uma consoante na língua árabe, aos ouvidos ocidentais o som que ela representa parece similar à vogal “a”, como se fosse pronunciada duas vezes, aa. Contudo, para a sua representação quando acompanhada dos sinais que representam as vogais breves, adotamos às (sic) seguintes transliterações:

- para ع (‘aīn) com ْ (sukūn), de valor consonântico [...], usamos apenas a com trema = ä.
- para ع (‘aīn) com َ (fatHa [...]), usamos a com trema mais a = äa.
- para ع (‘aīn) com ُ (Damma [...]), usamos u com trema mais u = üu.
- e para ع (‘aīn) com ِ (kasra [...]), usamos i com trema mais i = iī

Esse trabalho indica que, além da divisão em três categorias de correspondência entre árabe e português, há espaço para discutirmos a representação de algumas letras, sendo o caso da consoante ^ʿayn provavelmente o mais nítido.

4.2.8 O sistema leigo de Jubran

Em seu artigo, JUBRAN (2004) apresenta dois sistemas de romanização: um voltado para o público leigo, que chamaremos de ⟨romanização leiga⟩ e que veremos agora (tab. 46); um visando ao público acadêmico, que chamaremos de ⟨romanização acadêmica⟩ de que trataremos em 4.2.9 (p. 116).

A romanização leiga é «uma convenção que pode ser adequada aos propósitos da mídia escrita [...], que [...] não necessit[a] de muito rigor na transliteração, mas apenas de padronização». Esse menor rigor reflete-se no uso de um mesmo grafema para representar diferentes fonemas (e.g, ⟨T/t⟩ para /t/ e /tʰ/) e no fato de que, não tendo de recorrer a diacríticos incomuns ao português, os grafemas podem ser «con-

Tabela 45: O sistema de romanização da UNESP.

grafemas solares			grafemas lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ت	t	t	ء	ʔ	ʔ	ا	a:	ā
ث	θ	th	ب	b	b	ة		—
د	d	d	ج	ʒ	j	ـَ	a	—
ذ	ð	dh	ح	ħ	H	ـِ	i	—
ر	r	r	خ	x	kh	ـُ	u	—
ز	z	z	ع	ʕ	āa, üi, üu	ـِـ	an	—
س	s	s	غ	ɣ	gh	ـِـ	in	—
ش	ʃ	ch	ف	f	f	ـِـ	un	—
ص	sʕ	S	ق	q	q			
ض	dʕ	D	ك	k	k			
ط	tʕ	T	م	m	m			
ظ	ðʕ	Z	ه	h	h			
ل	l	l	و	u:/w	ū			
ن	n	n	ي	i:/j	ī			

[Graf]ema árabe; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semivogal.

seguidos facilmente por qualquer fonte existente no computador». E considera-se também o leitor, que está «geralmente mais interessado na notícia do que na pronúncia ou na grafia corretas dos nomes em sua língua de origem, nomes que muitas vezes são “pulados” para que não se prejudique o ritmo rápido da leitura».

Vale notar que, dentre todos os esquemas e sistemas analisados, ambos os sistemas de JUBRAN (2004) são os únicos a apresentar as formas em caixa-alta e caixa-baixa (maiúsculas e minúsculas). Via de regra, tal distinção pode parecer trivial. No entanto, quando se trabalha com grafemas derivados, nem sempre a colocação do diacrítico pode ser a mesma em caixa-alta como em caixa-baixa (*vide* 4.2.6, p. 110).

Tabela 46: O sistema de romanização leiga de JUBRAN (2004).

grafemas solares			grafemas lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ﺕ	t	T/t	ﺀ	ʔ	*	ﺍ	a:	Á/á
ﺙ	θ	Th/th	ﺏ	b	B/b	ﺓ		ah
ﺩ	d	D/d	ﺝ	ʒ	J/j	ﺍَ	a	—
ﺫ	ð	Th/th	ﺡ	ħ	H/h	ﺍِ	i	—
ﺭ	r	R/r	ﺦ	x	Kh/kh	ﺍُ	u	—
ﺯ	z	Z/z	ﻉ	ʕ	*	ﺍَٔ	an	—
ﺱ	s	S/s	ﻍ	ɣ	Gh/gh	ﺍَٔ	in	—
ﺶ	ʃ	Ch/ch	ﻑ	f	F/f	ﺍَٔ	un	—
ﺺ	s ^ʕ	S/s	ﻕ	q	Q/q			
ﺽ	d ^ʕ	D/d	ﻙ	k	K/k			
ط	t ^ʕ	T/t	ﻡ	m	M/m			
ظ	ð ^ʕ	Zh/zh	ﻩ	h	H/h			
ﻝ	l	L/l	ﻭ	u:/w	Ú/ú, W/w			
ﻥ	n	N/n	ﻱ	i:/j	Í/í, Y/y			

[Graf]ema árabe; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semivogal. Onde houver dois pares de romanização, o primeiro refere-se à vogal longa; o segundo, à semivogal. *A ⟨ﺀ⟩ /ʔ/ {hamza^t} será romanizada ⟨A/a, I/i, U/u⟩, conforme a vogal seguinte; o ⟨ﻉ⟩ /ʕ/ {ʕayn} será romanizado ⟨A, I, U⟩, conforme a vogal seguinte.

4.2.9 O sistema acadêmico de Jubran

A romanização acadêmica de JUBRAN (2004) «traz uma sugestão para os textos, acadêmicos ou não, sejam eles de teor científico ou literário». Seguindo o mesmo princípio encontrado nas romanizações de Nimer (4.2.5, p. 109) e de Sabbagh (4.2.6, p. 110), emprega-se sempre apenas um grafema simples (primitivo ou derivado) para representar cada fonema (tab. 47).

Tabela 47: O sistema de romanização acadêmica de JUBRAN (2004).

grafemas solares			grafemas lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ت	t	T/t	ء	ʔ	'	ا	a:	Ā/ā
ث	θ	T̄/t̄	ب	b	B/b	ة		ah
د	d	D/d	ج	ʒ	J/j	ـَ	a	—
ذ	ð	Ḍ/ḍ	ح	ħ	H/h	ـِ	i	—
ر	r	R/r	خ	x	Ḥ/ḥ	ـُ	u	—
ز	z	Z/z	ع	ʕ	ʿ	ـِـ	an	—
س	s	S/s	غ	ɣ	Ġ/ġ	ـِـ	in	—
ش	ʃ	Š/š	ف	f	F/f	ـِـ*	un	—
ص	sʕ	Ṣ/ṣ	ق	q	Q/q			
ض	dʕ	Ḍ/ḍ	ك	k	K/k			
ط	tʕ	Ṭ/ṭ	م	m	M/m			
ظ	ðʕ	Ḍ̣/ḍ̣	ه	h	H/h			
ل	l	L/l	و	u:/w	Ū/ū, W/w			
ن	n	N/n	ي	i:/j	Ī/ī, Y/y			

[Graf]ema árabe; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semivogal.

Onde houver dois pares de romanização, o primeiro refere-se à vogal longa; o segundo, à semivogal.

São bastante abrangentes e pormenorizadas as orientações para o uso da romanização acadêmica, no que se refere a diversos elementos ortográficos e morfosintáticos: as marcas de caso e indefinição, que na leitura fluente não se realizam; o artigo definido <ال> |al-| e sua assimilação frente às letras solares, bem como a supressão da vogal, em certos

contextos; os termos <عبد> |^cabd|, <ابن> |^qibn| e <أبو> |^qabū|, usados em nomes próprios.

4.3 em análise

Ao cotejarmos os esquemas e sistemas de romanização apresentados em 4.2 (pp. 103–117), observamos haver dois conjuntos de correspondência: o das **romanizações consensuais**, i-e, o dos fonemas ou grafemas árabes representados pelos mesmos grafemas latinos, independente do esquema ou sistema; os de **romanização não-consensual**, i-e, o dos fonemas ou grafemas árabes representados por diferentes grafemas latinos, dependendo do esquema ou sistema. Esses dois conjuntos podem ser cruzados com os três grupos apontados pelo GP da UNESP (4.2.7, p. 112), para orientar nossa análise das correspondências grafema–fonema a ser selecionadas em nossa proposta (capítulo 5, p. 122).

4.3.1 Correspondências consensuais — fonemas equivalentes

Dos 28 grafemas básicos¹³ do abjad árabe, 12 encontram consenso¹⁴ nos esquemas e sistemas de romanização analisados (tab. 48).

4.3.2 Correspondências não-consensuais — fonemas equivalentes

Dentre as correspondências não-consensuais, quatro referem-se a fonemas existentes no português. As romanizações variam entre grafemas simples e compostos, primitivos e derivados (tab. 49). E algumas dessas romanizações coincidem com as relações biunívocas selecionadas para o português (tab. 37, p. 94).

¹³As 14 letras solares, 14 lunares.

¹⁴Exceção única feita à correspondência de CUNHA (1987) para <◌◊>–<◌h> (tab. 40, p. 105).

Tabela 48: Romanização consensual — fonemas equivalentes.

plosivas			fricativas			nasais			líquidas		
ár.	fon.	rom.	ár.	fon.	rom.	ár.	fon.	rom.	ár.	fon.	rom.
ب	b	B, b	ف	f	F, f	م	m	M, m	ر	r	R, r
ت	t	T, t	س	s	S, s	ن	n	N, n	ل	l	L, l
د	d	D, d	ز	z	Z, z						
ك	k	K, k	ه	h	H, h						

Grafema [ár]abe; [fon]ema; [rom]anização.

Tabela 49: Romanização não-consensual de fonemas equivalentes.

graf·ár	fon.	romanizações		
		sem diacrítico	com diacrítico	
ش	ʃ	X ₁	CH ₃	Š ₄
ج	ʒ	J ₇		Ž ₁
ي	j	Y ₂	I ₃	(I) ₆
و	w	W ₃	U ₃	(U) ₅

[Graf]ema [ár]abe; [fon]ema.

Os índices representam a quantidade de ocorrências. Os parênteses () indicam que o grafema composto apresenta variantes diversas.

4.3.3 Correspondências não-consensuais — fonemas familiares

Os fonemas familiares — presentes em idiomas europeus como espanhol, francês e inglês —, quando representados por grafemas compostos, estão invariavelmente associados à composição de um grafema básico com o ⟨H⟩ — certamente por serem todos fonemas fricativos. Dentre os grafemas derivados, empregam-se diferentes diacríticos: traço inferior ⟨₋⟩; trema ⟨_.⟩; barra ⟨₋⟩; ponto ⟨_.⟩; bráquia ⟨_^⟩ (tab. 50).

Tabela 50: Romanização não-consensual de fonemas familiares.

graf·ár	fon.	romanizações				
		sem diacrítico		com diacrítico		
ث	θ	TH ₄		T ₃	T ₂	
ذ	ð	DH ₂	TH ₁	D ₂	D ₁	Z ₁
خ	x	KH ₄		K ₂	H ₁	
غ	ɣ	G ₁	GH ₃	G ₃	G ₁	

[Graf]ema [ár]abe; [fon]ema.

Os índices representam a quantidade de ocorrências.

Devemos ressaltar que o fonema /ɣ/ é normalmente descrito como similar ao /R/ francês, como neste excerto de WRIGHT (1859):

ξ é um *g* gutural, acompanhado por o som intermitente ou vibrado de um gargarejo, e não encontra similar em inglês. O γ do grego moderno, *r* da Nortúmbria¹⁵ e o *r grasséyé* francês são aproximações.¹⁶

e neste de LINGS (2010):

¹⁵*Northumbria*, região fronteira entre Inglaterra e Escócia.

¹⁶No original: «ξ is a guttural *g*, accompanied by a grating or rattling sound as in gargling, of which we have no example in English. The γ of the modern Greeks, the Northumbrian *r*, and the French *r grasséyé*, are approximations to it.» (p. 5)

Não é pronunciado como *g* de gato, mas do fundo da garganta, emitindo-se um *r* vibrante profundo, similar ao “gargarejo” ou o *r* francês de *ch[é]rie*.

E COWAN (2007) também afirma que «é exatamente o som emitido quando se faz gargarejo. Parece com o *r grasseyé* do francês». Por esse motivo, colocamo-lo entre os fonemas familiares.

4.3.4 Correspondências não-consensuais — fonemas não-familiares

As consoantes guturais tendem a ser representadas por um grafema composto com um ponto ⟨.⟩; exceção feita ao fonema /q/, normalmente romanizado com ⟨Q⟩. Outros diacríticos incluem: traço inferior ⟨_⟩; traço inferior e ponto ⟨_·⟩; cedilha ⟨,⟩; bráquia ⟨_⟩ (tab. 51).

Tabela 51: Romanização não-consensual de fonemas não-familiares.

graf·ár	fon.	romanizações				
		sem diacrítico		com diacrítico		
ط	t ^ʕ	T ₃		Ṭ ₅		
ض	d ^ʕ	D ₃		Ḍ ₅		
ق	q	Q ₇		Q̣ ₁		
ظ	ð ^ʕ	Z ₂	ZH ₁	Ẓ ₃	Ẓ ₁	Ḍ ₁
ص	s ^ʕ	S ₃		Ṣ ₁	Ṣ ₄	
ح	ħ	H ₃		Ḥ ₄	Ḥ ₁	

[Graf]ema [ár]abe; [fon]ema.

Os índices representam a quantidade de ocorrências.

4.3.5 Correspondências não-consensuais — outros fonemas

A romanização de ⟨ع⟩ /ʕ/ e ⟨ء⟩ /ʔ/ tende a ser feita por sinais que não derivam de letras, quase todos posicionados acima da altura de x, no espaço ocupado por caracteres sobrescritos.

A romanização do ʿayn ⟨ع⟩ é bastante diversificada; não há consenso entre os autores estudados, sendo os sinais gráficos os mais variados (tab. 52). A da hamza^t ⟨ء⟩ /ʔ/ tende a ser feita com o apóstrofo ⟨'⟩. E alguns autores sequer prevêm o registro de uma e/ou outra.

Tabela 52: Romanização não-consensual
de ⟨ع⟩ /ʕ/ e ⟨ء⟩ /ʔ/.

graf·ár	fon.	romanizações			
		sinais variados			letras
ع	ʕ	⟨'⟩	⟨`⟩	⟨¨⟩	⟨ ^c ⟩ ε
ء	ʔ	⟨'⟩			

[Graf]ema [ár]abe; [fon]ema.

Os parênteses angulares foram usados para melhor visualizar o posicionamento dos grafemas.

CAPÍTULO 5

Proposta de transposição ortográfica

Sapere aude.

«Ouse saber.»

Horácio

ESTABELECEMOS os princípios para a elaboração do sistema de transposição ortográfica e apresentamos seus quatro módulos, a saber: romanização leiga simplificada; romanização leiga comum; romanização acadêmica ampla; romanização acadêmica estrita. Apresentamos também alguns exemplos de uso desses módulos em palavras e expressões árabes, além de um breve texto.

5.1 A proposta

Nossa proposta de transposição ortográfica do árabe padrão (AP) para o português brasileiro (PB), na verdade, consta de quatro propostas de romanização: acadêmica ampla; acadêmica estrita; leiga comum; leiga simplificada. Cada uma tem público-alvo e/ou objetivo próprios, sendo distintas na abrangência e na pormenorização dos elementos abordados; todas têm como base comum a romanização acadêmica ampla, que estabelece as correspondências grafema–fonema a partir das quais as demais romanizações são derivadas.

5.1.1 Os princípios

Assim como a seleção de grafemas para representar os fonemas do português seguiu certos princípios, também as propostas seguem alguns:

- constituir grafemas somente a partir de letras;
- apresentar biunivocidade nas correspondências grafema–fonema;
- apresentar biunivocidade entre os dígrafos e as modificações por eles indicadas;
- usar caracteres básicos como grafemas simples primitivos (uma letra sem diacríticos);
- usar caracteres codificados no UNICODE.

5.2 A romanização acadêmica ampla

A romanização acadêmica ampla (RAA) constitui o cerne de nossa proposta. O público-alvo —como pretendemos explicitar na própria designação— são os usuários do meio acadêmico, que necessitem de um meio, no português brasileiro e em caracteres latinos, para grafar nomes, termos e expressões com abrangência e sem ambiguidades.

5.2.1 As vogais

As seis vogais do AP são de romanização bastante simples. Divididas em dois grupos —longas e breves—, a cada vogal corresponderá o grafema latino de seu fonema equivalente em PB; a distinção de qualidade será feita com o uso do macron para marcar as vogais longas (tabela 53).

Tabela 53: Romanização das vogais.

longas			breves		
graf·ár	fon.	rom.	graf·ár	fon.	rom.
ا	a:	ā	ا	a	a
ي	i:	ī	ي	i	i
و	u:	ū	و	u	u

[Graf]ema [ár]abe; [fon]ema; [rom]anização.

5.2.2 As consoantes — romanização consensual de fonemas equivalentes

Partindo do alfabeto latino de 21 letras —já excluídas as 5 usadas no PB para grafar as vogais—, podemos começar a estabelecer a romanização das consoantes do AP. As romanizações consensuais (*vide* 4.3.1, página 117) constituem um bom ponto de partida, pois empregam somente grafemas simples primitivos (uma letra sem diacrítico), ou seja, todas as 12 correspondências podem ser acomodadas dentre as 21 letras, restando 9 para uso posterior (tab. 54).

5.2.3 As consoantes — romanização não-consensual de fonemas equivalentes

Quatro das nove letras latinas disponíveis podem ser usadas para romanizar os fonemas equivalentes não contemplados pela romanização consensual (*vide* 4.3.2, p. 117). Se cruzarmos esses fonemas árabes com

Tabela 54: Seleção de grafemas simples primitivos — romanização consensual de fonemas equivalentes.

letras latinas	B	C	D	F	G	H	J	K	L	M	N	P	Q	R	S	T	V	W	X	Y	Z
grafemas selecionados	B		D	F		H		K	L	M	N			R	S	T					Z
fonemas árabes	b		d	f		h		k	l	m	n			r	s	t					z
grafemas árabes	ب		د	ف		ه		ك	ل	م	ن			ر	س	ت					ز
letras disponíveis		C				G	J						P	Q			V	W	X	Y	

as correspondências biunívocas selecionadas para o PB em 3.3.3 (p. 93), veremos que as consoantes árabes /ʃ, ʒ/ podem ser romanizadas com ⟨x, j⟩; as semivogais /j, w/, com ⟨y, w⟩ (tab. 55) — as correspondências de ⟨ي, > quando grafam /aː, uː/ estão estabelecidas em 5.2.1 (p. 124).

Tabela 55: Seleção de grafemas simples primitivos — romanização não-consensual de fonemas equivalentes.

letras latinas	B	C	D	F	G	H	J	K	L	M	N	P	Q	R	S	T	V	W	X	Y	Z
letras disponíveis		C				G	J						P	Q			V	W	X	Y	
grafemas selecionados							J											W	X	Y	
fonemas árabes							ʒ											w	ʃ	j	
grafemas árabes							ج											و	ش	ي	
letras disponíveis		C				G							P	Q			V				

5.2.4 As consoantes — romanização não-consensual de fonemas familiares

Os fonemas familiares (*vide* 4.3.3, p. 119) são todos fricativos; cada um dos quatro possui uma contraparte plosiva no PB (tab. 56). Não havendo grafemas simples primitivos (uma letra sem diacrítico) que possam representá-los, esses dois fatores permitem-nos tomar como base os grafemas de suas respectivas contrapartes não-fricativas, com o acréscimo de ⟨g⟩ (tab. 57).

Tabela 56: Romanização de fonemas familiares — grafemas simples e compostos.

plosivas			fricativas		
rom.	fon.	graf.ár	fon.	rom.	
T	t	ت ث	θ	?	
D	d	د ذ	ð	?	
K	k	ك خ	x	?	
(G)	g	– غ	ɣ	?	

[Rom]anização; [fon]ema; [graf]ema [ár]abe.

Tabela 57: Seleção de grafemas simples primitivos — romanização fonemas familiares.

letras latinas	B	C	D	F	G	H	J	K	L	M	N	P	Q	R	S	T	V	W	X	Y	Z
letras disponíveis		C			G								P	Q			V				
grafema selecionado					G																
fonema árabe					ɣ																
grafema árabe					غ																
letras disponíveis		C											P	Q			V				

Para evitarmos o uso de grafemas compostos (dígrafos), precisamos recorrer ao uso de um diacrítico. Considerando tratar-se unicamente de fricativas, um mesmo diacrítico aplicado a suas contrapartes plosivas serviria para marcar essa relação de maneira consistente. Em outras palavras, seria usado um mesmo sinal para indicar uma mesma modificação no valor do grafema em questão.

O diacrítico que adotaremos é o mesmo usado em algumas das romanizações estudadas (*vide* tab. 50, p. 119). Ele será usado sob as letras ⟨T, D, K⟩ e sobre a letra ⟨G⟩ (tab. 58), pois esses grafemas derivados já se encontram codificados no UNICODE. Trata-se do traço inferior ⟨_⟩:

traço inferior [...] Pode ser utilizado como uma alternativa ao ponto inferior na composição do árabe e do hebraico romanizados. Para desimpedir as descendentes, é preciso utilizar uma versão reposicionada do caractere em questão. (BRINGHURST, 2005, p. 348)

UNICODE:
codificação
padronizada
que reúne
a maioria
dos sistemas
de escrita,
para uso
em meios
eletrônicos.

Os pares plosivo–fricativo mostram as relações entre grafemas primitivos e derivados (tab. 59).

Tabela 58: Romanização de fonemas familiares — grafemas derivados.

	dentais			velares		
	g·ár	fon.	rom.	g·ár	fon.	rom.
desvozeados	ث	θ	Ṭ, ṭ	خ	x	Ḳ, ḳ
vozeados	ذ	ð	Ḍ, ḍ	غ	ɣ	G̣, g̣

[G]rafema [ár]abe; [fon]ema; [rom]anização.

5.2.5 As consoantes — romanização não-consensual de fonemas não-familiares

Os fonemas não-familiares (*vide* 4.3.4, p. 120) podem ser todos classificados como guturais. WATSON (2002) afirma que «[e]m árabe [...], as fri-

Tabela 59: Uso do traço inferior.

plosivas		fricativas		
rom.	fon.	graf·ár	fon.	rom.
T t	t	ت ث	θ	Ṭ ṭ
D d	d	د ذ	ð	Ḍ ḍ
K k	k	ك خ	x	Ḳ ḱ
(G g)	g	– غ	ɣ	Ḡ ḡ

[Rom]anização; [fon]ema; [graf]ema [ár]abe.

cativas laringais, faringais e uvulares formam um padrão que as agrupa fonologicamente em uma classe gutural[...]»¹, na qual incluem-se as coronais faringalizadas (tʕ, dʕ, sʕ, ðʕ). Com isso concorda DE LACY (2007):

Em muitos idiomas, as consoantes ‘guturais’ (i.e. glotais, faringais e uvulares) formam uma classe natural. Por exemplo, o árabe clássico possui muitos processos referentes às glotais /h ʔ/, faringais /h ʕ/ e uvulares /χ ʁ/ [...]. Esses sons têm sido classificados como [gutural] [...] ou FARINGAL [...]².

Aplicando o mesmo procedimento de 5.2.4 (p. 125), pareamos as guturais com suas contrapartes não-guturais (tab. 60).

A partir dessa comparação, ao grafema usado para representar as contrapartes não-guturais aplicamos um diacrítico que indique o traço gutural, qual seja, o ponto inferior ⟨.⟩:

ponto inferior Sinal diacrítico usado em consoantes [...] em árabe, hebraico e sânscrito romanizados [...]. (BRINGHURST, 2005, p. 344)

¹No original: «[i]n Arabic [...], the laryngeals, pharyngeals, and uvular fricatives pattern together phonologically as a guttural class [...]» (p. 37).

²No original: «In many languages the ‘guttural’ consonants (i.e. glottals, pharyngeals and uvulars) form a natural class. For example, Classical Arabic has many processes referring to glottals /h ʔ/, pharyngeals /h ʕ/ and uvulars /χ ʁ/ [...]. These sounds have been captured with [guttural] [...] or PHARYNGEAL [...].» (p. 329)

Tabela 6o: Romanização de fonemas não-familiares — grafemas primitivos e derivados.

plosivas e fricativas			dorsais e radicais		
rom.	fon.	graf·ár	fon.	rom.	
T t	t	ت ط	tʰ	?	
D d	d	د ض	dʰ	?	
K k	k	ك ق	q	?	
Ḍ ḍ	ḍ	ذ ظ	ḍʰ	?	
S s	s	س ص	sʰ	?	
H h	h	ه ح	h	?	

[Rom]anização; [fon]ema; [graf]ema [ár]abe.

Devemos fazer uma ressalva ao grafema derivado (com diacrítico) a ser usado para representar o fonema /ḍʰ/. Se aplicarmos o ponto ⟨.⟩ ao ⟨Ḍ⟩, poderemos esperar ao menos três diferentes posições para a colocação do ponto em relação ao traço inferior: sobre ⟨Ḍ⟩ ou ⟨Ḍ̣⟩; ao lado ⟨Ḍ̣⟩; abaixo ⟨Ḍ̣̣⟩³.

A fim de evitar essa diversidade de possíveis resultados, principalmente considerando-se a variedade de ferramentas de produção de texto —versões, programas e/ou sistemas operacionais—, preterimos o uso de ⟨Ḍ̣̣⟩ para representar /ḍʰ/, em favor de ⟨Ḍ̣̣⟩.

Há de se atentar para a diferença entre o ⟨Ḍ̣̣⟩ de caixa-baixa ⟨ḍ̣̣⟩, o ⟨Ḍ̣̣⟩ de caixa-baixa ⟨ḍ̣̣⟩ e o ⟨Ḍ̣̣⟩ de caixa-baixa ⟨ḍ̣̣⟩⁴. O primeiro par ⟨Ḍ̣̣, ḍ̣̣⟩ chama-se:

dyet Letra básica do alfabeto em servo-croata e vietnamita, também usada no macedônio romanizado. Sua forma maiúscula é, via

³JUBRAN (2004), em sua romanização acadêmica (*vide* 4.2.9, páginas 116–117) e WATSON (2002, p. 38) usam esta forma.

⁴Caso similar pode ser visto em 4.2.6 (pp. 110–111), a respeito do posicionamento do diacrítico na forma em caixa-baixa de ⟨Ḡ⟩.

de regra, a mesma do *eth* maiúsculo, mas seu conceito é diverso, e por isso seu endereço Unicode é diferente. (BRINGHURST, 2005, p. 337)

Sobre o *dyet*, NGUYỄN (1996)⁵ afirma que:

[a] oclusiva dental vozeada [d] escreve-se com *ḑ*, em vez de sê-lo com *d*, que representa a aspirante vozeada [z] [...] (o atual [z] < [d] *d* so século xvii, enquanto o [d] < [ḑ] *ḑ* (implosivo), eis o porquê das estranhas correspondências entre as letras).⁶

O segundo par ⟨Ḑ, ḑ⟩ chama-se *eth*:

eth Letra do alfabeto em anglo-saxão, feroês, em islandês e da IPA. O *eth* maiúsculo é igual ao *dyet*, mas as formas minúsculas não são intercambiáveis e as letras representam sons bastante diferentes (a pronúncia do nome *eth* aproxima-se do som de um z sibilado, como na palavra *azul* pronunciada com língua presa). (BRINGHURST, 2005, p. 338)

O terceiro par ⟨Ḑ, ḑ⟩ é chamado de ⟨Ḑ africano⟩, sendo usado no alfabeto de referência africano⁷.

Estando estabelecida a correspondência entre o fonema /ḑ/ e ⟨Ḑ, ḑ⟩, sua contraparte faringalizada /ḑ̤/ será romanizada com ⟨Ḑ̤, ḑ̤⟩ (tab. 61).

5.2.6 As consoantes — romanização não-consensual de ⟨ɛ⟩

/ʎ/ e ⟨ɛ̃⟩ /ʎ/

Podemos utilizar duas das quatro letras disponíveis na tab. 57 (p. 126). Para ⟨ɛ̃⟩ /ʎ/, selecionamos ⟨c⟩; para ⟨ɛ̃̃⟩ /ʎ̃/, ⟨q⟩ (tab. 62). Ambos deverão ser sobrescritos.

⁵Este autor possui o grafema em questão no início de seu prenome: Đình-Hoà.

⁶No original: «[t]he voiced dental stop [d] is written with barred *ḑ*, as opposed to non-barred *d*, which represents the voiced spirant [z] [...] (modern [z] < 17th-century [d] *d*, whereas [d] < implosive [ḑ] *ḑ*, hence the odd-seeming assignment of the letters).» (p. 693)

⁷Por se tratar de um elemento da tipografia africana, pouco pudemos verificar além da existência de um endereço UNICODE para ⟨Ḑ⟩ (0189) e outro para ⟨ḑ̤⟩ (0256).

Tabela 61: Romanização de fonemas não-familiares — grafemas primitivos e derivados.

plosivas e fricativas		dorsais e radicais			
rom.	fon.	graf. ár		fon.	rom.
T t	t	ت	ط	t ^ʕ	Ṭ ṭ
D d	d	د	ض	d ^ʕ	Ḍ ḍ
K k	k	ك	ق	q	Ḳ ḳ
Ḍ ḍ	ḍ	ذ	ظ	ḍ ^ʕ	Ḍ ḍ
S s	s	س	ص	s ^ʕ	Ṣ ṣ
H h	h	ه	ح	ḥ	Ḥ ḥ

[Rom]anização; [fon]ema; [graf]ema [ár]abe.

Tabela 62: Seleção de grafemas simples primitivos — romanização de <ع> /ʕ/ e <ء> /ʔ/.

letras latinas	B C D F G H J K L M N P Q R S T V W X Y Z																									
letras disponíveis	C													P Q		V										
grafemas selecionados	C													Q												
fonemas árabes	ʕ													ʔ												
grafemas árabes	ع													ء												
grafemas restantes														P		V										

5.2.6.1 As não-letras

Essa escolha diverge do habitual nas romanizações do árabe, que costumam recorrer à vírgula invertida e ao apóstrofo ⟨ ‘ ’ ⟩ (aspas simples de abertura e de fechamento):

vírgula invertida No árabe e no hebraico transliterados, [...] é usada para representar a letra *‘ain* (ع) ou *ayin* (ע), uma continuante faríngea (IPA ʕ), enquanto seu sinal oposto, o apóstrofo, representa a parada glotal. Os nomes Rei Ibn Sa‘ud e Vale de Beqa ‘a são exemplos de seu uso. Apenas a forma bem-curvada do glifo (‘ em vez de ’) é útil em trabalhos de transliteração. (BRINGHURST, 2005, p. 350)

parada glotal A parada glotal é um som em busca de um caractere, um som básico de muitas línguas que agora é escrito em letras latinas, mas para o qual os alfabetos latino, grego e cirílico não têm nenhum símbolo tradicional. Os lingüistas usam o caractere ʔ ou ʕ (um ponto de interrogação castrado) para representá-lo, mas o símbolo mais comumente usado em textos compostos é o apóstrofo. No árabe romanizado, por exemplo, a vírgula invertida ou a aspa de abertura (‘) é muitas vezes usada para representar a letra *‘ain* (ع), cujo símbolo fonético é ʕ ou ʕ, e o apóstrofo é usado para representar a *hamza* (ء, إ, ؤ, etc), a parada glotal árabe. Desse modo, a grafia romanizada da palavra árabe “alcorão” é *al-Qur‘an*, a palavra “árabe” é grafada como *‘arab*, e “a família” é escrita como *al-‘a‘ila*. (BRINGHURST, 2005, p. 343)

Note-se que o segundo ⟨A⟩ de ⟨al-‘a‘ila⟩ {a família} encontra-se tipograficamente isolado do restante da palavra. O mesmo pode ser observado em algumas romanizações de ⟨صنعاء⟩ /sʕanʕaːʔ/ (capital do Iêmen) e de ⟨عجام⟩ /ʕiʕʒaːm/ (sistema de pontuação dos grafemas árabes); isso não se observa caso usemos os grafemas sobrescritos (tab. 65).

Uma opção pouco comum em trabalhos não-especializados são ⟨ ˁ ⟩ e ⟨ ˀ ⟩, usados principalmente na romanização do hebraico bíblico (LAMBDIN, 2003, p. 19), mas também em algumas romanizações do árabe (BAUER, 1996, p. 560). Essa solução permite distinguir um grafema de outro com relativa clareza, ainda que se mantenha o problema do isolamento, parti-

cularmente devido à diferença de peso⁸ entre esses grafemas e os demais (tab. 63).

Tabela 63: Isolamento tipográfico decorrente do uso de ⟨‘ ’⟩ ou ⟨‘ ’⟩.

árabe	tr·fonol	rom·esp	rom·com
العائلة	ʔalʔa:ʔila	al-‘a’ila	al-‘a’ila
صنعاء	sʕanʔa:ʔ	San‘a’	San‘a’
إعجام	ʔiʔʒa:m	ʔi‘jām	ʔi‘jām

[Tr]anscrição [fonol]ógica;

[rom]anização [esp]ecializada, [com]um.

Além de não contar com formas visualmente muito distintas, o uso de ⟨‘ ’⟩ pode apresentar problemas na produção⁹. Uma solução para o problema da produção das aspas foi apresentada, para uso na romanização do hebraico, por KIRSCHBAUM *et al.* (2009): aspas simples para o alef ⟨א⟩; aspas duplas para o ain ⟨ע⟩ (tab. 64).

Tabela 64: Aspas simples e duplas na romanização do hebraico.

aspas	hebraico	romanização	tradução
simples (א)	אָמָא	‘íma’	mãe
	רֹאשׁ	ro’sh	cabeça
duplas (ע)	עֲמֻדָּה	“amúd	página, coluna
	נֶעִיִּם	na”ím	agradável

Exemplos de KIRSCHBAUM *et al.* (2009).

⁸Trata-se do «grau de escuridão (ou pretume) de um tipo, independentemente de seu tamanho» (BRINGHURST, 2005, p. 361).

⁹Normalmente, nos teclados dos computadores há somente uma tecla para produzir as aspas simples, sendo a escolha entre ⟨‘⟩ e ⟨’⟩ feita pelo *software*, conforme o contexto — se há ou não algum caractere antes das aspas. Em uma romanização menos cuidadosa, pode-se obter ⟨‘i’jām⟩, em vez de ⟨ʔi‘jām⟩.

Outro problema relacionado ao uso de ⟨‘ ’⟩ pode ser observado —por eventual que seja— nos casos em que esses grafemas são usados logo antes de ou após aspas (simples ou duplas) efetivas. Em NIMER (2005, p. 47) encontramos três exemplos neste trecho: «[...] há assimilação total do elemento “a” (ou, segundo os casos, “u” ou “i”) [...]». Se as palavras da tabela 64 forem mencionadas da mesma forma, teremos: ⟨“ima”⟩; ⟨“ro’sh”⟩; ⟨“amúd”⟩; ⟨“na’im”⟩.

E cabe atentar para a diferença gráfica entre essas formas de romanização e dois diacríticos visualmente semelhantes, usados na ortografia politônica do grego antigo. Não confundir ⟨ˊ⟩ ou ⟨ˋ⟩ com ⟨˘⟩, este o espírito áspero —do latim ⟨*spīritus asper*⟩, para o original ⟨δασὺ πνεῦμα⟩ |dasy pneuma|—, encontrado em ⟨ἥρως⟩ |hérōs| {herói} e ⟨ῥητορικός⟩ |rhētorikós| {retórica}; nem confundir ⟨˘⟩ ou ⟨˙⟩ com ⟨ˊ⟩, este o espírito brando —do latim ⟨*spīritus lēnis*⟩ para o original ⟨ψιλὸν πνεῦμα⟩ |psilon pneuma|— encontrado em ⟨ἐργον⟩ |ergon| {trabalho}, étimo de ⟨ergonomia⟩.

5.2.6.2 As letras

Por sua vez, empregar letras —em vez de outros sinais (vide 4.3.5, p. 120)— para constituir os grafemas para esses fonemas traz algumas vantagens, permitindo:

- alfabetação (organização por ordem alfabética);
- diferenciação entre caixa-alta e caixa-baixa (maiúsculas e minúsculas);
- inclusão desses grafemas, sem alterações que não a de posicionamento, na romanização acadêmica estrita (vide 5.3, pp. 140–141).

bem como:

- apresentação mais tipograficamente elegante (tab. 65);
- distinção melhor entre um e outro grafema;
- representação mais consistente das consoantes, sem o risco de acontecerem espelhamentos.

Tabela 65: Vantagem tipográfica do uso de ⟨^c q⟩.

árabe	tr·fonol	rom·esp	rom·com	RAA
العائلة	ʔalʔa:ʔila	al- ^c aʔila	al- ^c aʔila	al- ^c a ^q ila ^t
صنعاء	s ^ʔ anʔa:ʔ	San ^c aʔ	San ^c aʔ	Şan ^c ā ^q
إعجام	ʔiʔʒa:m	ʔi ^c jām	ʔi ^c jām	^q i ^c jām

[Tr]anscrição [fonol]ógica;

[rom]anização [esp]ecializada, [com]um.

E respaldo para o uso desses grafemas pode ser encontrado no uso que se faz do alfabeto latino em algumas línguas, como o somali e o maltês. O **somali** chegou a usar «um alfabeto [...] chamado Osmanyā, [que] foi substituído [...] por uma ortografia romana»¹⁰ (DANIELS, 1996b); desde 1972, quando «o governo somali promulgou a ortografia somali oficial usando o alfabeto romano sem letras adicionais»¹¹ (COULMAS, 1996), essa língua tem sido escrita com o alfabeto romano (KAYE, 1996, 744), alguns dígrafos e «letras que receberam novos valores, ⟨q⟩ para a parada glotal, ⟨c⟩ e ⟨x⟩ para as fricativas faríngeas vozeada e desvozeada.»¹² (COULMAS, 1996). Em **maltês** —dialeto árabe, falado na República de Malta, que adquiriu *status* autônomo— observa-se que «[o] uso de *q* para [ʔ] (derivado do árabe [q]) [...] é um reflexo das práticas de transliteração semíticas.»¹³ (COMRIE, 1996).

¹⁰No original: «an alphabet for Somali, called Osmanyā from the name of the inventor, ʔsman Yusuf, [...] has been replaced [...] by Roman orthography.» (p. 580).

¹¹No original: «In 1972 the Somali government promulgated an official Somali orthography using the Roman alphabet with no additional letters.» (p. 475).

¹²No original: «letters which are assigned new values, ⟨q⟩ for the glottal stop, and ⟨c⟩ and ⟨x⟩ for voiced and unvoiced pharyngeal fricatives.» (p. 475).

¹³No original: «The use of *q* for [ʔ] (deriving from Arabic [q]) [...] reflect[s] Semitic transliteration practices.» (p. 688).

5.2.7 Correspondências grafema–fonema da RAA

A RAA apresenta, portanto, dois conjuntos de grafemas. O conjunto dos grafemas primitivos (sem diacríticos) —escolhidos a partir do alfabeto latino (tab. 66)— abrange os fonemas equivalentes —tanto os 12 de romanização consensual quanto os 4 de não-consensual — e os 2 fonemas /ʔ, ʕ/, romanizados com grafemas sobrescritos; o dos grafemas derivados (com diacríticos) subdividem-se entre os 4 que são grafados com traço inferior ⟨_·⟩ e os 6 que o são com ponto inferior ⟨_·⟩ — colocados acima ou abaixo da letra, conforme o caso. Desse modo, a RAA permite representar, com grafemas simples (uma só letra, sem dígrafos), todos os 28 grafemas solares e lunares do abjad árabe. Além disso, permite registrar as vogais longas e breves, a nunação e o ta^q marbūṭa^q (tab. 67).

Tabela 66: Grafemas latinos selecionados para as respectivas correspondências grafema–fonema.

letras latinas	B	C	D	F	G	H	J	K	L	M	N	P	Q	R	S	T	V	W	X	Y	Z
rom·consensual	B		D	F		H		K	L	M	N			R	S	T					Z
fonemas árabes	b		d	f		h		k	l	m	n			r	s	t					z
grafemas árabes	ب		د	ف		ه		ك	ل	م	ن			ر	س	ت					ز
letras disponíveis		C			G		J						P	Q				V	W	X	Y
rom·não-consensual								J										W	X	Y	
fonemas árabes								ʒ										w	ʃ	j	
grafemas árabes								ج										و	ي	ش	
letras disponíveis		C			G								P	Q			V				
rom·familiar								G													
fonema árabe								ɣ													
grafema árabe								غ													
letras disponíveis		C											P	Q			V				
outras romanizações		C												Q							
fonemas árabes		ʔ												ʔ							
grafemas árabes		ع												ء							
letras restantes													P				V				

[Rom]anização.

Tabela 67: O sistema de romanização acadêmica ampla.

letras solares			letras lunares			outros		
graf.	fon.	rom.	graf.	fon.	rom.	graf.	fon.	rom.
ت	t	T/t	ء	ʔ	⟨Q/q⟩	ا	a:	Ā/ā
ث	θ	Ṭ/ṭ	ب	b	B/b	ة		A ^T /a ^t
د	d	D/d	ج	ʒ	J/j	ـَ	a	A/a, ⟨ ^{A/a} ⟩
ذ	ð	Ḍ/ḍ	ح	ħ	Ḥ/ḥ	ـِ	i	I/i, ⟨ ^{I/i} ⟩
ر	r	R/r	خ	x	Ḳ/k	ـُ	u	U/u, ⟨ ^{U/u} ⟩
ز	z	Z/z	ع	ʕ	⟨ ^{C/c} ⟩	ـِـ	an	⟨ ^{AN/an} ⟩
س	s	S/s	غ	ɣ	Ġ/ġ	ـِـ	in	⟨ ^{IN/in} ⟩
ش	ʃ	X/x	ف	f	F/f	ـِـ	un	⟨ ^{UN/un} ⟩
ص	s ^ʕ	Ṣ/ṣ	ق	q	Ḳ/ḳ			
ض	d ^ʕ	Ḍ/ḍ	ك	k	K/k			
ط	t ^ʕ	Ṭ/ṭ	م	m	M/m			
ظ	ð ^ʕ	Ḍ/ḍ	ه	h	H/h			
ل	l	L/l	و	u:/w	Ū/ū, W/w			
ن	n	N/n	ي	i:/j	Ī/ī, Y/y			

[Graf]ema; [fon]ema; [rom]anização.

Onde houver dois fonemas, o primeiro refere-se à vogal; o segundo, à semivogal. Onde houver dois pares de romanização, em ⟨grafemas lunares⟩, o primeiro refere-se à vogal longa e o segundo, à semivogal; em ⟨outros⟩, o primeiro refere-se à vogal breve e o segundo, às marcas de caso ^e/ou indefinição.

5.2.8 Alguns pontos pertinentes

5.2.8.1 à ortografia

A **geminção** é registrada simplesmente repetindo-se a consoante. Por exemplo, كَسَرَ |kasara|¹⁴ {quebrar} e كَسَّرَ |kassara| {fragmentar, estilhaçar}.

O تاء مربوطة, <ة>, |tā^q marbūṭa^t| {ta^q amarrado}, é sempre grafado sobrescrito, mesmo quando se registram as marcas de caso e/ou de indefinição (vide 5.2.8.2). Nos casos em que o tā^q marbūṭa “é aberto”, segue-se a ortografia original e grafa-se um <τ> comum. Por exemplo, مُدَرِّسَةٌ |mudarrisa^{tu}|¹⁵ {uma professora (NOMINATIVO)} e مُدَرِّسَاتٌ |mudarrisāt^{un}| {umas professoras (NOM.)}.

5.2.8.2 à morfossintaxe

As marcas de **caso** são grafadas sobrescritas, assim como as de **indefinição**. Por exemplo, بَيْتٌ |bayt^{un}| {uma casa (NOM.)} (vide tab. 18, p. 73).

Algumas **vogais longas** podem, especialmente em final de palavra, ser **abreviadas**. Essa diminuição da quantidade vocálica registraremos com o uso da bráquia <~>. Por exemplo, فِي بَيْتٍ |fī baytⁱⁿ| {numa casa}, mas فِي الْبَيْتِ |fī 'l-baytⁱ| {na casa}.

Na ortografia árabe, as partículas que constem de um único grafema são escritas unidas à palavra seguinte. Na romanização, elas serão ligadas por um hífen — a exemplo de como se registram alguns clíticos em português¹⁶. Por exemplo, وَكُتُبٌ |wa-kutub| {e livros}.

¹⁴Lembramos que as consoantes adotadas preservam suas correspondências grafema–fonema independente de sua posição na palavra, ou seja, o <s> intervocálico mantém-se associado a /s/, i-e, não passa a representar /z/.

¹⁵Para <R>, como para <s>, a correspondência grafema–fonema mantém-se mesmo entre vogais; o <RR> indica geminação.

¹⁶Essa diferença de registro é análoga à que se pode observar entre espanhol e português, como neste caso: <cuéntame> versus <conte-me>.

Da mesma forma, os **suffixos pronominais** —que podem indicar um objeto oblíquo ou ter uma função genitiva— são originalmente escritos unidos ao verbo ou ao nome a que se relacionam; serão romanizados separados por um hífen. Por exemplo, <ضربني> |ḍaraba-nī| {[ele] golpeou-me} e <كتابي> |kitāb-ī| {meu livro}.

5.2.8.3 ao |al-|

A romanização do artigo definido <ال> |al-| possui algumas particularidades. Na escrita árabe, ele é grafado unido à palavra que define —substantivo e/ou adjetivo (2.1.2, p. 60)— e permanece graficamente inalterado —salvo quando precedido pela partícula <ل> |li| {para} (vide 2.2.7.2, p. 75).

Em virtude da alta frequência de uso do artigo definido, suprimimos a indicação da parada glotal, i-e, <ال> /ʔal/ será romanizado não |ʔal|, mas |al|. E, assim como os demais clíticos (vide 5.2.8.2, p. 138), a grafia original (unida à palavra seguinte) será indicada pelo uso do hífen¹⁷ Por exemplo, <البيت> |al-bayt| {a casa}.

Na leitura fluente, em muitas situações elidem-se o ataque e o núcleo /ʔa/. Para indicar esse processo, usamos o apóstrofo <'>. Por exemplo, <في البيت> |fī 'l-bayt| {na casa} e <بِالبيت> |bi-'l-bayt| {com a casa} (vide 2.2.7.2, pp. 75–76). Não confundir com o valor que muitas romanizações do árabe atribuem ao apóstrofo <'>, nelas usado para representar /ʔ/:

apóstrofo Trata-se de uma marca de elisão em inglês, francês, italiano e em muitas outras línguas. [...] o apóstrofo frequentemente serve como um sinal de parada glotal. (BRINGHURST, 2005, p. 332)

Quando a palavra a que se liga começa com uma letra solar (vide 2.2.5.1, p. 70), o fonema /l/ sofre assimilação regressiva do fonema representado por essa letra. Isso é marcado opcionalmente com a <ʿ> |xaddaʔ|,

¹⁷ Ao hífen, portanto, fica reservada a função de separar os clíticos das palavras a que eles se associem; e de indicar-lhes a grafia original unida.

colocada sobre a letra solar. Na romanização, substituímos o ⟨L⟩ do artigo definido pelo grafema que representa a letra solar. Por exemplo, ⟨الشمس⟩ |ax-xams| {o sol}.

5.3 A romanização acadêmica estrita

Da RAA podemos derivar um sistema de romanização que, com base nos mesmos princípios, permita registrar mais elementos da ortografia do AP. Em outras palavras, trata-se de uma forma mais próxima da transliteração, em que se procura estabelecer, entre os grafemas da escrita de partida e os da de chegada, um tipo de relação que possibilite recuperar a grafia original. Para esse fim, devemos implementar algumas modificações¹⁸.

5.3.1 Linha de base

Da RAA a RAE difere principalmente no registro dos grafemas árabes não-opcionais —essencialmente, os que representam consoantes, semi-vogais e vogais longas. Todos são escritos sobre a **linha de base**, ou seja, tanto ⟨^c⟩ /ʕ/ quanto ⟨^q⟩ /ʔ/ passam a ser registrados ⟨c, q⟩. Isso ratifica a necessidade de termos selecionado grafemas a partir de letras, não de outros sinais gráficos.

E, uma vez que as vogais breves serão grafadas sobrescritas (5.3.2), as vogais longas não necessitarão do macron ⟨[~]⟩, sendo grafadas ⟨a, i, u⟩.

5.3.2 Sobrescrito

As vogais breves, a nunação, o ta^q marbūṭa^t e uns poucos registros do ^qalif são registrados sobrescritos. No caso das vogais breves, são todas assim registradas, não somente as que marcam caso (*vide* 2.1.2, p. 60; e tab. 18, p. 73).

¹⁸Os exemplos serão mostrados em 5.7 (p. 146).

5.3.3 Subescrito

Na RAA registra-se a hamza^t ⟨ء⟩ simplesmente com um ⟨^q⟩. No entanto, essa romanização não permite distinguir-lhe o suporte (*vide* tab. 15, p. 72). Na RAE, por sua vez, designamos grafemas subscritos para indicar se a hamza^t —neste sistema romanizada ⟨q⟩— está apoiada em ^qalif ⟨إ!⟩, yā^q ⟨ئ⟩ ou wāw ⟨ؤ⟩. Para o ^qalif, selecionamos ⟨_a⟩; para yā^q e wāw selecionamos ⟨_i⟩ e ⟨_u⟩¹⁹.

O mesmo conceito será aplicado ao tā^q marbūṭa^t ⟨ة⟩. Para indicar que se trata de um ⟨o⟩ com os dois pontos do ⟨ت⟩, ele será grafado ⟨_ht⟩.

5.3.4 Caixa-baixa e caixa-alta

Em vez de, como na RAA, seguir a ortografia do português e aplicar caixa-alta (maiúsculas) à primeira letra de um período ou à de nomes próprios, na RAE grafa-se, basicamente, em caixa-baixa (minúsculas). Às versais reserva-se a função de indicar a gemação das consoantes.

5.3.5 Outros

Duas distinções não se fazem na RAE. Considerando-se que o objetivo principal de usar uma transliteração é grafar possibilitando recuperar a ortografia da escrita de partida, não separaremos os clíticos das palavras a que porventura estejam ligados. Pelo mesmo motivo não indicaremos, com a bráquia ⟨˘⟩, a redução na quantidade das vogais longas, em determinados casos.

5.4 A romanização leiga comum

A RLC, segundo o próprio nome sugere, tem como público-alvo as pessoas que não têm conhecimento do idioma árabe ou que com ele têm

¹⁹No abjad árabe, yā^q e wāw podem representar vogais longas ou semivogais. A fim de manter uniforme a representação do suporte da hamza^t, optamos por selecionar as variantes que tivessem um ponto em comum entre ^qalif, yā^q e wāw.

contato que não exija os pormenores da RAA nem a minúcia da RAE. As correspondências grafema–fonema mantêm-se basicamente as mesmas, porém valendo-se de uma forma de produção mais simples e direta; ou para, como KIRSCHBAUM *et al.* (2009) escrevem, em sua proposta de romanização do hebraico moderno, «[p]ermitir a escrita de palavras transliteradas em processadores de texto usuais, sem necessidade de recorrer a artifícios.»

5.4.1 Grafemas compostos

Na romanização de fonemas familiares (*vide* 5.2.4, p. 125) e na de não-familiares (*vide* 5.2.5, p. 127), a RAA emprega grafemas simples derivados (uma letra com diacrítico). Conquanto permitam eles estabelecer correspondências um a um com os grafemas e/ou fonemas que representam, esses diacríticos —ou os caracteres que os contenham— podem ser de utilização trabalhosa.

Meio mais fácil de obter as mesmas correspondências, o uso de grafemas compostos (duas ou mais letras) mostra-se uma opção mais prática com o benefício extra de manter as relações observadas, com as respectivas contrapartes, nos fonemas familiares (tab. 59, p. 128) e nos não-familiares (tab. 61, p. 131).

Para os fonemas familiares, todos fricativos, formaremos os grafemas compostos em ⟨H⟩ (tab. 68). Além de ser uma letra comum e familiar na composição de dígrafos²⁰, ela é usada nos esquemas e sistemas estudados (*vide* 4.2, pp. 103–117).

Para os fonemas não-familiares, todos guturais, formá-lo-emos em ⟨X⟩ (tab. 69). Apesar de ⟨Z⟩ ser relativamente familiar —por ser usado na ortografia do húngaro e do polonês—, preterimos esta opção em favor de uma combinação que dê menos margens a interpretações errôneas —como a de que ⟨DZ⟩ representaria a africada /dz/, em vez da faringalizada

²⁰Em português, forma ⟨CH, LH, NH⟩; em inglês, ⟨CH, SH, TH⟩ e ⟨ZH⟩, usado na romanização de /z/ em nomes estrangeiros (e.g., ⟨Zhivago⟩).

/dʰ/. A seleção de ⟨x⟩ inspira-se no «sistema x» do esperanto²¹.

No caso de um grafema composto estar associado a uma consoante geminada, a reduplicação será indicada pela repetição somente da primeira letra²². Isso evita produzir sequências como ⟨DHXDHX⟩, havendo, em seu lugar, ⟨DDHX⟩ (vide 5.7.1, p. 149).

Tabela 68: Romanização de fonemas familiares — grafemas compostos.

plosivas		fricativas			
rom.	fon.	graf. ár		fon.	rom.
T	t	ت	ث	θ	TH
D	d	د	ذ	ð	DH
K	k	ك	خ	x	KH
(G)	g	–	غ	ɣ	GH

[Rom]anização; [fon]ema; [graf]ema [ár]abe.

5.4.2 O ponto médio

Apesar de serem de produção menos trabalhosa, os grafemas compostos apresentam a desvantagem de gerar ambiguidades — especialmente no árabe, idioma em que os encontros consonantais são relativamente irrestritos.

²¹Em esperanto, os grafemas ⟨ĉ, ĝ, ĥ, ĵ, ŝ⟩, devido à dificuldade de produção de sua composição pouco usual, podem ser representados por grafemas compostos em H ⟨CH, GH, HH, JH, SH⟩ ou em x ⟨CX, GX, HX, JX, SX⟩, tendo este sistema a vantagem de empregar uma letra sem valor grafemático, ao contrário do H.

²²Esse procedimento baseia-se parcialmente no sistema Hepburn de romanização do japonês. Comparem-se estas duas romanizações: ⟨ まち ⟩ /matʃi/ |machi| {cidade}; ⟨ マッチ ⟩ /matʃi/ |matchi| {palito de fósforo}. No segundo caso, a geminação de ⟨CH⟩ não é ⟨CHCH⟩, mas ⟨TCH⟩. O uso de t, em vez de c —como se poderia esperar—, deve-se ao modo como o sistema Hepburn romaniza determinados alofones em distribuição complementar.

Tabela 69: Romanização de fonemas não-familiares — grafemas compostos.

plosivas e fricativas			dorsais e radicais		
rom.	fon.	graf·ár	fon.	rom.	
T	t	ت ط	t ^ʕ	TX	
D	d	د ض	d ^ʕ	DX	
K	k	ك ق	q	KX	
DH	ð	ذ ظ	ð ^ʕ	DHX	
S	s	س ص	s ^ʕ	SX	
H	h	ه ح	h	HX	

[Rom]anização; [fon]ema; [graf]ema [ár]abe.

A fim de evitar que um grafema composto seja confundido com uma sequência de dois grafemas (simples e/ou compostos), recorreremos a uma solução de uso corrente no catalão: o ponto médio ⟨·⟩. Esse sinal é usado

quando um *l* junta-se a outro (em catalão, assim como em espanhol, o *ll* é tratado como uma letra única: quando um *l* é adjacente mas separado de outro, escreve-se *l·l*. Exemplos disso são as palavras catalãs *cel·les* [células], *col·lecció* [coleção] e *paral·lel* [paralelo]. (BRINGHURST, 2005, p. 345)

E pode ser facilmente obtido tanto no sistema operacional WINDOWS, usando o comando + , quanto no MAC OS X, usando + + ²³.

5.5 A romanização leiga simplificada

A RLS é a forma mais simplificada da RAA. Não se usam grafemas derivados (com diacríticos) nem compostos (dígrafos e trígrafos). Consequentemente, por haver somente grafemas simples primitivos, não é

²³ Agradecimentos ao colega Daniel Oliveira Peres, que testou e confirmou o comando.

possível selecionar grafemas exclusivos para todos os grafemas e/ou fonemas árabes. Isso afeta a representação tanto dos fonemas familiares (tab. 70) quanto a dos não-familiares (tab. 71), pois esses grafemas passam a representar de dois a quatro diferentes fonemas (tab. 72).

Tabela 70: Romanização simplificada de fonemas familiares.

plosivas		fricativas			
rom.	fon.	graf.ár		fon.	rom.
T	t	ت	ث	θ	T
D	d	د	ذ	ð	D
K	k	ك	خ	x	K
(G)	g	–	غ	ɣ	G

[Rom]anização; [fon]ema; [graf]ema [ár]abe.

Tabela 71: Romanização simplificada de fonemas não-familiares.

plosivas e fricativas			dorsais e radicais		
rom.	fon.	graf.ár	fon.	rom.	
T	t	ت ط	t ^ʕ	T	
D	d	د ض	d ^ʕ	D	
K	k	ك ق	q	K	
D	ð	ذ ظ	ð ^ʕ	D	
S	s	س ص	s ^ʕ	S	
H	h	ه ح	ħ	H	

[Rom]anização; [fon]ema; [graf]ema [ár]abe.

Tabela 72: Ambiguidades geradas pelo uso de um mesmo grafema para diferentes fonemas.

graf.	fonemas representados			
	não-guturais		guturais	
	plos.	fric.	plos.	fric.
T	t	θ	t ^ʕ	
D	d	ð	d ^ʕ	ð ^ʕ
K	k	x	q	
S		s		s ^ʕ
H		h		ħ

[Graf]emas; [plos]ivas; [fric]ativas.

5.6 Resumo das propostas

A tabela 73 (p. 147) apresenta todas as romanizações das letras solares e das lunares. A tabela 74 (p. 148), os grafemas árabes, em todas as suas formas, juntamente com seus nomes e acompanhados de todas as romanizações.

5.7 Exemplos de aplicação

Nas páginas seguintes (149–159), mostraremos alguns exemplos de aplicação dos quatro sistemas de transposição ortográfica do árabe padrão para o português brasileiro. As eventuais melhorias, ainda por ser estudadas, serão abordadas no capítulo 6 (pp. 160–165).

Tabela 73: Letras solares e lunares (consoantes e semivogais).

modo	letras solares						letras lunares					
	graf.	fon.	romanização				graf.	fon.	romanização			
			RAA	RAE	RLC	RLS t.			RAA	RAE	RLC	RLS t.
plos.	ت	t	t			T	ب	b	b			
	ط	t ^s	ṭ		tx	t T	ك	k	k			κ
	د	d	d			D	ق	q	ḳ		kx	k κ
	ض	d ^s	ḍ		dx	d D						
fric.	ث	θ	ṯ		th	t T	ف	f	f			
	ذ	ð	ḏ		dh	d D	ج	ʒ	j			
	ظ	ð ^s	ḏ		dhx	d D	خ	x	ḵ		kh	k κ
	س	s	s			s	غ	ɣ	ḡ		gh	g
	ص	s ^s	ṣ		sx	s S	ه	h	ḥ			H
	ز	z	z				ح	ħ	ḥ		hx	h H
	ش	ʃ	x									
nas.	ن	n	n				م	m	m			
etc.	ر	r	r				ع	ʕ	⟨ ^c ⟩	c	–	–
	ل	l	l				ء	ʔ	⟨ ^q ⟩	q	–	–
							ا	–	–	–	–	–
							ي	j	y			
							و	w	w			

[Graf]ema, [fon]ema, [t]ipo de grafema básico (letra). [Plos]ivo, [fric]ativo, [nas]al. Nas colunas de RAE, RLC e RLS, as células em branco indicam que se usa o o mesmo grafema da RAA. Os colchetes angulares ⟨ ⟩ são usados apenas para melhor indicar o posicionamento dos grafemas sobrescritos.

Tabela 74: Grafemas do abjad árabe — nome, fonema e romanizações.

formas				nome	fon.	romanizações			
final	med.	inic.	isol.			RAA	RAE	RLC	RLS
ا			ا	ʿalif	a:	Ā/ā	a	aa	a
ب	ب	ب	ب	bā ^q	b	B/b			
ت	ت	ت	ت	tā ^q	t	T/t			
ث	ث	ث	ث	ṯā ^q	θ	T̤/t̤		TH/th	t
ج	ج	ج	ج	jīm	ʒ	J/j			
ح	ح	ح	ح	ḥā ^q	ħ	H/h		HX/hx	H/h
خ	خ	خ	خ	kā ^q	x	K/ḵ		KX/kx	K/k
د			د	dāl	d	D/d			
ذ			ذ	ḏāl	ð	Ḍ/ḏ		DH/dh	D/d
ر			ر	rā ^q	r	R/r			
ز			ز	zayn	z	Z/z			
س	س	س	س	sīn	s	S/s			
ش	ش	ش	ش	xīn	ʃ	X/x			
ص	ص	ص	ص	ṣād	s ^ʕ	Ṣ/ṣ		SX/sx	S/s
ض	ض	ض	ض	ḏād	d ^ʕ	Ḍ/ḏ		DX/dx	D/d
ط	ط	ط	ط	ṭā ^q	t ^ʕ	Ṭ/ṭ		TX/tx	T/t
ظ	ظ	ظ	ظ	ḏā ^q	ð ^ʕ	Ḍ̣/ḏ̣		DHX/dhx	D/d
ع	ع	ع	ع	ʿayn	ʕ	⟨ ^{C/c} ⟩	c	–	–
غ	غ	غ	غ	ḡayn	ɣ	Ġ/ḡ		GH/gh	G/g
ف	ف	ف	ف	fā ^q	f	F/f			
ق	ق	ق	ق	kāf	q	K/ḵ		KX/kx	K/k
ك	ك	ك	ك	kāf	k	K/k			
ل	ل	ل	ل	lām	l	L/l			
م	م	م	م	mīm	m	M/m			
ن	ن	ن	ن	nūn	n	N/n			
ه	ه	ه	ه	hā ^q	h	H/h			
و			و	wāw	w	W/w			
و			و	wāw	u:	Ū/ū	u	UU/uu	U/u
ي	ي	ي	ي	yā ^q	j	Y/y			
ي	ي	ي	ي	yā ^q	i:	Ī/ī	i	II/ii	I/i
			ء	hamza ^t	ʔ	⟨ ^{Q/q} ⟩	q	–	–

[Med]ial, [inic]ial, [isol]ada; [fon]ema. Nas colunas das formas, registraram-se as variantes conforme sua ligação (ou não) com grafemas adjacentes, não segundo sua posição na palavra. Nas colunas de RAE, RLC e RLS, as células em branco indicam que se usa o o mesmo grafema da RAA. Os colchetes angulares ⟨ ⟩ são usados apenas para melhor indicar o posicionamento dos grafemas sobrescritos.

5.7.1 Geminação

Tabela 75: Exemplos de romanização — consoantes geminadas.

	grafema primitivo		grafema derivado	
trad.	quebrar	triturar	ver; olhar	comparar
árabe	كَسَرَ	كَسَّرَ	نَظَرَ	نَظَّرَ
tr·fon	kasara	kas:ara	naðˤara	naðˤ:ara
rom·com	kasara	kassara	naẓara	naẓẓara
RAA	kasara	kassara	naḏara	naḏḏara
RAE	k ^a s ^a r ^a	k ^a s ^a r ^a	n ^a ḏ ^a r ^a	n ^a ḏ ^a r ^a
RLC	kasara	kassara	nadhxara	naddhxara
RLS	kasara	kasara	nadara	nadara

[Trad]ução; [tr]anscrição [fon]ológica; [rom]anização [com]um.

Figura 1: O registro das consoantes e das vogais breves (a); e o registro, na RAE, da consoante geminada (b).

ا	ا:	Ā
ب	ب	B
ت	ت	T
ث	ث	Ṭ
ج	ج	J
ح	ح	H
خ	خ	K
د	د	D
ذ	ذ	Ḍ
ر	ر	R
ز	ز	Z
س	س	S
ش	ش	X
ص	ص	Ṣ
ض	ض	Ḍ
ط	ط	Ṭ
ظ	ظ	Ḍ
ع	ع	ʿ
غ	غ	Ġ
ف	ف	F
ق	ق	Q
ك	ك	K
ل	ل	L
م	م	M
ن	ن	N
ه	ه	H
و	و:	Ū
و	و	W
ي	ي:	Ī
ي	ي	Y
ء	ء	ʿ
ة	ة	Ṭ

5.7.2 ^Qalif nuloTabela 76: Exemplos de romanização — correspondência nula de ⟨^Q⟩.

	suporte (acus.)		final (3·m·p)	
trad.	casa	uma casa	escrever	escreveram
árabe	بَيْت	بَيْتاً	كَتَبَ	كَتَبُوا
tr·fon	bajt	bajtan	kataba	katabu:
rom·com	bayt	baytan	kataba	katabū
RAA	bayt	bayt ^{an}	kataba	katabū
RAE	b ^a yt	b ^a yt ^a _a ^{an}	k ^a t ^a b ^a	k ^a t ^a b ^a _a
RLC	bayt	baytan	kataba	katabuu
RLS	bayt	baytan	kataba	katabu

[Acus]ativo; [3] terceira pessoa; [m]asculino [p]lural.

[Trad]ução; [tr]anscrição [fon]ológica; [rom]anização [com]um.

Figura 2: O ^Qalif nulo — valor meramente ortográfico.

ا	a:	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	ʃ	X
ص	sʕ	Ṣ
ض	dʕ	Ḍ
ط	tʕ	Ṭ
ظ	ḏʕ	Ḍ
ع	ʕ	ʕ
غ	ɣ	Ġ
ف	f	F
ق	q	Q
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	ʔ
ة		Ṭ

5.7.3 ^Qalif sobrescritoTabela 77: Exemplos de romanização — maddat^t e ^Qalif ḥanjariyya^t.

	maddat ^t	^Q alif ḥanjariyya ^t	
trad.	alcorão	este	Deus
árabe	القرآن	هَذَا	الله
tr·fon	ʔalqurʔa:n	ha:ða:	ʔal:a:h
rom·com	al-qur'an	hadha	Allah
RAA	al-ḵur ^Q ān	hādā	Al-lāh
RAE	_a q ^a lḵ ^u r _a qān	h ^A ḏa	_a q ^a l ^l h ^A
RLC	al-kxuraan	haadhaa	Al-laah
RLS	al-kuran	hada	Al-lah

[Trad]ução; [tr]anscrição [fon]ológica;
[rom]anização [com]um.

Figura 3: Maddat^t (a) e ^Qalif sobrescrito (b).

ا	a:	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḑ
ر	r	R
ز	z	Z
س	s	S
ش	ʃ	X
ص	s ^ʕ	Ṣ
ض	d ^ʕ	Ḍ
ط	t ^ʕ	Ṭ
ظ	ḏ ^ʕ	Ḑ
ع	ʕ	Ḑ
غ	ɣ	Ġ
ف	f	F
ق	q	Q
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	Q
ة		Ṭ

5.7.4 Hamza^tTabela 78: Exemplos de romanização — suporte da hamza^t.

	sem	com			
	ء	و	ئ	إ	أ
trad.	coisa	uma parte	vendedor, vendendo	nome	alif
árabe	شَيْء	جُزْء	بَائِع	إِسْم	أَلِف
tr-fon	ʃajʔ	ʒuzʔun	ba:ʔiʃun	ism	ʔalif
rom-com	shay	juz'un	bā'i'un	ism	alif
RAA	xay ^q	juz ^{qun}	bā ^q i ^{cun}	^q ism	^q alif
RAE	x ^a yq	j ^u z _u q ^{un}	ba _i q ⁱ c ^{un}	a ^q ism	a ^q alif
RLC	xay	juzun	baaiun	ism	alif
RLS	xay	juzun	baiun	ism	alif

[Trad]ução; [tr]anscrição [fon]ológica; [rom]anização [com]um.

Figura 4: A hamza^t sem suporte (a), apoiada no ^qalif (b, c), no yā^q (d) e no wāw (e).

ا	a:	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	ʃ	X
ص	sʕ	Ṣ
ض	dʕ	Ḍ
ط	tʕ	Ṭ
ظ	ḏʕ	Ḍ
ع	ʕ	ʕ
غ	ɣ	Ġ
ف	f	F
ق	q	Q
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	ʔ
ة		Ṭ

5.7.5 Tā^q marbūṭa^tTabela 79: Exemplos de romanização —
tā^q marbūṭa^t.

	tā ^q marbūṭa ^t		ha ^q
trad.	árvore	moça	Deus
árabe	شَجَرَة	فتاة	الله
tr·fon	ʃaʒarah	fata:h	?al:a:h
rom·com	shajarah	fatāh	Allah
RAA	xajara ^t	fatā ^t	Al-lāh
RAE	x ^a j ^a r ^a _h t	f ^a ta _h t	a ^q al ^l Ah
RLC	xajarat	fataat	Al-laah
RLS	xajara	fata	Al-lah

[Trad]ução; [tr]anscrição [fon]ológica;
[rom]anização [com]um.

Figura 5: Registro da tā^q marbūṭa^t (a) *versus* hā^q final (b).

ا	a:	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	ʃ	X
ص	s ^ʕ	Ṣ
ض	d ^ʕ	Ḍ
ط	t ^ʕ	Ṭ
ظ	ḏ ^ʕ	Ḍ
ع	ʕ	Ḥ
غ	ɣ	G
ف	f	F
ق	q	Q
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	Q
ة		Ṭ

5.7.6 Assimilação

Tabela 8o: Exemplos de romanização — assimilação regressiva em <ال> |al-|.

tradução árabe tr·fon	sol	
	sol	o sol
	شمس	الشَّمْس
	ʃams	ʔaf:ams
rom·com	shams	al-shams
RAA	xams	ax-xams
RAE	x ^a ms	aq ^a l ^x x ^a ms
RLC	xams	ax-xams
RLS	xams	ax-xams

[Trad]ução; [tr]anscrição [fon]ológica;
[rom]anização [com]um.

Figura 6: Registro da assimilação de /l/, na RAA (a) e na RAE (b).

ا	a:	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ʒ	J
ح	ħ	H
خ	x	K
د	d	D
ذ	ð	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	ʃ	X
ص	s ^ʕ	Ṣ
ض	d ^ʕ	Ḍ
ط	t ^ʕ	Ṭ
ظ	ð ^ʕ	Ḍ
ع	ʕ	^c
غ	ɣ	Ġ
ف	f	F
ق	q	Q
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	Q
ة		Ṭ

5.7.7 Artigo definido e a hamza^t al-waṣl.Tabela 81: Exemplos de romanização
— hamza^t al-waṣl.

	casa		
trad.	casa	a casa	e a casa
árabe	بَيْت	الْبَيْت	وَالْبَيْت
tr·fon	bajt	?albajt	walbajt
rom·com	bayt	al-bayt	wa-albayt
RAA	bayt	al-bayt	wa-'l-bayt
RAE	b ^a yt	a ^q l ^a b ^a yt	w ^a _a 'l ^a b ^a yt
RLC	bayt	al-bayt	wa-'l-bayt
RLS	bayt	al-bayt	wa-'l-bayt

[Trad]ução; [tr]anscrição [fon]ológica;
[rom]anização [com]um.

Figura 7: Hamza^t waṣl.

ا	a:	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	š	X
ص	s ^ʕ	Ṣ
ض	d ^ʕ	Ḍ
ط	t ^ʕ	Ṭ
ظ	ḏ ^ʕ	Ḍ
ع	ʿ	^c
غ	ǧ	Ġ
ف	f	F
ق	q	Q
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	Q
ة		Ṭ

5.7.8 Artigo definido e suas elisões

Tabela 82: Exemplos de romanização — elisões em <ال> |al-|.

	homem			noite		
trad.	homem	o homem	p/o homem	noite	a noite	p/a noite
árabe	رَجُل	الرَّجُل	لِلرَّجُل	لَيْل	اللَّيْل	لِللَّيْل
tr·fon	raʒul	?ar:aʒul	li:r:aʒul	laɟl	?al:ayl	li:l:ayl
rom·com	rajul	al-rajul	li-l-rajul	layl	al-layl	li-l-layl
RAA	rajul	ar-rajul	li-'r-rajul	layl	al-layl	li-'l-layl
RAE	r ^a j ^u l	a ^q l ^r r ^a j ^u l	li ^r r ^a j ^u l	l ^a yl	a ^q l ^l l ^a yl	li ^l l ^a yl
RLC	rajul	ar-rajul	li-'r-rajul	layl	al-layl	li-'l-layl
RLS	rajul	ar-rajul	li-'r-rajul	layl	al-layl	li-'l-layl

[Trad]ução; [tr]anscrição [fon]ológica; [rom]anização [com]um.

Figura 8: Supressão ortográfica de grafemas do artigo definido..

ا	a:	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	ʃ	X
ص	s ^ʕ	Ṣ
ض	d ^ʕ	Ḍ
ط	t ^ʕ	Ṭ
ظ	ḏ ^ʕ	Ḍ
ع	ʕ	Ḍ
غ	ɣ	G
ف	f	F
ق	q	K
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	Q
ة		Ṭ

5.7.9 Sufixos pronominais

Tabela 83: Exemplos de romanização — sufixos pronominais.

	acusativo		genitivo			
trad.	ele me	ele o	para	para	a minha	a casa
árabe	golpeou	golpeou	mim	ele	casa	dele
tr·fon	ضَرَبَنِي	ضَرَبَهُ	لِي	لَهُ	بَيْتِي	بَيْتُهُ
	ðarabani:	ðarabahu	li:	lahu	bajti:	bajtuhu
rom·com	dharabani	dharabahu	ly	lahu	bayty	baytuhu
RAA	ðaraba-nī	ðaraba-hu	l-ī	la-hu	bayt-ī	bayt ^u -hu
RAE	ð ^a r ^a b ^a ni	ð ^a r ^a b ^a hu	li	l ^a h ^u	b ^a yt ⁱ	b ^a yt ^u h ^u
RLC	dharaba-nii	dharaba-hu	l-ii	la-hu	bayt-ii	baytu-hu
RLS	daraba-ni	daraba-hu	l-i	la-hu	bayt-i	baytu-hu

[Trad]ução; [tr]anscrição [fon]ológica; [rom]anização [com]um.

Figura 9: Sufixos pronominais no acusativo (a) e no genitivo (b, c).

ا	ā	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	š	X
ص	s ^ʕ	Ṣ
ض	d ^ʕ	Ḍ
ط	t ^ʕ	Ṭ
ظ	ḏ ^ʕ	Ḍ
ع	ʿ	Ḷ
غ	ǧ	Ġ
ف	f	F
ق	q	Q
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	Q
ة		Ṭ

5.7.10 Alguns nomes próprios

Tabela 84: Exemplos de romanização — nomes próprios.

	topônimos		antropônimo	n·próprio
árabe	صَنَعَاء	بَغْدَاد	مَعْمَرُ الْقَذَافِي	الْجَزِيرَة
tr·fon	s ^ʰ anʕa:ʔ	bayda:d	muʕam:ar ʔalqað:a:fi:	ʔalzazi:rah
rom·com	San‘a’	Bagdá	(várias)	Al-jazira
RAA	Şan ^c ā ^q	Bağdād	Mu ^c ammar Al-Ḳaddāfi	Al-Jazīra ^t
RAE	ş ^a ncaq	b ^a ğdad	m ^u c ^a M ^a r a ^q l ^k āḌafi	a ^q l ^j a ^z ir ^{at}
RLC	Sanaa	Baghdaad	Muammar Al-Kxaddhaafii	Al-Jaziira
RLS	Sana	Bagdad	Muamar Al-Kadafi	Al-Jazira

[N]ome próprio. [Trad]ução; [tr]anscrição [fon]ológica; [rom]anização [com]um.

Figura 10: O nome da capital do Iêmen.

Figura 11: O nome do chefe de Estado da Líbia.

ا	a:	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	ʃ	X
ص	s ^ʰ	Ṣ
ض	d ^ʰ	Ḍ
ط	t ^ʰ	Ṭ
ظ	ḏ ^ʰ	Ḍ
ع	ʕ	Ḥ
غ	ɣ	G
ف	f	F
ق	q	Ḳ
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	ʕ
ة		Ṭ

5.7.11 O título

O título «A tradução da escrita — sistema de transposição ortográfica do árabe padrão para o português brasileiro» na RAA (tab. 85) e na RAE (tab. 86).

Tabela 85: Exemplos de romanização — o título (RAA).

árabe	romanização
تَرْجَمَةُ الْكِتَابَةِ:	Tarjama ^t al-kitāba ^t :
نِظَامٌ إِمْلَائِيٌّ	niḍām ʿimlāʿiyy
لِلنَّقْلِ الْحَرْفِيِّ لِلْعَرَبِيَّةِ	li-ʿn-naql al-ḥarfiyy li-ʿl-ʿarabiyya ^t
فِي اللُّغَةِ الْبُرتُغَالِيَّةِ الْبَرَازِيلِيَّةِ.	fī ʿl-luġa ^t al-burtuġāliyya ^t al-barāzīliyya ^t .

Tabela 86: Exemplos de romanização — o título (RAE).

árabe	romanização
تَرْجَمَةُ الْكِتَابَةِ:	t ^a rj ^a m ^{at} aq ^a lk ⁱ tab ^{at} :
نِظَامٌ إِمْلَائِيٌّ	n ⁱ ḍam aq ⁱ mla ⁱ q ⁱ Y
لِلنَّقْلِ الْحَرْفِيِّ لِلْعَرَبِيَّةِ	l ⁱ l ⁿ akl aq ^a lh ^a rfiY l ⁱ lc ^a rb ⁱ Y ^{at}
فِي اللُّغَةِ الْبُرتُغَالِيَّةِ الْبَرَازِيلِيَّةِ.	fī ʿl ⁱ l ^u ġ ^{at} aq ^a lb ^u rt ^u ġal ⁱ Y ^{at} aq ^a lb ^a razil ⁱ Y ^{at} .

ا	ā	Ā
ب	b	B
ت	t	T
ث	θ	Ṭ
ج	ǰ	J
ح	ḥ	H
خ	x	K
د	d	D
ذ	ḏ	Ḍ
ر	r	R
ز	z	Z
س	s	S
ش	š	X
ص	s ^ʿ	Ṣ
ض	d ^ʿ	Ḍ
ط	t ^ʿ	Ṭ
ظ	ḏ ^ʿ	Ḍ
ع	ʿ	C
غ	ġ	G
ف	f	F
ق	q	Q
ك	k	K
ل	l	L
م	m	M
ن	n	N
ه	h	H
و	u:	Ū
و	w	W
ي	i:	Ī
ي	j	Y
ء	ʔ	Q
ة		T

CAPÍTULO 6

Considerações finais

Adhuc tua messis in herba est.

«Ainda verde tua seara está.»

Ovídio, Heroides

CONSIDERAMOS algumas das atuais limitações e das futuras possibilidades de desenvolvimento referentes ao sistema proposto. Das propostas verificamos o que precisa ser consolidado, especialmente na representação de determinados fenômenos. Dos testes e exemplos em que pudemos aplicar as romanizações —particularmente a RAA e a RAE—, levantamos o que pudemos perceber acerca de algumas questões, para as quais procuramos apontar soluções e possíveis alternativas. Do restante analisamos o que falta abordar, em termos de contextos de aplicação, e para quais áreas mais podemos expandir, em relação ao escopo de aplicação.

6.1 Letras

Na RAA, ⟨Ḍ⟩ e ⟨Ḑ⟩ fogem a dois princípios da proposta: uso consistente dos diacríticos; uso de caracteres presentes no UNICODE. Trata-se, a bem da verdade, de uma adaptação que parte da necessidade de representar /ð/ com os grafemas básicos e os dígrafos de que dispomos (*vide* 5.2.5, página 128). De qualquer modo, uma solução seria usar um grafema composto adotado por alguns esquemas e sistemas de romanização: ⟨ẓ⟩. Além de ser um caractere existente no UNICODE, ele libera o uso de ⟨ḏ⟩ para representar /ð/.

Ainda na RAA, há uma questão que lida com o grau de representação da romanização. Trata-se de registrar ou não as supressões que se operam sobre o ال |al-| quando precedido pela preposição ⟨لِ| |li| {para} em casos como ⟨لِلرَّجُلِ| |li-'r-rajul| e ⟨لِلَّيْلِ| |li-'l-layl| (*vide* tabela 82, p. 156). Com vistas a preservar a identidade ortográfica do artigo definido romanizado, optamos por manter o conjunto ⟨'l-⟩ mesmo quando, na escrita de partida, ele se encontra diminuído de um ou de ambos os grafemas, uma vez que se trata de um fenômeno da ortografia do árabe, apenas. Na RAE, por outro lado, registramos essas supressões, posto que esta romanização visa a registrar elementos e fenômenos ortográficos.

Uma distinção que não contemplamos é a da oposição /l, l̥/ (JUBRAN, 1996). Ela é possível tanto na RAA quanto na RAE: ⟨l, l̥⟩. E, de certo modo, na RLC: ⟨lx⟩. Além desse par mínimo, no dialeto caiota identificam-se outros: /r, r̥/; m, m̥/; b, b̥/ (WATSON, 2002, pp. 42–43, 274–276). Todas essas distinções podem ser registradas¹ —usando-se o mesmo diacrítico das demais consoantes faringalizadas, o que lhe atesta a consistência e a produtividade— restando verificar-lhes a necessidade e a abrangência.

Nas duas romanizações leigas (RLC e RLS), o registro do ʿayn ⟨ع⟩ e da hamza^t ⟨ء⟩ foi abolido, pelo menos nos exemplos. Sendo necessário ou desejável registrá-los, uma solução terá de ser elaborada.

Faz-se necessária, também, uma solução que permita indicar o acento tônico dos sintagmas romanizados.

¹Na RAA e na RAE, ⟨Ṛ ṘḐ⟩; na RLC, ⟨RX, MX, BX⟩.

caiota: do Cai

6.2 Palavras

O resumo romanizado (p. x) foi o primeiro teste maior de nossa proposta; nele pudemos aplicar a RAA em mais do que palavras isoladas e pequenos sintagmas. Essa nova dimensão gerou algumas situações novas.

Um ponto que não havíamos previsto é o da translineação, i.e, como seria feita a mudança de linha, caso fosse necessário manter parte de uma unidade ortográfica em uma linha e continuar na seguinte. Uma opção seria simplesmente mover a unidade inteira; o problema seriam os grandes espaços em branco resultantes da justificação de sucessivas linhas em que se não fizesse a translineação. Outra opção seria permitir a translineação nos pontos em que houvesse fronteira de morfemas, mas persistiria o problema da justificação das linhas, uma vez que esses morfemas tendem a ser monossilábicos. Uma terceira opção seria realizar a translineação independentemente das fronteiras morfológicas, o que resolveria o problema das linhas justificadas.

Mas isso levou a esta outra questão: como registrar a translineação. Se fosse usado o hífen em todos os casos, em algumas situações não ficaria claro se a quebra realizara-se em fronteira de morfemas ou não. Para solucionar esse impasse, nos casos de fronteira de morfemas adotamos o hífen repetido na linha seguinte²; nos demais casos adotamos o travessão baixo (_). Com isso, a translineação pôde ser aplicada tanto em fronteira de morfemas quanto em outras posições.

Um último caso, ainda que pouco frequente, seria o da translineação das consoantes geminadas. Diferentemente do que se faz em português, optamos por não separar o grafema composto (linhas 16–17 do resumo romanizado, p. x).

²Em português também há tal recurso, que permite translinear, por exemplo, (sentem-se). Entretanto, isso parece ser de uso corrente apenas em Portugal, tendo caído em desuso no Brasil.

6.3 Nomes

No caso do ⟨al-⟩ preposto a nomes próprios, o uso de caixa-alta ou caixa-baixa ainda não se encontra consolidado. As combinações possíveis são: acompanhar o núcleo do sintagma (e.g, Al-Ḳaddāfi); assumir a caixa-alta pelo sintagma todo (e.g, Al-Ḳaddāfi); manter-se em caixa-baixa (e.g, al-Ḳaddāfi). Cada uma dessas possibilidades encontra um análogo em sobrenomes escritos em letras latinas, por exemplo: Della Torre; Dellatorre; della Torre. E outro fator a considerar é sua posição, se precedido por outro nome ou não: Vincent van Gogh; Van Gogh.

Quanto à romanização dos antropônimos de figuras históricas, personalidades atuais e autores de publicações literárias —juntamente com os títulos de suas obras— ou acadêmicas, preferimos pouco abordá-la por ora. O sistema de JUBRAN (2004) prevê a romanização de alguns elementos comuns em antropônimos árabes, como estes: ⟨عبد⟩ |^cabd|; ⟨أبو⟩ |^Qabū|; ⟨ابن⟩ |^Qibn|. A partir desse modelo poderemos estudar melhor a questão.

E a romanização de topônimos não pode ser negligenciada; de maneira mais específica, à geonímia³ poderemos voltar maior atenção, em um trabalho vindouro, que abordará discussões acerca da padronização internacional de endônimos e exônimos.

6.4 Bibliotecas

Um dos princípios da presente proposta é empregar somente letras para constituir grafemas. A principal vantagem dessa opção é a possibilidade de proceder à alfabetação do ^cayn ⟨ع⟩ e da hamza^t ⟨ء⟩, que em alguns outros esquemas e sistemas de romanização encontram-se representados por sinais variados.

³A geonímia ocupa-se do estudo de referentes únicos, quer sejam eles topônimos, hidrônimos ou acidentes geográficos. Enquanto um mesmo topônimo pode estar associado a dois referentes, como ⟨São Paulo⟩ (cidade) e ⟨São Paulo⟩ (estado), um geônimo possui apenas um único referente.

Mas colocar em ordem alfabética pode não ser tão simples quanto colocar c depois de b e antes de d; ou q depois de p e antes de r. Um estudo mais detido será necessário para analisarmos não somente a alfabetação de ⟨^c q⟩, mas também a dos grafemas derivados, a dos grafemas compostos e a de alguns elementos a compor antropônimos. É preciso decidir se os grafemas derivados serão elencados após os grafemas primitivos (e.g. ⟨açude⟩ depois de ⟨acudir⟩) ou indistintamente (e.g. ⟨açude⟩ antes de ⟨acudir⟩); e se os grafemas compostos serão tratados como os derivados ou como sequências de duas letras⁴, e em que casos — em dinamarquês, por exemplo, na alfabetação não se decompõem os dígrafos, que são tratados como grafemas com posição específica, a respeitar, inclusive, distinções etimológicas ou afins⁵.

Deriva deste último caso o registro de abreviaturas e de variantes pertinentes a certos elementos a compor antropônimos. Embora seja uma prática esquecida, que só se observe em publicações mais antigas, já se abreviou ⟨Charles → Ch.⟩ e ⟨Theodoro → Th.⟩, para diferenciá-los de ⟨Carlos → C.⟩ e ⟨Teodoro → T.⟩. E, em escocês, as variantes ⟨Mac⟩ e ⟨Mc⟩ são tratadas ambas como ⟨Mac⟩, para efeito de alfabetação⁶. Se o mesmo princípio for aplicado ao árabe, pode ser necessário assumir ⟨إِبْن⟩ |^qibn| {filho de} e ⟨بِنْ⟩ |bin| como variantes a ser alfabetadas em uma mesma posição — outra questão a ser devidamente estudada.

6.5 Desdobramentos da proposta

Propusemo-nos a estabelecer um sistema de romanização do árabe padrão para o português brasileiro. Ao longo do trabalho de sua elaboração, percebemos-lhe o potencial para desdobrar-se em quatro sistemas: **romanização acadêmica ampla**, ou RAA, que estabelece princípios de

⁴Trata-se da diferença entre organizar ⟨KA-KI-KU-KH⟩ ou ⟨KA-KH-KI-KU⟩.

⁵A letra ⟨Å⟩, que pode ser grafada ⟨AA⟩, vem após o ⟨z⟩. Por isso, tanto ⟨Århus⟩ quanto ⟨Aarhus⟩ aparecem depois de ⟨z⟩; mas ⟨Aachen⟩ (cidade da Alemanha) vem em ⟨A⟩.

⁶Logo ⟨McKinley⟩ vem antes de ⟨MacKinney⟩; e ⟨McAlister⟩, antes de ⟨MacArthur⟩.

operação consistentes e produtivos, construindo a base para os três outros sistemas; **romanização acadêmica estrita**, ou RAE, que representa mais minuciosamente a ortografia da escrita de partida, servindo como forma de transliteração; **romanização leiga comum**, ou RLC, que faz uso de grafemas compostos, visando a uma maior facilidade de produção; **romanização leiga simplificada**, ou RLS, que nenhum uso faz de grafemas derivados ou compostos, sendo o mais simples dos quatro sistemas.

Com a experiência adquirida no desenvolvimento desses sistemas, existe a possibilidade de ampliarmos a abrangência dos princípios norteantes adotados, expandindo-lhes a aplicação a outros sistemas de escrita. E o Departamento de Letras Orientais é campo fértil para isso, pois todas as cadeiras são de línguas cujo sistema de escrita difere do latino. Além do árabe e do hebraico —ambas línguas semíticas—, encontramos armênio, chinês, coreano, japonês e russo. Se o chinês já conta com o pinyin e o japonês, com três sistemas de romanização⁷, cotejar-lhes as romanizações atuais e passadas pode-nos prover de um melhor entendimento acerca de como os sistemas de romanização são desenvolvidos e implementados, além de como e por que são suplantados.

Mas não temos a pretensão de ter abrangido a totalidade do escopo dos quatro sistemas desta pesquisa. Além das questões pontuais abordadas anteriormente, testes de aplicação ainda precisam ser efetuados, pois neles surgem impasses a evidenciar os limites e as limitações dos sistemas utilizados. Conhecidos os atuais limiares, poderemos trabalhar com vistas a aprimorar os quatro sistemas ora apresentados.

⁷São eles: Nihon-shiki (Nihon-siki); Kunrei-shiki (Kunreisiki); Hepburn, mencionado na p. 143 (nota de rodapé 22).

Posfácio

A reflexão é do pesquisador a fornalha do ferreiro;
são-lhe a palavra e a escrita o martelo e a bigorna.

do autor

TODA DISSERTAÇÃO é produto de uma reflexão — de pesquisa e de labor, mas, sobretudo, de reflexão. Forjamos esta dissertação no sutil equilíbrio entre a adesão a padrões e a quebra de paradigmas.

Reflexão

Nosso principal objetivo —ao selecionar o programa, planejar os tópicos, escolher os exemplos, etc.— sempre foi o de melhor apresentar as informações aqui contidas, em uma diagramação organizada e com um visual agradável aos olhos e à leitura. E aí se inclui a “desobediência civil” que permeia esta dissertação em diferentes estratos.

Normas

Procuramos seguir as normas de formatação, de diagramação, de citação e referência, entre outras. Havendo-lhes substitutos, entretanto, foram abandonadas ou subvertidas, caso fosse favorecida uma melhor experiência de leitura. De modo geral, tentamos delas não desviar tanto a ponto de no leitor causar espécie ou lhe dificultar o reconhecimento da informação pertinente.

A primeira norma de que divergimos foi a NBR 14724:2005, que em 5.2 diz que «as folhas devem apresentar margem esquerda e superior de 3 cm; direita e inferior de 2 cm». Posto que nenhum livro bem diagramado⁸ apresenta tal sorte de proporções⁹, preterimos essa norma em favor de margens mais adequadas¹⁰.

Outra norma foi a NBR 6027:2003, que em 5.4.4, alínea *b*, a respeito da paginação dos elementos do sumário, afirma poderem-se apresentar os «números das páginas inicial e final, separadas [*sic*] por hífen (exemplo: 91-143)». Novamente, diverge a ABNT de um padrão tipográfico; BRINGHURST (2005) diz textualmente que «[o] hífen é curto demais para funcionar assim» (p. 92) e recomenda o uso do travessão eno ⟨–⟩¹¹.

Por fim, em nosso contato com as normas APA e Vancouver, pudemos perceber-lhes as vantagens em alguns aspectos da apresentação de citações e referências. Da APA trazemos o ano de publicação subsequente aos nomes dos autores; da Vancouver, a ausência de vírgula e ponto(s) no nome dos autores. Nas referências bibliográficas, não usar pontos

⁸As dimensões da mancha tipográfica e sua relação com as margens dependem de variados fatores, sendo-lhe as proporções baseadas em diferentes procedimentos, que não nos cabe aqui pormenorizar.

⁹A bem da verdade, trata-se do exato oposto, sendo menores as margens interna (esquerda) e superior; e invariavelmente maiores as margens externa (direita) e inferior.

¹⁰Margem superior de 3cm; externa de 4cm; inferior de 5cm. A margem interior mede 5cm, já contados os 2cm reservados para a encadernação; os 3cm restantes foram calculados para abrigar a numeração das seções e alguns recursos que não nos foi possível implementar.

¹¹Em vez de ⟨91-143⟩, temos ⟨91–143⟩.

mancha tipográfica: o espaço ocupado, na página, pelo texto principal.

em línguas como dinka, eve e tuí. A intenção primeira era a de poder diferenciar, por meio outro que não o do contexto, pares de palavras como ⟨sede, sēde⟩ e ⟨corte, cōrte⟩; e de não se alterar significativamente a ortografia, mantendo-se os diacríticos porventura necessários: ⟨étimo, ótimo⟩. Mas a tarefa demasiado laboriosa e a interferência —ainda que passageira— na leitura restringiram-lhes a aplicação a este posfácio.

E, no texto como um todo, procuramos atribuir funções específicas a diferentes variantes tipográficas¹³, sinais de pontuação¹⁴ e notas (de rodapé e de margem). Pontuais variações e diversos reajustes foram necessários, em extenso número, mas em diminuta relevância. E, nesses casos, também houve de prevalecer a razão sobre os padrões por nós mesmo estabelecidos.

Em «nós mesmo» ocorre silepse de número.

L^AT_EX

Para preparar esta dissertação, escolhemos o L^AT_EX, por suas qualidades e vantagens. Dentre as tantas características consideradas, podemos citar:

- a separação entre conteúdo e apresentação, qual em HTML e CSS;
- os padrões tipográficos nativos;
- as adaptações e os ajustes finos que o sistema permite fazer;
- a capacidade de padronizar a apresentação de um sem-número de elementos (e.g, títulos, citações, etc.);
- a facilidade e o funcionamento automático do mecanismo de referências (internas e bibliográficas);
- uma muito boa capacidade de mesclar escritas dos mais variados tipos, inclusive esquerda–direita (e.g, português) com direita–esquerda (e.g, árabe).

HTML:
*hypertext
mark-up
language*,
usado para
estruturar
o conteúdo
das páginas
da *internet*.

CSS:
*cascading
style sheets*,
usado para
formatar
a aparência
das páginas
da *internet*

Apesar d'ele nem sempre funcionar a contento, o produto final facilmente compensa os inúmeros problemas e contratempos vivenciados.

¹³ Vide lista de variantes tipográficas (p. xxvi).

¹⁴ Vide lista de sinais de pontuação (p. xxv).

X₃TEX

E em conjunto usamos o X₃TEX, que é, *grosso modo*, uma forma de compilar o $\text{\TeX}$ necessária quando o *input* inclui caracteres UNICODE. Uma vez que trabalhamos extensivamente com diacríticos diversos e sistemas de escrita distintos, a escolha não poderia ser outra. Infelizmente, alguns recursos que pretendíamos utilizar neste trabalho não estão disponíveis¹⁵, notadamente dois de microtipografia: ajuste automático de espacejamento e entreletra, que permitiriam parágrafos mais bem justificados e menor quantidade de hifenações; e, principalmente, alinhamento óptico da margem, que tornaria as laterais dos parágrafos mais visualmente bem alinhadas¹⁶.

espacejamento
ou *letterspacing*
entreletra
ou *tracking*

Sumários

O leitor mais atento deve ter notado que usamos $\langle$ sumário de tabelas $\rangle$ e $\langle$ sumário de figuras $\rangle$ por $\langle$ lista de tabelas $\rangle$ e $\langle$ lista de figuras $\rangle$ — mas utilizamos $\langle$ lista de sinais de pontuação $\rangle$ e $\langle$ lista de variantes tipográficas $\rangle$. Nossa opção por $\langle$ sumário $\rangle$ em vez de $\langle$ lista $\rangle$ baseou-se em uma observação simples: o sumário relaciona as páginas; as listas, não. Entendemos ser o sumário, na verdade, a organização dos tópicos do texto, motivo pelo qual se usa $\langle$ lista $\rangle$ para a relação de tabelas e figuras. Entretanto, no intuito de marcar o que havíamos observado — e na falta de um termo mais específico —, por aproximação optamos por $\langle$ sumário $\rangle$, ainda que ele não seja plenamente apropriado.

A respeito do sumário propriamente dito, por abranger quatro níveis (capítulo, seção, subseção e subsubseção) e incluir todos os elementos pós-textuais, tornou-se ele muito extenso (sete páginas), o que poderia

¹⁵Em vez de X₃TEX, teríamos de compilar o texto usando pdf $\text{\TeX}$, que não trabalha com UNICODE.

¹⁶Uma amostra de como seria um parágrafo opticamente alinhado encontra-se nos resumos em português (p. vii) e em inglês (p. viii), alinhados manualmente, para emular o mesmo efeito.

resultar em uma consulta desnecessariamente demorada, em alguns casos. Para contornar esse inconveniente, decidimos antepor-lhe um sumário breve, de página única, contendo tão-somente os capítulos e os níveis mais altos dos elementos pós-textuais, todos seguidos das respectivas primeira e última páginas.

A propósito, nos sumários os números de página estão posicionados logo ao final de cada entrada. Isso visa a seguir um padrão tipográfico empregado nos livros mais bem diagramados — curiosamente, não se trata de um padrão do $\text{\LaTeX}$. Em vez de alinhar todos os números à direita —o que força o leitor a traçar uma linha imaginária ou a fazer de régua o objeto retilíneo mais próximo—, os números estão junto às entradas a que se referem, e em itálico, para que se não confundam com o título da entrada. Em termos tipográficos, o ideal seria que estivessem eles todos alinhados à esquerda, ao início de cada entrada. No entanto, testes realizados ainda no começo da diagramação deste trabalho mostraram que, por serem praticamente todos os títulos precedidos por números (e.g., 5.2.8.3), mesmo o itálico não bastava para que a apresentação ficasse visualmente aceitável. Acabamos optando pela solução de meio-termo.

Capítulos, seções, subseções e subsubseções

Faz parte da *praxis* tipográfica que o fólio da abertura de um capítulo seja baixado —a fim de não ocupar o claro de abertura—, sendo os seguintes posicionados no alto da página. Essa alternância parece-nos desnecessária, motivo pelo qual preferimos registrar baixados todos os fólhos.

Quanto à organização dos níveis, ela por vezes está imbuída de informação que não simplesmente a de hierarquia. Em alguns casos, as subdivisões formam uma continuação do título de nível imediatamente superior. Por exemplo:

2.2.2 Origens... da escrita; das consoantes; das vogais.

2.2.7 Ortografia... e morfologia; e suas regras; e o alif.

fólio: o número da página.

baixado: disposto ao pé da página.

claro de abertura ou <de entrada> é o espaço deixado intencionalmente entre o alto da página e o início do texto.

4 A romanização do árabe... nos jornais; na academia; em análise.

5.2.8 Alguns pontos pertinentes... à ortografia; à morfossintaxe;
ao |al-|.

Isso fica mais aparente no sumário, onde se podem visualizar essas relações uma a uma.

Fundição e prensa

A escolha dos tipos —especialmente Linux Libertine (serifado) e Linux Binolium (não-serifado), da fundição de Philipp H. Poll¹⁷ — foi fundamental. Precisávamos de uma fonte que tivesse suporte extensivo ao UNICODE, pois era essencial não ser necessário recorrer a outras fontes para usarmos caracteres com diacríticos. Outras características incluem numerais antigos e versaletes verdadeiros, que se combinam, ambos, melhor com o texto em caixa-baixa.

A escolha do papel —normalmente tão negligenciada— foi motivada pela leitura de um ensaio do tipógrafo Tschichold¹⁸ e a decorrente observação de que livros de texto (sem figuras) bem planejados via de regra não se imprimem em papel branco.

⟨fundição⟩
designa a empresa
ou pessoa que cria
tipos (fontes).
O termo remete
aos tipos móveis,
feitos de chumbo,
fabricados
em fundições.

Por fim...

Muito ainda poderia ser feito. Mas preferimos fechar este trabalho da forma como está. Acreditamos que a experiência aqui adquirida servirá para que a tese de doutorado seja melhor, na forma como no conteúdo.

¹⁷<http://www.linuxlibertine.org>

¹⁸«Papel de impressão: branco ou mate?» In: TSCHICHOLD J (2007) **A forma do livro: ensaios sobre tipografia e estética do livro**. Cotia, SP: Ateliê.

APÊNDICE A

Terminologia

Para que a descrição e a análise fossem tão precisas quanto possível, precisáramos adotar termos adequados ao tratamento do objeto de estudo em questão. Os termos relacionados a seguir agrupam-se, *grosso modo*, nestas categorias: letras e sinais gráficos; grafemas; combinações; grafema–fonema; alfabetos; transposição ortográfica.

A.1 Letras e sinais gráficos

Alguns dos menores elementos constituintes de um sistema ortográfico:

letra Unidade básica de um alfabeto ou de um abjad¹.

grafema Unidade básica² de um sistema ortográfico, podendo ser uma letra, uma letra acrescida de diacrítico(s) ou um conjunto de letras.

¹No caso do abjad árabe, as letras seriam as formas básicas de cada grupo de grafemas, como ⟨ب⟩ para ⟨ب, ت, ث⟩.

²Para evitarmos o uso constante de “grafema (ou dígrafo)”, o termo ⟨grafema⟩ poderá ser usado para referir tanto grafemas quanto dígrafos.

glifo Representação particular de um numeral ou um grafema.
Cada variante³ conta como um glifo separado.

diacrítico Sinal gráfico acrescido a uma letra, sendo-lhe colocado à volta⁴ ou dentro⁵, alterando-lhe o valor ortográfico ou constituindo um grafema independente⁶.

A.2 Grafemas

Em GRAFEMAS, os diferentes tipos de grafema, conforme usados nesta dissertação:

simples Formado por uma letra.

composto Formado por duas ou mais letras.

primitivo Sem diacríticos.

derivado Com diacríticos.

dobrado Formado pela reduplicação
de um grafema simples ou composto.

estendido Formado pela repetição de um dos elementos
de um grafema composto.

Os **grafemas simples primitivos** são a base dos grafemas derivados e dos grafemas compostos. Os **grafemas dobrados** distinguem-se

³Variantes são versões graficamente diferentes, mas ortograficamente indiferentes, como ⟨2, 4, 6⟩ *versus* ⟨2, 4, 6⟩ ou ⟨a⟩ vs. ⟨a⟩.

⁴Normalmente, sobre ou sob a letra. Mais raros são os casos em que o diacrítico é colocado ao lado —ou dentro, a depender da análise—, como no hebraico ⟨וְ⟩, em que ao VAV ⟨וְ⟩ [v] pode-se acrescentar um SHURUK ⟨שׁוּרוּק⟩ (para representar a vogal [u]) ou um DAGESH ⟨דָּגֵשׁ⟩ (para representar a consoante [v]).

⁵Em hebraico, por exemplo, o DAGESH ⟨דָּגֵשׁ⟩ é um ponto colocado no centro de algumas letras —principalmente nas BEGADKEFAT ⟨ב ג ד כ פ ת⟩.

⁶Em dinamarquês, norueguês e sueco, por exemplo, a vogal ⟨Å, å⟩ (um ⟨A⟩ com anel) ocupa posição própria na ordem alfabética, após a letra ⟨Z⟩.

dos compostos por serem apenas reduplicações de um grafema simples (primitivo ou derivado), sinalizando aumento na quantidade vocálica ou geminação da consoante, conforme o caso. Os **grafemas compostos** distinguem-se dos derivados apenas no modo adotado para indicar as diferenças entre o fonema que representam e o que é representado pelo grafema primitivo.

A.3 Combinações

Em COMBINAÇÕES, as maneiras como grafemas ou fonemas podem formar sequências:

dígrafo Conjunto de duas letras que representa um fonema.

trígrafo Conjunto de três letras que representa um fonema.

dífono Um grafema (simples ou composto) representa uma sequência de dois fonemas.

ligaduras Caracteres compostos por dois ou mais glifos. Podem ter função ortográfica, como o ESZETT ⟨ß⟩, formado a partir de ⟨f + s = ß⟩ (*vide* A.5.2, p. 177); função tipográfica, como ⟨fi⟩, formado a partir de ⟨f + i⟩ e usado para evitar colisão; ou função decorativa, como as ligaduras históricas ⟨ct⟩ e ⟨st⟩. ^{a que ocorre em} ⟨fi⟩

A.4 Grafema–fonema

Em GRAFEMA–FONEMA, os tipos básicos⁷ de relação entre grafemas e fonemas:

⁷As relações múltiplas, em que um grafema representa mais de um fonema ou um fonema pode ser representado por mais de um grafema, não contam com termo específico, sendo designadas por sua descrição, e-g, “um grafema, vários fonemas” ou “um fonema, vários dígrafos”.

biunívoco O grafema representa apenas um fonema, que é representado somente por aquele grafema.

unívoco O grafema representa um único fonema, que pode ser representado por outros grafemas; e *vice versa*, i-e, um certo fonema é representado por um só grafema, que pode representar mais fonemas.

A.5 Alfabetos

Em ALFABETOS, o modo como letras e sinais gráficos (e suas combinações) organizam-se em uma língua:

abjad Conjunto de letras ou grafemas que representam, essencialmente, fonemas consonantais, sendo as vogais marcadas opcionalmente.

alfabeto Conjunto de letras ou grafemas que representam fonemas vocálicos e consonantais.

ortografia Conjunto de regras e convenções de uma língua que estabelecem os grafemas adotados e que prescrevem como as palavras devem ser grafadas, estipulando, por exemplo, relações grafema–fonema fixas ou dependentes do contexto (e.g, início de palavra, posição de coda, intervocálico, pretônico ou postônico).

A.5.1 Abjad

Adotamos a distinção entre abjad e alfabeto estabelecida por DANIELS & BRIGHT (1996), qual seja, a de que:

[em] um consonantário, aqui chamado *abjad* à semelhança de “alfabeto” [...], os caracteres denotam consoantes (somente). Em um alfabeto, os caracteres denotam consoantes e vogais.⁸ (p.4)

devendo ser feita a ressalva de que o abjad árabe contém grafemas (ʔalif ⟨ا⟩, wāw ⟨و⟩ e yāʔ ⟨ي⟩) que podem representar semivogais ou vogais longas, conforme o contexto.

Sobre o termo ‘abjad’, DANIELS & BRIGHT (1996) afirmam que «a palavra é formada pelas quatro primeiras letras do exemplo mais conhecido, a escrita árabe, em sua ordem histórica»⁹. No entanto, uma consulta ao abjad árabe (*vide* tab. 74, p. 148) pode não confirmar essa afirmação, uma vez que os quatro primeiros grafemas são, na ordem, ʔalif ⟨ا⟩, bāʔ ⟨ب⟩, tāʔ ⟨ت⟩ e ṭāʔ ⟨ث⟩, formando a sequência A-B-T-T. Na verdade, deve-se tomar o primeiro grafema de cada uma das quatro primeiras formas básicas (⟨ا⟩, ⟨ب⟩, ⟨ح⟩ e ⟨د⟩), i.e, ʔalif ⟨ا⟩, bāʔ ⟨ب⟩, jīm ⟨ج⟩ e dāl ⟨د⟩, formando então a sequência A-B-J-D, que forma ⟨abjad⟩.

A.5.2 Alfabeto latino

O alfabeto latino, em sua forma mais difundida, contém 26 letras¹⁰. Algumas línguas empregam letras adicionais, como as consoantes THORN ⟨Þ, þ⟩ (islandês) e EDH ⟨Ð, ð⟩ (islandês e feroês), bem como as vogais E ABERTO ⟨Ê, ê⟩ e O ABERTO ⟨Ô, ô⟩, usadas em algumas línguas africanas (e.g, baule, krio e twi); outras adotam ligaduras, a exemplo do ESZETT ⟨ß, ß⟩¹¹ (alemão) e do ⟨Œ, œ⟩ francês.

⁸No original: «[in] a consonantary, here called *abjad* as a parallel to “alphabet” [...], the characters denote consonants (only). In an *alphabet*, the characters denote consonants and vowels.»

⁹No original: «the word is formed from the first letters of the most widespread example, the Arabic script, in their historic order».

¹⁰A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z

¹¹Trata-se de uma ligadura do s LONGO ⟨f⟩ com ⟨z⟩. Tem valor de ⟨ss⟩, podendo ser, em alguns casos, substituído por este grafema composto.

A.6 Transposição ortográfica

Em TRANSPOSIÇÃO ORTOGRÁFICA, as diferentes formas de transpor itens lexicais e gramaticais de um a outro sistema de escrita:

transposição ortográfica Serve como hiperônimo de todas elas, podendo ser usado nas situações em que mais de um procedimento tiver sido utilizado ou o processo não estiver bem definido.

transcrição Estabelece relações fonema–grafema biunívocas, permitindo recuperar a pronúncia original a partir da grafia na língua de chegada¹².

transliteração Estabelece relações grafema–grafema biunívocas, permitindo recuperar a escrita original a partir da grafia na língua de chegada¹³.

romanização Pode empregar um misto de transliteração e transcrição. O termo é específico no sentido, i.e, tem a escrita latina como sistema de chegada.

latinização Modifica e adapta o original, para conformá-lo aos moldes da língua de chegada —tanto na ortografia quanto na fonotática—, fazendo com que o produto final pareça-se com um elemento vernáculo¹⁴. Normalmente opera sobre itens lexicais, especialmente topônimos e antropônimos.

¹²Por exemplo, a palavra russa ⟨что⟩ {o que} compõe-se dos grafemas ⟨ч⟩–/tʃ/, ⟨т⟩–/t/ e ⟨о⟩–/o/, mas pronuncia-se /tʃto/, como se fosse ⟨што⟩. Uma transcrição para essa palavra (e.g, |shto|) deve-se aproximar da pronúncia, não da forma ortográfica.

¹³Por exemplo, em japonês os *kana* ⟨ ず ⟩ e ⟨ づ ⟩ representam /zu/, não permitindo a transcrição |zu| identificar os grafemas originais. Por sua vez, as transliterações ⟨ ず ⟩ |zu| e ⟨ づ ⟩ |du| possibilitam discernir um e outro.

¹⁴Por exemplo, ⟨Confúcio⟩ (vide 1.2.3, p. 47) parece-se menos com ⟨Chiang Kai-shek⟩ ou ⟨Máo Zédōng⟩ do que com ⟨Lucrécio⟩ ou ⟨Suplício⟩.

kana: grafemas que compõem os silabários *hiragana* e *katakana*.

ANEXO A

MUAMMAR GADDAFI OU MU'AMMAR AL-QADHDHAFI?

UMA POLÊMICA ÁRABE-BRASILEIRA

A Líbia, país do norte da África, está em crise. As forças rebeldes pedem a renúncia do governante líbio. O líder árabe, há 42 anos no poder, continua irredutível. Esse impasse projetou o nome do ditador mundo afora: MUAMMAR GADDAFI ou seria MU'AMMAR AL-QADHDHAFI? Como era de se esperar, a discussão sobre as diferentes grafias do nome do magistrado árabe chegou ao Brasil: MUAMAR KADDAFI, MUAMMER KHADAFI, MUAMMAR AL-GADDAFI, MUAMMAR AL-GATHAFI ou MOAMAR KADAFFI? Existem, ainda, as grafias adotadas pela imprensa brasileira: MUAMAR KADAFI (O Globo / revista Veja) e MUAMMAR GADDAFI (Folha de São Paulo). Esses são apenas alguns dos registros divulgados na mídia impressa e eletrônica. Dois fatores contribuíram significativamente para a ocorrência dessa polêmica árabe-brasileira. Primeiramente, a dificuldade de se estabelecer uma correspondência fonológica (de sons) exata entre alguns sons de nomes árabes

e fonemas de nosso alfabeto latino. Em segundo lugar, o fato de cada uma das formas gráficas tentar reproduzir, por escrito, diferentes pronúncias regionais do nome do ditador árabe. Ao que tudo indica, essa diversidade gráfica ainda vai render muito: MUAMMAR QUATHAFI, MUAMMAR GHA AWWAL, MUHAMMAD ABU MINYAR AL-QADHAFI também são palavras que constam da lista. Pelo que se propaga, as formas gráficas que nomeiam o governante líbio já chegam a mais de cem.

[...]

A RELAÇÃO FALA E ESCRITA

Para refletirmos sobre as variações do nome “Muamar Gaddafi”, bem como sobre as demais variantes gráficas anteriores, estamos recorrendo a um campo de estudo chamado Fonética. Segundo essa ciência, é possível compreender variações gráficas por meio de hipóteses feitas sobre a escrita das palavras. Estudos fonéticos seguem esse curso. Por outro lado, se considerarmos os estudos discursivos, a pergunta seria outra: As variações do nome do ditador árabe seriam dotadas de neutralidade? Certamente não, porque nomear é uma forma de interpretar. Para compreender a significação das palavras, temos aí dois caminhos. E há muitos outros. O leitor pode escolher qual deles trilhar.

Glória Dias Soares Vitorino. Doutora em Língua Portuguesa e Linguística. Mestre e especialista em Leitura e Produção de Textos. É pesquisadora da FAPEMIG e professora de Língua Portuguesa e Linguística UnilesteMG. [...]

do jornal *Diário do Aço*
<http://www.diariodoaco.com.br/colunas.aspx?cd=1705>
Postado em 2011-03-04, às 22h22.
Último acesso em 2011-07-24, 16h15.

ANEXO B

Kadafi ou Gaddafi: como se escreve o nome do ditador da Líbia?

Ditador do Líbano: 112 nomes para a mesma pessoa

Como não poderia deixar de ser, até a grafia correta do nome do mais famoso ditador do momento é motivo de controversa. [*sic*]

Momar Kadafi, Muammar Qaddafi, Muammar Gadhafi, Moamar Gaddafi, Moammar Kadhafi... Afinal, qual é o jeito certo de escrever o nome do coronel? Por causa da dificuldade em traduzir o árabe para o inglês, a Biblioteca do Congresso norte-americano fez uma lista de 72 grafias alternativas quando o presidente líbio visitou os Estados Unidos em setembro de 2009.

O jornal The New York Times e as agências de notícias Associated Press e Xinhua, só para complicar, adicionaram na época outras 40 maneiras (usadas entre 1998 e 2008) e a relação chegou a incríveis 112 grafias diferentes. Hoje, com Kadafi dominando o noticiário mundial dia após dia, só Deus (ou Alá) sabe quantas realmente são.

A Associated Press e as redes de TV CNN e MSNBC escolheram Moammar Gadhafi. The New York Times decidiu por Muammar el-

Qaddafi. Para o Los Angeles Times, é Moammar Kadafi. A agência de notícias Reuters, o diário The Guardian, a BBC e a Al-Jazeera vão com Muammar Gaddafi, enquanto o Irish Times fechou com Muammar G-dafy.

No jornal Christian Science Monitor, é Muammar Qaddafi. A rede de TV ABC News, que soletra Moammar Gaddafi, divulgou em seu site a lista com as 112 variações. [vide anexo F, p. 189]

[...]

do jornal *Diário do Comércio*
<http://www.jornalpequeno.com.br/blog/robertlobato/?p=15907>
Postado em 2011-02-25, por Robert Lobato.
Último acesso em 2011-07-24, 16h15.

Entenda as variações de fonética e grafia do nome do ditador da Líbia

Muammar Kadhafi tem 112 registros de grafias diferentes. ‘Transliteração mais correta seria Mu’amar Al Kadhafi’, diz especialista.

Muammar Kadhafi, Muamar Gaddafi, Mu’ ammar al-Qadhdhafi... A grafia ocidental do nome do ditador da Líbia encontra 112 variações diferentes em jornais, livros e documentos. A Library of Congress, órgão do governo dos Estados Unidos, por exemplo, lista 72 variações do nome do político. O G1 adota a grafia Muammar Kadhafi.

As variações foram notícia em reportagem do Jornal da Globo [...] e em outros veículos de comunicação. Para entender as razões de tantas diferenças de grafia e fonética, o G1 ouviu a especialista de estudos de cultura e literatura árabe Muna Omran, professora do departamento de Letras da Universidade Plínio Leite, do Rio, e editora da revista acadêmica Litteris, onde escreve sobre estudos árabes e islâmicos.

“Esta polêmica existe por conta da transliteração e dos fonemas árabes que não existem no nosso alfabeto”, explica Muna Omran.

Segundo ela, a transliteração mais correta seria Mu’amar Al Kadhafi,

assim, com apóstrofo no prenome e mais um “dh” e um “f” no sobrenome. “Na grafia em árabe do nome do ditador [...], a letra que se encontra depois do Al (que inicia o nome dele) seria na língua portuguesa o nosso ‘Q’, mas para podermos enfatizar a força deste fonema usa-se o ‘K’”, explica a professora.

Ainda de acordo com Muna, a letra árabe seguinte representa o “dh” do sobrenome de Kadhafi. “Não há necessidade de se pronunciar o ‘f’ [sic] dobrado”, diz. Sobre o primeiro nome, a professora explica que uma das letras representaria o apóstrofo em Mu’amar.

“É importante salientar que existe uma variação da oralidade da língua árabe de acordo com as diversas regiões”, explica Muna. “Em países do Norte da África se pronuncia o K como G, por isso há registros do nome Gaddafi. “Já a oralidade da Síria, Jordânia, Líbano e Arábia Saudita se aproxima do árabe clássico, com outras variações.”

Veja como alguns meios de comunicação grafam o nome do ditador:

BBC , Reino Unido - Gaddafi

Clarín , Argentina - Kadafi

El Pais , Espanha - Gadafi

Folha de S. Paulo , Brasil - Gaddafi

La Repubblica , Itália - Gheddafi

Le Monde , França - Kadhafi

New York Times , EUA - Qaddafi

Público , Portugal - Khadafi

O Globo , Brasil - Kadafi

do portal de notícias G1 (escrito por Paulo Guilherme)

<http://m.g1.globo.com/revolta-arabe/noticia/2011/02/entenda-variacoes-de-fonetica-e-grafia-do-nome-do-ditador-da-libia.html>

Postado em 2011-02-25, às 06h50; atualizado em 2011-02-25, às 12h34.

Último acesso em 2011-07-24, 16h15.

How Do You Spell ‘Gaddafi?’

The Linguistics Behind Libya’s Leader

Better hope you don’t get stuck with this one in your next spelling bee.

You may have noticed reading the news recently, that there is no agreed upon way to spell the name of the current Libyan leader, Muammar Gaddafi. The controversy is well illustrated by a May 1986 letter to Minnesota second graders; the leader signed his name in English “Moammar El-Gadhafi.” The AP reported the event, “Second-Graders Get Letter From Khadafy.” In 2009, ABC News listed 112 different ways to spell Gaddafi [*vide* anexo F, p. 189], which have appeared in various news outlets. The leader’s name was even the topic of a 1981 Saturday Night Live sketch [*vide* anexo E, p. 187], offering the most creative spelling a one-way ticket to Tripoli.

The explanation, you may have read, is due to discrepancies transliterating Arabic. While true, this is not the whole story. Gaddafi (Google’s most frequent spelling) is spelled القذافي in Arabic. The first letter “ق,” qoph, is pronounced as a “k” sound and usually transliterated as

a “q.” Likewise, the second letter of his name, *thal*, “ث”, is pronounced as a deep “d” or “th” (“the” not “with”) and transliterated “dh.” Given these standardized spellings, the Libyan leader’s name should be spelled “Qadhafi.”

The reason the most common spelling of Gaddafi begins with a “g” is due to the Libyan dialect, which pronounces *qoph* like a “g” sound. English translations of Arabic leaders names generally are consistent with their local dialects, rather than standard translation. The best example of this would be late Egyptian leader, Gamal Abdel Nasser. His first name beginning with a *jim* “ج” would be traditionally transliterated *Jamal*. Instead, the Egyptian pronunciation of his name is the norm for English transliteration.

So, while the Libyan dialect certainly isn’t as widely spoken as the Egyptian, there should at least be a standardization of the first letter of the leader’s last name. As for the “dh” versus “dd,” “i” versus “y,” including an “el” or an “al,” and spelling “Muammar”, we’ll have to save that for the next time NewsFeed writes about linguistics.

da revista *Time*

<http://newsfeed.time.com/2011/02/23/how-do-you-spell-gaddafi-the-linguistics-behind-libyas-leader/>

Postado em 2011-02-23.

Último acesso em 2011-07-24, 16h15.

FROM SATURDAY NIGHT LIVE 12/12/1981

QADAFI, KADAFI, AL-GADDAFI? MY SPELLCHECK JUST FROZE

Announcer: And now, “SNL Newsbreak”, with anchor Brian Doyle-Murray.

Brian Doyle-Murray: Good evening, I’m Brian Doyle-Murray.

Our top story tonight: This man, Libyan leader, Col. Moammar Kadaffi, has been the study of intense news coverage this week by every major news organization in America. However, every time his name appears in print, it has a different spelling. The Chicago Tribune spells it K-H-A-D-A-F-Y; The Los Angeles Times spells it K-A-D-A-F-I; Newsweek Magazine, K-A-D-D-A-F-I; Time Magazine, G-A-D-D-A-F-I; The Wall Street Journal, Q-A-D-H-A-F-I; The Washington Post, Q-A-D-D-A-F-I; The New York Times, el-Qaddafi. My personal favorite is from the comic book publishers - Kadaffy Duck¹.

¹Trocadilho com ⟨Daffy Duck⟩, nome original do personagem conhecido no Brasil como ⟨Patolino⟩ (Warner Bros.).

Brian Doyle-Murray: How do you spell Kadaffi? Let us know. [news screen scroll many weird spellings of Kadaffi] Our news research department has determined that no two people spell it alike. Send us your spelling of Kadaffi, and remember, it can't be the same as any of these spellings you're seeing on the screen right now. The most original spelling of the Libyan leader's name will be awarded a one-way ticket to Tripoli - that is, if your passport allows you to go there. so, send that in, let us know how you spell Kadaffi!

Figura 12: Programa *Saturday Night Live* de 1981-12-12.

do blog *The news that matters*

<http://thenewsthatmatters.blogspot.com/2011/02/qadaffi-kadaffi-al-gaddafi-my.html>

Postado em 2011-02-22, às 09h35.

Último acesso em 2011-07-24, 16h15

How Many Different Ways Can You Spell 'Gaddafi'?

Libyan leader Moammar Gaddafi has been in the headlines a lot this week – he's visiting the U.S. for the first time since he took power 40 years ago, and his arrival is not without controversy.

But politics aside, we've come across a list of the many different English spellings of Gaddafi's name. Because of the difficulty translating Arabic to English, there are several different translations – the Library of Congress lists 72 alternate spellings, and the New York times, Associated Press and Xinhua news sources used 40 additional spellings between 1998 and 2008. We've posted all 112 of them below...

1. Qaddafi, Muammar	7. El Kadhafi, Moammar
2. Al-Gathafi, Muammar	8. El Kazzafi, Moamer
3. al-Qadhafi, Muammar	9. El Qathafi, Mu'Ammar
4. Al Qathafi, Mu'ammar	10. Gadafi, Muammar
5. Al Qathafi, Muammar	11. Gaddafi, Moamar
6. El Gaddafi, Moamar	12. Gadhafi, Mo'ammar

-
13. Gathafi, Muammar	44. Muammar Al-Gathafi
14. Ghadafi, Muammar	45. Muammar al-Khaddafi
15. Ghaddafi, Muammar	46. Mu' ammar al-Qadafi
16. Ghaddafy, Muammar	47. Mu' ammar al-Qaddafi
17. Gheddafi, Muammar	48. Muammar al-Qadhafi
18. Gheddafi, Muhammar	49. Mu' ammar al-Qadhdhafi
19. Kadaffi, Momar	50. Mu' ammar al-Qadhdhāfi
20. Kad'afi, Mu'amar al-	51. Mu' ammar Al Qathafi
21. Kaddafi, Muamar	52. Muammar Al Qathafi
22. Kaddafi, Muammar	53. Muammar Gadafi
23. Kadhafi, Moammar	54. Muammar Gaddafi
24. Kadhafi, Mouammar	55. Muammar Ghadafi
25. Kazzafi, Moammar	56. Muammar Ghaddafi
26. Khadafy, Moammar	57. Muammar Ghaddafy
27. Khaddafi, Muammar	58. Muammar Gheddafi
28. Moamar al-Gaddafi	59. Muammar Kaddafi
29. Moamar el Gaddafi	60. Muammar Khaddafi
30. Moamar El Kadhafi	61. Mu' ammar Qadafi
31. Moamar Gaddafi	62. Muammar Qaddafi
32. Moamer El Kazzafi	63. Muammar Qadhafi
33. Mo' ammar el-Gadhafi	64. Mu' ammar Qadhdhafi
34. Moammar El Kadhafi	65. Muammar Quathafi
35. Mo' ammar Gadhafi	66. Mulazim Awwal
36. Moammar Kadhafi	Mu' ammar Muhammad
37. Moammar Khadafy	Abu Minyar al-Qadhafi
38. Moammar Qudhafi	67. Qadafi, Mu' ammar
39. Mu' amar al-Kad'afi	68. Qadhafi, Muammar
40. Mu' amar al-Kadafi	69. Qadhdhāfi, Mu' ammar
41. Muamar Al-Kaddafi	70. Qathafi, Mu' Ammar el
42. Muamar Kaddafi	71. Quathafi, Muammar
43. Muamer Gadafi	72. Qudhafi, Moammar
-

73. Moamar Al Kadafi [sic ¹]	94. Muammar Gaddhafi
74. Maummar Gaddafi	95. Muammar Gadhafi
75. Moamar Gadhafi	96. Muammar Ghadaffi
76. Moamer Gaddafi	97. Muammar Qadthafi
77. Moamer Kadhafi	98. Muammar al Gaddafi
78. Moamma Gaddafi	99. Muammar el Gaddafy
79. Moammar Gaddafi	100. Muammar el Gaddafi
80. Moammar Gadhafi	101. Muammar el Qaddafi
81. Moammar Ghadafi	102. Muammer Gadaffi
82. Moammar Khadaffy	103. Muammer Gaddafi
83. Moammar Khaddafi	104. Mummar Gaddafi
84. Moammar el Gadhafi	105. Omar Al Qathafi
85. Moammer Gaddafi	106. Omar Mouammer
86. Mouammer al Gaddafi	Al Gaddafi
87. Muamar Gaddafi	107. Omar Muammar
88. Muammar Al Ghaddafi	Al Ghaddafi
89. Muammar Al Qaddafi	108. Omar Muammar
90. Muammar Al Qaddafi	Al Qaddafi
91. Muammar El Qaddafi	109. Omar Muammar Al Qathafi
92. Muammar Gadaffi	110. Omar Muammar Gaddafi
93. Muammar Gadafy	111. Omar Muammar Ghaddafi
	112. Omar al Ghaddafi

da rede *ABC News*

<http://blogs.abcnews.com/theworldnewser/2009/09/how-many-different-ways-can-you-spell-gaddafi.html>

Postado em 2009-09-22, às 19h03.

Último acesso em 2011-07-24, 16h15.

¹A depender do tipo utilizado, não se nota estar grafado <AI>.

Nemo solus satis sapit.

«Sozinho,
ninguém sabe o bastante.»

ditado latino

Referências bibliográficas

ATTIE M Filho (2002). *Falsafa — a filosofia entre os árabes*. São Paulo: Palas Athena.

BAUER T (1996). Arabic writing. In: DANIELS PT; BRIGHT W (Eds.), *The world's writing systems*, New York: Oxford University. 559–568.

BRINGHURST R (2005). *Elementos do estilo tipográfico*. São Paulo: Cosac Naify.

CAGLIARI LC (2009). *A história do alfabeto*. São Paulo: Paulistana.

CHESTERMAN A; WAGNER E (2002). *Can theory help translators? — a dialogue between the ivory tower and the wordface*. Translation theories explained. Saint Jerome.

COMRIE B (1996). Languages of eastern and southern Europe. In: DANIELS PT; BRIGHT W (Eds.), *The world's writing systems*, New York: Oxford University. 663–689.

COULMAS F (1991). *The writing systems of the world*. Oxford: Basil Blackwell.

_____ (1996). *The Blackwell encyclopedia of writing systems*. Oxford: Basil Blackwell.

COWAN D (2007). *Gramática do árabe moderno: uma introdução*. São Paulo: Globo.

CRYSTAL D (1997). *The Cambridge encyclopedia of language*. Cambridge: Cambridge University.

CUNHA AGd (1987). *Dicionário etimológico Nova Fronteira da língua portuguesa*. Rio de Janeiro: Nova Fronteira.

DANIELS PT (1996a). The first civilizations. In: DANIELS PT; BRIGHT W (Eds.), *The world's writing systems*, New York: Oxford University. 21–32.

_____ (1996b). The invention of writing. In: DANIELS PT; BRIGHT W (Eds.), *The world's writing systems*, New York: Oxford University. 579–586.

_____ (1996c). The study of writing systems. In: DANIELS PT; BRIGHT W (Eds.), *The world's writing systems*, New York: Oxford University. 3–17.

DANIELS PT; BRIGHT W (1996). *The world's writing systems*. New York: Oxford.

DE LACY P (2007). *The Cambridge handbook of phonology*. New York: Cambridge University.

HIGOUNET C (2003). *História concisa da escrita*. São Paulo: Parábola.

JAROUCHE MM (2002). *Livro das mil e uma noites*. São Paulo: Globo.

JUBRAN SAC (1996). *Análise contrastiva do sistema fonológico do árabe e do português — nível segmental e suas implicações pedagógicas*. Dissertação de mestrado, Universidade de São Paulo.

_____ (2004). Para uma romanização padronizada de termos árabes em textos de língua portuguesa. *Tiraz*, 16–29.

- KAYE AS (1996). Adaptations of Arabic script. In: DANIELS PT; BRIGHT W (Eds.), *The world's writing systems*, New York: Oxford University. 743–762.
- KIRSCHBAUM S; BLATYTA DF; LANGER ER; ARAÚJO RGd (2009). *Transliteração do hebraico para leitores brasileiros*. São Paulo: Ateliê.
- LAMBDIN TO (2003). *Gramática do hebraico bíblico*. São Paulo: Paulus.
- LINGS M (2010). *Muhammad, a vida do Profeta do Islâm segundo fontes antigas*. São Paulo: Attar.
- MACHADO JP (1990). *Dicionário etimológico da língua portuguesa*. Lisboa: Livros Horizonte.
- NASR H (2005). *Dicionário árabe-português*. São Paulo: Câmara de Comércio Árabe Brasileira.
- NGUYỄN DH (1996). Vietnamese. In: DANIELS PT; BRIGHT W (Eds.), *The world's writing systems*, New York: Oxford University. 691–695.
- NIMER M (2005). *Influências orientais na língua portuguesa — os vocábulos árabes, arabizados, persas e turcos*. São Paulo: EDUSP.
- SABBAGH AN (1988). *Dicionário árabe-português-árabe*. Rio de Janeiro: UFRJ.
- SACKS D (2003). *Language visible — unraveling the mystery of the alphabet from A to Z*. New York: Broakway.
- SILVA TC (2003). *Fonética e fonologia do português — roteiro de estudos e guia de exercícios*. São Paulo: Contexto.
- WATSON JC (2002). *The phonology and morphology of Arabic*. New York: Oxford University.
- WRIGHT W (1859). *A grammar of the Arabic language*. Leipzig: Williams and Norgate.

Índice de citações

- ARAÚJO, Reginaldo Gomes de
(2009), 132, 141
- ATTIE, Miguel, Filho
(2002), 53
- BAUER, Thomas
(1996), 64, 65, 74, 75, 131
- BLATYTA, Dora Fraiman
(2009), 132, 141
- BRIGHT, William
(1996), 176
- BRINGHURST, Robert
(2009), 110, 126, 127, 129, 131, 132, 138, 143, 166
- CAGLIARI, Luiz C
(2009), 70
- CHESTERMAN, Andrew
(2002), 45, 52
- COMRIE, Bernard
(1996), 70, 134
- COULMAS, Florian
(1991), 64, 65, 67
- (1996), 134
- COWAN, David
(2007), 60, 61, 64, 70, 75–77, 118
- CRYSTAL, David
(1997), 60, 61, 64, 66
- CUNHA, Antônio G da
(1987), 102–104, 116
- DANIELS, Peter T
(1996), 176
(1996a), 66
(1996b), 134
(1996c), 64
- DE LACY, Paul
(2007), 127
- HIGOUNET, Charles
(2003), 65–67, 69, 70, 74
- JAROUCHE, Mamede M
(2005), 50
- JUBRAN, Safa AA-C
(1996), 160

- (2004), 103, 112, 114, 115, 128, WRIGHT, William
162 (1859), 118
- KAYE, Alan S
(1996), 68, 70, 134
- KIRSCHBAUM, Saul
(2009), 132, 141
- LAMBDIN, Thomas O
(2003), 131
- LANGER, Eliana Rosa
(2009), 132, 141
- LINGS, Martin
(2010), 118
- MACHADO, José P
(1990), 38, 103, 107
- NASR, Helmi
(2005), 102, 103, 105
- NGUYỄN, Đình-Hoà
(1996), 129
- NIMER, Miguel
(2005), iv, 103, 108, 133
- SABBAGH, Alphonse N
(1988), 70, 103, 109, 110
- SACKS, David
(2003), 68, 69
- SILVA, Thaïs C
(2003), 79
- WAGNER, Emma
(2002), 45, 52
- WATSON, Janet CE
(2002), 126, 128, 160