



A latinidade do léxico romeno pela lista de Swadesh

The latinity of the Romanian lexicon by means of Swadesh's list

Rogério Augusto Monteiro Cardoso

Universidade Federal do Amazonas (UFAM), Manaus, Amazonas / Brasil

rogerackbar@gmail.com

<https://orcid.org/0000-0003-4953-6175>

Resumo: Este artigo visa a apresentar a latinidade do romeno por meio da *Lista de Swadesh*, que reúne de 100 a 200 itens lexicais do vocabulário básico de uma língua qualquer, com o intuito de identificar possíveis parentescos linguísticos. Embora a filiação do romeno à família românica seja hoje indiscutível, as suas feições peculiares o fazem parecer, à primeira vista, um idioma muito apartado da realidade neolatina. No entanto, por meio da análise etimológica dos cerca de 100 itens lexicais aqui listados, chega-se à conclusão de que o léxico romeno primitivo não só é fortemente latino, como também possui muitas palavras em comum com as línguas românicas ocidentais. Com efeito, 90% dos itens arrolados têm origem latina, demonstrando que o vocabulário básico de uma língua é, de fato, mais resistente a empréstimos. Os cognatos românicos nem sempre são tão óbvios porque certos fenômenos fonéticos típicos da história do romeno podem ter alterado drasticamente a forma dos vocábulos, tornando-os às vezes irreconhecíveis.

Palavras-chave: léxico; etimologia; Swadesh; latinidade; romeno.

Abstract: This article aims to present Latinity of the Romanian language by means of the *Swadesh's List*, which gathers from 100 to 200 lexical items of the basic vocabulary of a certain language, in order to identify possible linguistic kinship. Although the Romanian's affiliation to the Romance family is nowadays unquestionable, its peculiar features make it seem, at first glance, a language far apart from neolatin reality. However, through the etymological analysis of the approximately 100 lexical items here listed, it is concluded that the early Romanian lexicon is not only strongly Latin-based but also

has many words in common with Western Romance languages. As a result, 90% of the listed items has Latin origin, demonstrating that the basic vocabulary of a language is, in fact, more resistant to borrowing. Romance cognates are not always that obvious because certain phonetic phenomena typical of Romanian history may have drastically altered the shape of the words, making them sometimes unrecognizable.

Keywords: lexicon; etymology; Swadesh; latinity; romanian.

Introdução

No prefácio da obra *História breve da língua romena*, de Alexandru Niculescu (1983), o filólogo brasileiro Sílvio Elia, em tom de lamentação, afirmou à época que o romeno era uma espécie de primo pobre da Filologia Românica, porque, segundo ele, “contam-se pelos dedos os brasileiros que dominam esse idioma” (Niculescu, 1983, p. 9). Décadas depois, a situação parece não ser muito distinta: o romeno continua sendo um idioma pouco acessível, haja vista o ínfimo número de gramáticas, livros didáticos, dicionários e cursos voltados para o público lusófono. Quem quiser adentrar nos meandros desse parente românico terá de recorrer, ou a obras mais antigas, há muito ausentes do mercado editorial, ou a obras editadas noutras línguas, como inglês, francês ou castelhano. O insulamento geográfico do romeno no Leste Europeu, que por muito tempo o privou de um contato mais direto com os membros ocidentais da família românica, e o desenvolvimento tardio de sua tradição escrita, a partir do século XVI, ajudam a explicar o conhecimento limitado que se tem desse idioma no Brasil e alhures (Bassetto, 2013, p. 188-193). Dentre os principais entusiastas do romeno no campo da Romanística, está o filólogo brasileiro Theodoro Henriques Maurer Jr., cuja obra *A unidade da România ocidental* (1951) enfatiza o papel corroborativo desse parente linguístico na detecção de fenômenos, construções ou palavras panromânicas. Segundo ele, o antigo dialeto daco-romano, falado em terras dácias (atual Romênia), apartou-se dos dialetos ocidentais há cerca de 1500 anos, de sorte que boa parte do seu patrimônio latino supérstite tenha sido preservada desde os tempos antigos, sem a influência do latim medieval ou do eclesiástico (Maurer Jr., 1951, p. 9-15). Advirta-se, porém, que muitos eruditos da Valáquia e da Transilvânia promoveram uma “relatinização” lexical do romeno nos

séculos XVIII e XIX, por influência do francês, do italiano e do próprio latim científico, num afã de reaproximá-lo dos seus parentes ocidentais (Niculescu, 1983, p. 98-106) – um fato que uma pesquisa nesta área **não** poderia negligenciar. Enfim, este artigo visa a mostrar ao público leitor que a língua romena, não obstante as suas peculiaridades morfossintáticas e lexicais, conservou vestígios indeléveis das suas origens latinas, ainda que estas nem sempre sejam tão óbvias ou reconhecíveis a um falante lusófono. Uma boa maneira de demonstrá-lo é pela *Lista de Swadesh*, na qual se arrolam de 100 a 200 itens lexicais pertencentes ao vocabulário básico de uma língua.

Antes de aplicar a lista ao romeno, convém trazer alguns esclarecimentos teóricos e metodológicos. Na seção 1.1, discute-se, por meio de excertos, o papel do léxico na identificação e no reconhecimento de parentescos linguísticos. Na seção 1.2, expõem-se os fundamentos da *Lista de Swadesh*, consoante os apontamentos de Fox (1995). Na seção 1.3, enumeram-se, de modo mais pormenorizado, as etapas de aplicação da lista e alguns cuidados a serem tomados durante o processo. Enfim, na seção 2, aplica-se a *Lista de Swadesh*. Em cada subseção, há uma tabela na qual constam: 1) a numeração e a exibição dos itens lexicais em inglês e português; 2) a tradução destes para romeno; 3) a etimologia dos itens lexicais romenos, com base nas informações fornecidas pelo dicionário virtual e monolíngue de romeno, o *dexonline*¹. Logo abaixo das tabelas, há alguns comentários pontuais para esclarecer ou destacar peculiaridades etimológicas. Por último, na seção 3, apresentam-se, em números absolutos e percentuais, as línguas que mais contribuíram para a formação do vocabulário básico romeno, além de algumas conclusões interessantes que se podem tirar dos dados obtidos e de certas palavras constantes na lista. A aplicação da *Lista de Swadesh*, em sua versão de 100 itens lexicais, demonstra que aproximadamente 90% do vocabulário básico do romeno têm origem latina, contrastando com apenas 20% de vocabulário latino primitivo em relação ao léxico total da língua, segundo os dados apontados por Cojocar (2003, p. 11). Esses 90% corroboram a **latinidade** do léxico romeno, conforme expressa o título do artigo, na medida em que 9 a cada 10 palavras presentes na lista foram herdadas do latim há mais de mil anos, perpassando inúmeras gerações de falantes.

¹ Disponível em: <<https://dexonline.ro/>>

1 Fundamentos e métodos

1.1 O léxico e os parentescos linguísticos

O léxico, ao lado da morfologia e da sintaxe, tem papel fortemente corroborativo na detecção de parentescos linguísticos (Fox, 1995). Com efeito, um falante monoglota de português poderia identificar no castelhano, no francês e no italiano um grande número de itens lexicais cognatos, de modo que ele tenha, ao menos, uma compreensão parcial de textos escritos nessas línguas neolatinas. Provam-no, por exemplo, estas versões do preâmbulo da famosa Declaração Universal dos Direitos Humanos, que já foi traduzida para centenas de idiomas:

Castelhano: “*Considerando que la libertad, la justicia y la paz en el mundo tienen por base el reconocimiento de la dignidad intrínseca y de los derechos iguales e inalienables de todos los miembros de la familia humana...*”²;

Francês: “*Considérant que la reconnaissance de la dignité inhérente à tous les membres de la famille humaine et de leurs droits égaux et inaliénables constitue le fondement de la liberté, de la justice et de la paix dans le monde...*”³;

Italiano: “*Considerato che il riconoscimento della dignità inerente a tutti i membri della famiglia umana e dei loro diritti, uguali ed inalienabili, costituisce il fondamento della libertà, della giustizia e della pace nel mondo...*”⁴.

Um olhar rápido sobre os trechos já permite constatar a existência de itens cognatos mutuamente inteligíveis com as formas portuguesas.

² Disponível em: <<https://www.un.org/es/universal-declaration-human-rights/>>. Acesso em: abr. 2023.

³ Disponível em: <<https://www.un.org/fr/universal-declaration-human-rights/>>. Acesso em: abr. 2023.

⁴ Disponível em: <<https://www.ohchr.org/en/human-rights/universal-declaration/translations/italian>>. Acesso em: abr. 2023.

Alguns podem ser identificados de imediato: *familia, famille, famiglia* (i.e. “família”); *inalienables, inaliénables, inalienabili* (i.e. “inalienáveis”). Outros são um pouco menos óbvios, mas podem ser deduzidos por meio de uma análise comparativa dos três trechos: *derechos, droits, diritti* (i.e. “direitos”); *iguales, égaux, uguali* (i.e. “iguais”).

Veja-se agora esta versão do preâmbulo da Declaração Universal dos Direitos Humanos:

Romeno: “*Considerînd că recunoaşterea demnităţii inerente tuturor membrilor familiei umane şi a drepturilor lor egale şi inalienabile constituie fundamentul libertăţii, dreptăţii şi păcii în lume...*”⁵.

À primeira vista, um falante monoglota de português estranharia a combinação de certas letras com alguns diacríticos, como a braquia em *ă*, o circunflexo em *î*, além de uma cedilha subscrita em *ş* ou em *ş̃*⁶. Também lhe dificultaria a compreensão o fato de o romeno utilizar o caso genitivo-dativo em contextos sintáticos nos quais as outras línguas neolatinas utilizariam a preposição *de* seguida de um sintagma nominal: *reconocimiento de la dignidad, reconnaissance de la dignité, riconoscimento della dignità, recunoaşterea demnităţii*⁷. Outra diferença morfosintática marcante é a posposição dos artigos definidos aos respectivos nomes: por exemplo, em *membrilor*, a partícula enclítica *-lor* é o artigo masculino no genitivo-dativo plural, logo *membrilor* significa “dos membros”. Não obstante essas diferenças, é possível identificar os mesmos itens cognatos já citados nos parágrafos anteriores: *familiei, inalienabile, drepturilor, egale*. Desse modo, as semelhanças lexicais entre o português e o romeno já dariam de antemão boas evidências de

⁵ Disponível em: <<https://childrenandarmedconflict.un.org/keydocuments/romanian/universaldeclarat.html>>. Acesso em: abr. 2023.

⁶ O *ă* representa uma vogal central média [ə]; o *î* representa uma vogal central alta [i]; o *ş* representa uma fricativa pós-alveolar surda [ʃ], tal como o dígrafo *ch* do português; por fim, o *ş̃* representa a consoante africada surda [tʃ].

⁷ Note-se que a palavra romena para dignidade (*demnitate*) possui vogal tônica aberta [a] no caso nominativo-acusativo, mas, quando se troca o *-e* final pela desinência *-i* do genitivo-dativo, ocorre um alçamento vocálico de [a] para [ə] e uma fricativização do [t] em [tʃ]: *demnităţi*. O segundo *-i* de *demnităţii* é o artigo enclítico. Logo: *demnităţii* significa “da dignidade”.

um parentesco linguístico entre os dois idiomas. É justamente disso que trata a *Lista de Swadesh*.

1.2 Os fundamentos da *Lista de Swadesh*

Morris Swadesh (1909-1967) foi um linguista estadunidense que se notabilizou pelas duas áreas gêmeas que desenvolveu desde a década de 1950 até a sua morte: a Lexicostatística e a Glotocronologia (Grant, 2010, p. 1). Ambas as áreas possuem abordagens quantitativas, mas têm objetivos levemente distintos: a primeira visa a estabelecer, por meio do chamado **vocabulário básico**, o grau de similaridade entre as línguas, enquanto a segunda busca determinar, com base no grau de diferenciação vocabular, quando duas línguas da mesma família se separaram (Fox, 1995, p. 281). Por exemplo, o romeno e o castelhano têm a mesma origem latina, porém é empiricamente constatável que este último é mais similar ao português devido ao ininterrupto contato linguístico entre ambos na Península Ibérica e nas Américas. Calcula-se que o castelhano e o português compartilhem entre si quase 90% do seu léxico (García; Souza, 2014, p. 150), porém, entre o romeno e o português, tal percentagem é decerto menor devido à penetração de muitos vocábulos eslavos, húngaros, neogregos e turcos nos dialetos daco-românicos, distanciando-os dos dialetos românicos ocidentais (Bassetto, 2010, p. 161-167).

A *Lista de Swadesh* foi desenvolvida num artigo intitulado *Towards Greater Accuracy in Lexicostatistic Dating* (“Para uma maior precisão em datação lexicostatística”), vindo a lume em 1955. Para entender melhor o emprego da referida lista nos estudos linguísticos, é necessário demonstrar o que vem a ser o dito vocabulário básico (*core vocabulary*). Por exemplo, ao se compararem línguas como o inglês e o alemão, que pertencem à família germânica, notam-se significativas diferenças vocabulares devido ao grande número de vocábulos latinos existentes na língua inglesa: *recognition-Anerkennung* (“reconhecimento”); *inherent-angeborenen* (“inerente”); *dignity-Würde* (“dignidade”) *etc.* Porém, ao se compararem vocábulos de outros campos semânticos – partes do corpo, meio ambiente, palavras de classe fechada e outros –, as raízes germânicas das duas línguas ficam em evidência. Veja-se esta pequena amostra:

Tabela 1 – Itens básicos em inglês e alemão

| N.º | Português | Inglês | Alemão |
|-----|-----------|--------------|---------------|
| 1. | cabeça | <i>head</i> | <i>Kopf</i> |
| 2. | mão | <i>hand</i> | <i>Hand</i> |
| 3. | joelho | <i>knee</i> | <i>Knie</i> |
| 4. | pé | <i>foot</i> | <i>Fuß</i> |
| 5. | água | <i>water</i> | <i>Wasser</i> |
| 6. | terra | <i>land</i> | <i>Land</i> |
| 7. | sol | <i>sun</i> | <i>Sonne</i> |
| 8. | lua | <i>moon</i> | <i>Mond</i> |
| 9. | eu | <i>I</i> | <i>ich</i> |
| 10. | tu | <i>you</i> | <i>du</i> |

Fonte: Elaboração própria

Nota-se que apenas dois pares não são cognatos: *you-du* (“tu”) e *head-Kopf* (“cabeça”). No primeiro caso, o correspondente etimológico de *du* no inglês é *thou*, cujo uso é arcaico nos países anglófonos (Algeo, 2010, p. 165-166). No segundo caso, o correspondente etimológico de *head* no alemão é *Haupt*, cujo uso, no sentido de “cabeça”, é também arcaico (Kluge, 1899, p. 165). Em sendo assim, por que certas áreas do léxico inglês parecem ora se distanciar, ora se aproximar das suas origens germânicas? Porque os itens lexicais do vocabulário básico são mais resistentes a empréstimos e substituições. Por vocabulário básico entenda-se o conjunto de itens lexicais a princípio existentes em qualquer língua, independentemente da época, da cultura, da localização geográfica ou do desenvolvimento tecnológico dos seus falantes (Fox, 1995, p. 282). Portanto, todas as línguas humanas, vivas ou mortas, possuem algum termo para designar partes do corpo (*cabeça, mão pé*), elementos do meio ambiente (água, terra, sol), além de categorias gramaticais (*eu, tu, nós*). A *Lista de Swadesh* nada mais é do que uma tentativa de arrolar esses termos culturalmente **não marcados**, a fim de determinar graus de similaridade linguística.

Uma das previsíveis dificuldades de aplicação do método proposto por Swadesh consiste em estabelecer qualitativa e quantitativamente os itens lexicais que devem entrar na lista. Segundo Fox (1995):

A lista original de Swadesh continha 200 itens básicos, mas ela foi depois (Swadesh, 1955) reduzida a 100, com o intuito de eliminar vários itens suspeitos de serem culturalmente dependentes, ou, de outro modo, não confiáveis. Esses incluem nomes de animais ('peixe'), itens dependentes do clima ('neve'), alguns itens que pudessem ser expressos por sinônimos ('mulher'/'esposa'), assim como palavras para algumas atividades que foram consideradas insatisfatórias ('cortar', 'empurrar', 'cavar', 'apertar'), e outras. (Fox, 1995, p. 282, tradução nossa)⁸.

A versão utilizada nesta pesquisa consta de 100 itens lexicais⁹, que cobrem as seguintes áreas do léxico: palavras de classe fechada; adjetivos de tamanho; seres vivos; plantas e suas partes; partes do corpo dos seres vivos; verbos de ação, percepção e declaração; elementos do meio ambiente, cores, entre outras coisas.

1.3 As etapas de aplicação da *Lista de Swadesh*

A aplicação da *Lista de Swadesh* ao romeno ou a qualquer língua de tradição escrita é relativamente simples, mas requer alguns cuidados por parte do pesquisador.

O primeiro passo é traduzir os 100 itens com base em critérios semânticos, e não com base em possíveis elos etimológicos. Na pequena amostra exibida na seção anterior, traduziu-se o substantivo inglês *head* ("cabeça") por *Kopf*, ainda que em alemão haja o substantivo cognato *Haupt*. Noutros casos, é necessário dar mais de uma tradução ao termo arrolado na lista, seja pela existência de sinônimos muito próximos na língua pesquisada, seja por eventuais dissimetrias entre a língua pesquisada e o inglês. Por exemplo, um item como *short* se pode traduzir em português por *baixo* ou por *curto* a depender do substantivo qualificado pelo adjetivo: *short man*, *short answer* (i.e. homem baixo,

⁸ No original: "Swadesh's original list contained 200 core items, but this was later (Swadesh, 1955) pruned to 100, in order to eliminate a number of items that were suspected of being culture-dependent, or otherwise unreliable. These included animal names ('fish'), items dependent on climate ('snow'), some items that could be expressed by synonyms ('woman'/'wife'), as well as words for some activities which were found to be unsatisfactory ('cut', 'pull', 'dig', 'squeeze'), and some others."

⁹ Disponível em inglês e francês na página: <<http://comparalex.org/index.php?page=stdlist&id=19>> Acesso em: jan. 2023).

resposta curta). É forçoso lembrar que a opção por traduzir um item da lista de duas ou mais maneiras depende, em alguma medida, de um juízo do pesquisador, de modo que a aplicação da *Lista de Swadesh* ao romeno possa gerar resultados levemente variáveis. Todavia, com vistas a tornar a tradução dos termos originais em inglês a mais objetiva possível, esta pesquisa recorreu aos seguintes critérios: 1) só se utilizam duas ou mais palavras na tradução, quando estas forem sinônimas muito próximas e de uso corrente em romeno, a exemplo de *arbore* e *copac*, que significam “árvore”; 2) se duas palavras quaisquer forem sinônimas próximas, sendo uma de uso corrente e a outra de uso apenas regional, popular, figurado ou científico, só se utilizará, para efeito de contagem, a de uso corrente. Por exemplo, os adjetivos *negru* e *lai* significam “negro” ou “preto”, mas, como o adjetivo *lai* só se emprega em variantes populares (Buescu, 2003, p. 248), não foi incluído na tradução. A ideia, pois, é tentar buscar formas **não marcadas** por nenhum tipo de registro mais específico.

O segundo passo é indicar a etimologia dos itens da língua-alvo – no caso em tela, da língua romena. A principal fonte de informações etimológicas desta pesquisa é um dicionário virtual e monolíngue de romeno, o *dexonline*¹⁰, cujo cerne é o *Dicționarul explicativ al limbii române* (Coteanu, 2009), da Academia Romena. Ao longo do tempo, a equipe de voluntários do *dexonline* acresceu informações oriundas de dezenas de dicionários – incluindo dicionários técnicos ou temáticos –, de sorte que o consulente tenha à sua disposição um copioso número de acepções, de expressões idiomáticas, de informações etimológicas, além dos paradigmas flexionais. Em havendo divergências acerca da origem de uma palavra, o material traz os étimos propostos, indicando-lhes os proponentes e fazendo comentários críticos sobre as mudanças fonéticas, morfológicas e/ou semânticas pelas quais o termo passou ou teria passado.

O terceiro passo é fazer breves comentários etimológicos acerca de alguns itens lexicais romenos. Por ser demasiado fastidioso e até desnecessário trazer adendos sobre todos os itens da lista, dá-se destaque apenas àqueles que apresentarem alguma peculiaridade digna de nota na fonética, na semântica ou no seu étimo. Alusões comparativas aos cognatos românicos são frequentes e servem para ilustrar a latinidade do léxico romeno. Por fim, também para mitigar o fastio da leitura e para

¹⁰ Disponível em: <<https://dexonline.ro/>>

facilitar o trabalho de consulta, cada tabela da próxima seção destina-se a apenas um ou, no máximo, dois campos semânticos.

O quarto e último passo é apresentar, em números absolutos e percentuais, as línguas que mais contribuíram para a formação do vocabulário básico romeno, com o devido respaldo nas informações etimológicas fornecidas pelos dicionários do *dexonline* e, de modo esporádico, por outros autores. Cumpre aqui ressaltar que, no caso dos itens de base eslava mais remota, as fontes utilizadas pelo dicionário virtual titubeiam em atribuir-lhes origem no **antigo eslavo comum**, datado das primeiras incursões eslavas no território da Romênia moderna no fim do século VI d.C. (Niculescu, 1983, p. 43-54), ou no **eslavônico eclesiástico**, utilizado pela Igreja Ortodoxa a partir do século IX no Leste Europeu (Huntley, 1993, p. 125-126) – pelo que convém usar o genérico rótulo de *eslavo*. Por fim, nos casos em que um item lexical de determinada origem remota chegou ao romeno passando antes por uma língua intermediária, considera-se, para efeito de contagem, a língua intermediária. Assim, o substantivo *piatră* (“pedra”), que tem origem remota no grego πέτρα, mas que passou pelo latim *petram* antes de chegar ao romeno, é contado como latino.

2 Lista de Swadesh aplicada ao romeno¹¹

Tabela 2 – Palavras de classe fechada

| N.º | Inglês | Português | Romeno | Etimologia |
|-----|-------------------|-----------|----------------------------|--------------------------------|
| 1. | <i>I</i> | eu | <i>eu</i> | < lat. <i>ego</i> |
| 2. | <i>thou (you)</i> | tu | <i>tu</i> | < lat. <i>tu</i> |
| 3. | <i>we</i> | nós | <i>noi</i> | < lat. <i>nos</i> |
| 4. | <i>this</i> | este | <i>acest</i> ¹² | < lat. <i>*eccum istum</i> |
| 5. | <i>that</i> | aquele | <i>acel</i> | < lat. <i>*eccum illum</i> |
| 6. | <i>who</i> | quem | <i>cine</i> | < lat. <i>*quene < quem</i> |
| 7. | <i>what</i> | o que | <i>ce</i> | < lat. <i>quid</i> |
| 8. | <i>not</i> | não | <i>nu</i> | < lat. <i>non</i> |
| 9. | <i>all</i> | todos | <i>toți</i> | < lat. <i>toti</i> |
| 10. | <i>many</i> | muitos | <i>mulți</i> | < lat. <i>multi</i> |
| 11. | <i>one</i> | um | <i>unu</i> | < lat. <i>unum</i> |
| 12. | <i>two</i> | dois | <i>doi</i> | < lat. <i>*dui ≈ duo</i> |

Fonte: Elaboração própria

¹¹ Índice de símbolos (cf. Viaro, 2011, p.13-14):

*X X é reconstruído

X > Y X se transforma em Y

X < Y X provém de Y

X >> Y o significado X se transforma no significado Y

X << Y o significado X provém do significado Y

X → Y X deriva (morfologicamente) Y

X ← Y X deriva (morfologicamente) de Y

X ≈ Y X é variante de Y.

¹² À semelhança do italiano, o grafema c, antes de e ou de i, designa uma consoante africana [tʃ] em romeno. Logo, *acest* se pronuncia [a'tʃest], *acel* se pronuncia [a'tʃel], *cine* se pronuncia [tʃine], etc.

Observações:

- Nas palavras de classe fechada, há notável identidade etimológica entre o português e o romeno, que surgiram, vale lembrar, em extremidades opostas do território românico. Isso não só reforça a latinidade do romeno, como também confirma uma das premissas da *Lista de Swadesh*: as palavras de classe fechada são menos suscetíveis a empréstimos ou substituições.
- Nos pronomes demonstrativos e nos interrogativos, ocorreu um consonantismo típico da fonética diacrônica do romeno: a fricativização da consoante velar [k^w] em [tʃ] diante de vogais anteriores: *quem* > **quene* > *cine*; *quid* > *ce* (Dobrinescu, 1992, p. 228).

Tabela 3 – Adjetivos de tamanho

| N.º | Inglês | Português | Romeno | Etimologia |
|-----|--------------|-----------|-------------|--------------------------------------|
| 13. | <i>big</i> | grande | <i>mare</i> | < lat. <i>marem</i> |
| 14. | <i>long</i> | longo | <i>lung</i> | < lat. <i>lōngum</i> |
| 15. | <i>small</i> | pequeno | <i>mic</i> | < lat. * <i>miccus</i> < <i>mīca</i> |

Fonte: Elaboração própria

Observações:

- Em (13), o adjetivo *mare* (“grande”) vem provavelmente do latim *marem* (“macho”). O semantismo¹³ aqui se explica pelo fato de os indivíduos machos serem geralmente maiores e mais robustos que as respectivas fêmeas (Ciorănescu, 1966). Nesse caso, o romeno divergiu das línguas românicas ocidentais: *grande* (português, castelhano e italiano) e *grand* (francês);
- Em (15), o adjetivo *mic* (“pequeno”) teria provindo da forma hipotética **miccus*, e esta, do latim *mīca* (“parcela”), cognata do grego dórico μικός (Șăineanu, 1929);

¹³ Entenda-se por semantismo a mudança ou as mudanças semânticas pelas quais um termo passou na história de uma língua.

Tabela 4 – Seres vivos

| N.º | Inglês | Português | Romeno | Etimologia |
|-----|------------------|-----------|-----------------|--|
| 16. | <i>woman</i> | mulher | <i>femeie</i> | < lat. <i>familiam</i> |
| | | | <i>muiere</i> | < lat. <i>muliërem</i> |
| 17. | <i>man (vir)</i> | homem | <i>bãrbat</i> | < lat. <i>barbãtum</i> |
| 18. | <i>person</i> | peessoa | <i>persoanã</i> | < fr. <i>personne</i> |
| 19. | <i>fish</i> | peixe | <i>pește</i> | < lat. <i>piscem</i> |
| 20. | <i>bird</i> | pãssaro | <i>pasãre</i> | < lat. pop. <i>passãrem</i> < <i>passërem</i> |
| 21. | <i>dog</i> | cão | <i>cãine</i> | < lat. <i>canem</i> |
| 22. | <i>louse</i> | pioelho | <i>pãduche</i> | < lat. <i>peducũlum</i> |

Fonte: Elaboração própria

Observações:

- Em (16), o substantivo *femeie* (“mulher”) provém do latim *familiam* (“família”), segundo os dicionários compilados no *dexonline*, como o *Dicționarul explicativ al limbii române* (COTEANU, 2009). A princípio, poder-se-ia pensar que o étimo de *femeie* seria o latim *femīnam*, do qual provém, por exemplo, o francês *femme* e o português *fêmea*. Porém, seguindo as tendências do consonantismo romeno, a consoante lateral [l] sofreu iodização, de sorte que *familiam* tenha, de fato, originado *femeie* (Dobrinescu, 1992, p. 228). Aliás, o substantivo romeno *famīlie* (“família”) chegou à língua pelo francês *famille*. Por fim, em *muiere* (< latim *muliërem*), ocorreu a mesma iodização na consoante lateral [l], além da diástole do acento tônico. Esse segundo substantivo se refere geralmente a mulheres casadas (Buescu, p. 283, 2003);
- Em (17), o substantivo *bãrbat* (“homem”, “indivíduo do sexo masculino”) provém do latim *barbãtum* (“barbado”). O semantismo aqui é óbvio uma vez que a barba é comum aos homens. O romeno possui também o substantivo *om* (“ser humano”), que provém da forma nominativa latina *homo*, enquanto a forma cognata portuguesa *homem* provém do

acusativo: *hominem* > *homêe* > *homem* (Huber, p. 85, §142, 1986).

- Em (18), o substantivo *persoană* (“pessoa”) tem origem remota no latim *persōna*, mas, segundo o *Dicționarul explicativ al limbii române* (Coteanu, 2009), teria havido uma influência conjunta do francês *personne* e do alemão *Person*, logo não se trataria de um vocábulo herdado ao latim desde os tempos antigos. O substantivo *om* acima referido também pode significar “pessoa”.
- Em (19), o substantivo *pește* (“peixe”) nada tem a ver com o substantivo latino *pestem* (“peste”). O consonantismo romeno evidencia que o grupo consonantal latino *-sc*, seguido de *-e* ou *-i*, torna-se *-șt*, logo: [sk] > [ʃt]. Isso explica a mudança *pisces* > *pește* (Dobrinescu, 1992, p. 228);

Tabela 5 – Plantas e suas partes

| N.º | Inglês | Português | Romeno | Etimologia |
|-----|-------------|-----------|-----------------|--|
| 23. | <i>tree</i> | árvore | <i>arbore</i> | < lat. culto <i>arbōr</i> , <i>-is</i> |
| | | | <i>copac</i> | (incerta) |
| 24. | <i>seed</i> | semente | <i>sămânță</i> | < lat. pop. * <i>se-</i> <i>mentīam</i> |
| 25. | <i>leaf</i> | folha | <i>foaie</i> | < lat. <i>folīam</i> |
| | | | <i>frunză</i> | < lat. <i>frondia</i> ← <i>frons</i> , <i>frondis</i> |
| 26. | <i>root</i> | raiz | <i>rădăcină</i> | < lat. <i>radīcīna</i> ← <i>radīcem</i> |
| 27. | <i>bark</i> | casca | <i>scoarță</i> | < lat. <i>scortea</i> ← <i>scortum</i> |

Fonte: Elaboração própria

Observações:

- Em (23), há dois itens lexicais para “árvore”: *arbore* (< latim culto *arbōr*, *-is*) e *copac*. A semelhança física entre *arbore* e o seu étimo latino sugere tratar-se de um provável latinismo, que inclusive sobrepujou uma variante mais antiga dessa palavra:

arbure (Ciorănescu, 1966). O item *copac* tem origem incerta e talvez tenha vindo do albanês *kopač* (Coteanu, 2009). Por fim, o romeno tem ainda a palavra *pom*, que não entrou na lista por designar especificamente árvores frutíferas (Buescu, 2003, p. 343);

- Em (24), vê-se que a palavra romena *sămânță* (“semente”) provém de uma forma popular hipotética **sementiam*, de que também se originaram a correspondente italiana *semenza* e a francesa *semence* (Ciorănescu, 1966). A palavra portuguesa *semente*, por seu turno, provém da forma culta *sementem*, de *sementis*, *-is*.
- Em (25), há dois itens lexicais para “folha”: *foaie* (< latim *foliam*) e *frunză* (< lat. *frondia*). De acordo com alguns dicionários citados pelo *dexonline*, a forma *frondia* teria vindo do substantivo feminino *frons*, *frondis*, conforme o indicado, porém, segundo Scriban (1939), *frondia* seria uma forma neutra plural que teria surgido paralelamente à forma regular *frondes*, no feminino plural.

Tabela 6 – Partes do corpo dos seres vivos

| N.º | Inglês | Português | Romeno | Etimologia |
|-----|---------------------|-------------|----------------|---|
| 28. | <i>skin</i> | pele | <i>piele</i> | < lat. <i>pēllem</i> |
| 29. | <i>flesh (meat)</i> | carne | <i>carne</i> | < lat. <i>carnem</i> |
| 30. | <i>blood</i> | sangue | <i>sânge</i> | < lat. <i>*sanguem</i> < <i>sanguinem</i> |
| 31. | <i>bone</i> | osso | <i>os</i> | < lat. <i>ossum</i> |
| 32. | <i>grease, oil</i> | gordura | <i>grăsime</i> | (<i>gras</i> + <i>ime</i>) <i>gras</i> < lat. <i>grassum</i> |
| 33. | <i>egg</i> | ovo | <i>ou</i> | < lat. <i>ovum</i> |
| 34. | <i>horn</i> | chifre | <i>corn</i> | < lat. <i>cornu</i> |
| 35. | <i>tail</i> | cauda, rabo | <i>coadă</i> | < lat. <i>*coda</i> < <i>caudam</i> |
| 36. | <i>feather</i> | pena | <i>pană</i> | < lat. <i>pinnam</i> |
| | | | <i>fulg</i> | (desconhecida) |
| 37. | <i>hair</i> | cabelo | <i>păr</i> | < lat. <i>pīlum</i> |

| | | | | |
|-----|------------------|---------|-----------------|---|
| 38. | <i>head</i> | cabeça | <i>cap</i> | < lat. * <i>capum</i> < <i>caput</i> |
| 39. | <i>ear</i> | orelha | <i>ureche</i> | < lat. pop. <i>oricla</i> < <i>auricūlam</i> ← <i>auris</i> |
| 40. | <i>eye</i> | olho | <i>ochi</i> | < lat. <i>oc(u)lum</i> |
| 41. | <i>nose</i> | nariz | <i>nas</i> | < lat. <i>nasum</i> |
| 42. | <i>mouth</i> | boca | <i>gură</i> | < lat. <i>gulam</i> |
| 43. | <i>tooth</i> | dente | <i>dinte</i> | < lat. <i>dentem</i> |
| 44. | <i>tongue</i> | língua | <i>limbă</i> | < lat. <i>linguam</i> |
| 45. | <i>claw</i> | garra | <i>unghie</i> | < lat. <i>ung(u)lam</i> |
| 46. | <i>foot</i> | pé | <i>picior</i> | < lat. <i>petiōlum</i> |
| 47. | <i>knee</i> | joelho | <i>genunchi</i> | < lat. <i>genucūlum</i> ← <i>genu</i> |
| 48. | <i>hand</i> | mão | <i>mână</i> | < lat. * <i>mana</i> < <i>manum</i> |
| 49. | <i>belly</i> | barriga | <i>burtă</i> | (desconhecida) |
| | | | <i>pântece</i> | < lat. <i>pantīcem</i> |
| 50. | <i>neck</i> | pescoço | <i>gât</i> | (incerta) |
| 51. | <i>breast(s)</i> | seio | <i>sân</i> | < lat. <i>sīnum</i> |
| 52. | <i>heart</i> | coração | <i>inimă</i> | < lat. <i>animam</i> |
| 53. | <i>liver</i> | fígado | <i>ficat</i> | < lat. <i>ficātum</i> , da loc. <i>iecur ficātum</i> |

Fonte: Elaboração própria

Observações:

- Em (37), o substantivo *păr* (“cabelo”, “pelo”) vem do latim *pīlum*, que é também o étimo português *pele*. Em *păr* (< *pīlum*), houve o típico rotacismo do *l* intervocálico: [l] > [r] (Dobrinescu, 1992, p. 228);
- Em (42), o substantivo *gură* (“boca”) vem do latim *gulam*, cujo significado original era o de “garganta”, “goela” e, menos frequentemente, o de “boca” ou “gula” (Gaffiot, 2016, p. 648). Nele também ocorreu o rotacismo do *l* intervocálico: [l] > [r].
- Em (46), o romeno usa a mesma palavra para designar “perna” e “pé”: *picior*, que vem do latim *petiōlum* (“pezinho”,

“pedúnculo”, “peciolo”). Para diferenciar essas duas partes do corpo, pode-se recorrer à locução *laba piciorului* (literalmente: “a pata da perna”) para designar a ideia de “pé”. A substituição de um substantivo primitivo latino por uma forma diminutiva é fenômeno bastante conhecido nos estudos românicos, como em: *aurem* → *auricūlam* (“orelhinha”) > *oricla* > *orelha* (português), *ureche* (romeno) (Bassetto, 2010, p. 94).

- Em (48), o substantivo *mână* (“mão”) remete ao latim *manus*, mas pode ter vindo de uma forma intermediária hipotética **mana* (Ciorănescu, 1966). A mudança na terminação talvez se explique por se tratar de nome feminino;
- Em (49), há dois substantivos para “barriga”: *burtă* e *pânțec*. O primeiro tem origem desconhecida, enquanto o segundo vem do latim *pantīcem*, de onde se originou o português *pança*;
- Em (50), o substantivo *gât* (“pescoço”) tem etimologia obscura. Sugerem-se o latim *guttūra*, neutro plural de *guttus*, *-uris* (“goela”, “garganta”), e o termo eslavo *glūtŭ* (Ciorănescu, 1966);
- Em (52), o substantivo *inimă* (“coração”) vem do latim *anīmam* (“sopro”), do qual provém o cognato português *alma*. Segundo o supradito Ciorănescu (1966), o semantismo *sopro* >> *coração* ocorreu apenas em romeno.

Tabela 7 – Verbos de ação, percepção e declaração

| N.º | Inglês | Português | Romeno | Etimologia |
|-----|--------------|-----------|----------------|---|
| 54. | <i>drink</i> | beber | <i>a bea</i> | < lat. <i>bibĕre</i> |
| 55. | <i>eat</i> | comer | <i>a mânca</i> | < lat. * <i>manucare</i> < <i>manducāre</i> |
| 56. | <i>bite</i> | morder | <i>a mușca</i> | (incerta) |
| 57. | <i>see</i> | ver | <i>a vedea</i> | < lat. <i>vidĕre</i> |
| 58. | <i>hear</i> | ouvir | <i>a auzi</i> | < lat. <i>audĭre</i> |

| | | | | |
|-----|---------------------|--------------|-------------------|--|
| 59. | <i>know</i> | saber | <i>a ști</i> | < lat. <i>scīre</i> |
| | | conhecer | <i>a cunoaște</i> | < lat. * <i>connoscēre</i> < <i>cognoscēre</i> |
| 60. | <i>sleep (verb)</i> | dormir | <i>a dormi</i> | < lat. <i>dormīre</i> |
| 61. | <i>die</i> | morrer | <i>a muri</i> | < lat. * <i>morire</i> < <i>mori</i> |
| 62. | <i>kill</i> | matar | <i>a ucide</i> | < lat. <i>occidēre</i> |
| | | | <i>a omorī</i> | < esl. <i>umoriti</i> |
| 63. | <i>swim</i> | nadar | <i>a înota</i> | < lat. * <i>innotare</i> ← * <i>notare</i> < <i>natāre</i> |
| 64. | <i>fly (verb)</i> | voar | <i>a zbura</i> | < lat. * <i>exvolare</i> < <i>volāre</i> |
| 65. | walk | andar | <i>a merge</i> | < lat. <i>mergēre</i> |
| 66. | <i>come</i> | vir | <i>a veni</i> | < lat. <i>venīre</i> |
| 67. | <i>lie</i> | jazer | <i>a zăcea</i> | < lat. <i>iacēre</i> |
| 68. | <i>sit</i> | sentar | <i>a ședea</i> | < lat. <i>sedēre</i> |
| | | | <i>a așeza</i> | < lat. * <i>assēdiāre</i> ← <i>sedēre</i> |
| 69. | <i>stand</i> | estar, ficar | <i>a sta</i> | < lat. <i>stāre</i> |
| 70. | <i>give</i> | dar | <i>a da</i> | < lat. <i>dāre</i> |
| 71. | <i>say</i> | dizer | <i>a zice</i> | < lat. <i>dicēre</i> |
| | | | <i>a spune</i> | < lat. <i>exponēre</i> |

Fonte: Elaboração própria

Observações:

- Em (55), o verbo *a mânca* (“comer”) é cognato da forma francesa *manger* e da italiana *mangiare*, que provêm do latim *manducāre* (“devorar”);
- Em (56), o verbo *a mușca* (“morder”) tem origem incerta. Talvez venha da forma hipotética latina **muticāre*, e esta, de *mutīre* (“murmurar”), cuja origem é onomatopaica (Ciorănescu, 1966). Outros étimos foram propostos;
- Em (58) e (71), percebe-se, nos verbos *a auzi* (latim < *audīre*) e *a zice* (latim < *dicēre*), um fenômeno bem típico do

consonantismo romeno: a assibilação de [d] diante de vogais anteriores. Logo: *di* > *zi* (Dobrinescu, 1992, p. 228);

- Em (59), foi necessário utilizar dois verbos distintos para traduzir o verbo inglês *to know*: *a ști* (“saber”) e *a cunoaște* (“conhecer”). É interessante notar que o romeno tenha herdado ao latim o verbo *scīre* (> *a ști*), pois, na România Ocidental, ele foi substituído por *sapĕre* (“ter sabor de”), como em: *saber* (português e castelhano), *savoir* (francês) e *sapere* (italiano). O verbo *a cunoaște*, por seu turno, é cognato do português *conhecer*, do castelhano *conocer*, do italiano *conoscere* e do francês *connaître* – todos oriundos de uma forma vulgar hipotética **connoscĕre*, derivada de *cognoscĕre* (Ciorănescu, 1966);
- Em (62), há dois verbos com a ideia de “matar”: *a ucide* e *a omorî*. O primeiro vem do latim *occidĕre*, que é também o étimo do italiano *uccidere*. O segundo vem do eslavo *umoriti*, de que se originou o verbo russo homossemântico: *уморить* (*umorit*) (Ciorănescu, 1966);
- Em (63), o verbo romeno *a înota* (“nadar”) provém de uma forma hipotética, **innotāre*, e esta, de outra forma hipotética, **notāre*, advinda da forma clássica *natāre*. É desta última que provêm as formas portuguesa e castelhana (*natāre* > *nadar*) (Ciorănescu, 1966).
- Em (64), o verbo romeno *a zbură* (“voar”) provém do latim **exvolare* (< *volare*), de cujo radical também se originou o português *voar*. Chama a atenção o fato de *a zbură* e *voar* terem o mesmo radical, pois o consonantismo das duas línguas os diferenciou muito na pronúncia, tornando-os mutuamente irreconhecíveis. No caso em tela, nota-se outra vez o já citado rotacismo do *l* intervocálico: [l] > [r] (Dobrinescu, 1992, p. 228);
- Em (65), o verbo *a merge* (“andar”) provém do latim *mergĕre* (“mergulhar”). Supõe-se que, no semantismo *mergulhar* >> *andar*, teria havido um significado intermediário de “cair” (Ciorănescu, 1966);

- Em (68), o verbo romeno *a ședea* significa “estar sentado”, mantendo aí quase intacto o significado etimológico do verbo latino *sedēre*. Em português, como é sabido, o infinitivo *sedēre* substituiu o infinitivo latino irregular *esse* (do verbo *sum, es, esse, fui*) e tornou-se o hodierno *ser* (< *seer* < *sedēre*) (Cunha, 2010, p. 590). O verbo *a așeza* designa o ato ou a ação de se sentar e provém de um hipotético **assēdiāre*, derivado do mesmo verbo *sedēre* acima mencionado (Coteanu, 2009);
- Em (69), o verbo *a sta* (“ficar”, “permanecer”) manteve em boa medida o significado etimológico do verbo latino *stāre* (“estar de pé”, “estar fixado”). Convém lembrar que o romeno não faz a distinção entre *ser* e *estar*, à maneira do português e do castelhano. Em ambos os casos, usa-se apenas o verbo *a fi* (“ser”, “estar”);¹⁴
- Em (71), há dois verbos romenos para o verbo *dizer*: *a zice* e *a spune*. O primeiro é cognato da forma portuguesa *dizer*, oriunda do latim *dicēre*, enquanto o segundo vem de *exponēre*, étimo da forma portuguesa *expor*, que também pode, às vezes, empregar-se como verbo declarativo.

Tabela 8 – Elementos do meio ambiente

| N.º | Inglês | Português | Romeno | Etimologia |
|-----|--------------|-----------|---------------|---|
| 72. | <i>sun</i> | sol | <i>soare</i> | < lat. <i>sōlem</i> |
| 73. | <i>moon</i> | lua | <i>lună</i> | < lat. <i>lūnam</i> |
| 74. | <i>star</i> | estrela | <i>stea</i> | < lat. <i>stēllam</i> |
| 75. | <i>water</i> | água | <i>apă</i> | < lat. <i>aquam</i> |
| 76. | <i>rain</i> | chuva | <i>ploaie</i> | < lat. <i>*pluvia</i> < <i>pluvīam</i> |
| 77. | <i>stone</i> | pedra | <i>piatră</i> | < lat. <i>petram</i> < gr. <i>πέτρα</i> |
| 78. | <i>sand</i> | areia | <i>nisip</i> | < búlg. <i>nasip</i> ou esl. <i>nasŭpŭ</i> (?) |
| 79. | <i>earth</i> | terra | <i>pământ</i> | < lat. <i>*pauimentum</i> < <i>pavīmentum</i> |

¹⁴ Com efeito, as sentenças portuguesas *Ele é doente* e *Ele está doente* se traduzem da mesma maneira para o romeno: *El este bolnav*.

| | | | | |
|-----|--------------------|--------------------|---------------|---|
| 80. | <i>cloud</i> | nuvem | <i>nor</i> | < lat. <i>nubĭlum</i> |
| 81. | <i>smoke</i> | fumaça | <i>fum</i> | < lat. <i>fumum</i> |
| 82. | <i>fire</i> | fogo | <i>foc</i> | < lat. <i>focum</i> |
| 83. | <i>ash</i> | cinzas | <i>cenușă</i> | < lat. * <i>cinusia</i> ← <i>cinis</i> |
| | | | <i>scrum</i> | (desconhecida) |
| 84. | <i>burn (verb)</i> | queimar | <i>a arde</i> | < lat. <i>ardĕre</i> |
| 85. | <i>path</i> | caminho, via | <i>drum</i> | < esl. <i>drumŭ</i> < gr. <i>δρόμος</i> |
| | | | <i>cale</i> | < lat. <i>callem</i> |
| 86. | mountain | montanha, monte | <i>munte</i> | < lat. <i>montem</i> |

Fonte: Elaboração própria

Observações:

- Em (75), o substantivo *apă* (“água”) apresenta o típico consonantismo romeno *qu* > *p* diante de *a*. Logo: *aquam* > *apă* (Dobrinescu, 1992, p. 228);
- Em (76), o substantivo *ploaie* vem de uma forma hipotética **plovĭa*, e esta da forma clássica *pluvĭam*, étimo de *chuva*, em português. A palatalização do encontro consonantal [pl] em [ʃ] obscurece, à primeira vista, os elos etimológicos entre a palavra romena *ploaie* e a sua cognata portuguesa *chuva*, de que também são cognatas as palavras *lluvia* (castelhano), *pluie* (francês) e *pioggia* (italiano) (Bassetto, 2010, p. 86-87);
- Em (78), o substantivo *nisip* (“areia”) chegou ao romeno pelo búlgaro *nacun* (*nasip*) (Coteanu, 2009) ou, mais remotamente, pelo eslavo *nasŭpŭ* (Ciorănescu, 1966);
- Em (79), o substantivo *pământ* (“terra”) provém do latim *pavimentum*, de que se originou o português *pavimento*. O romeno possui também o substantivo *țară* (< lat. *terram*), que se emprega mais no sentido de “país”, “nação” ou “território” (Buescu, 2003, p. 448);

- Em (82), encontra-se a palavra *foc* (< latim *focum*), cognata da forma portuguesa *fogo*. O romeno tem ainda a palavra *jar*, que significa “brasa” ou figurativamente “fogo”.
- Em (83), o substantivo *cenușă* (“cinza”) teria vindo do latim **cinusia* (← *cinus*), do qual também proviria a forma cognata portuguesa *cinza*, porém *scrum* tem origem incerta. Ciorănescu (1966) supõe que tenha vindo da palavra cumana¹⁵ *kurum* (“fuligem”);
- Em (85), um dos substantivos para “caminho” ou “via”, *drum*, provém do grego *δρόμος* (“pista”), mas chegou ao romeno através do eslavo *drumŭ*. Em português, *dromo* se usa como substantivo pleno e na formação de vários compostos: *hipódromo*, *autódromo*, etc. Já o substantivo *cale* (latim < *callem*) é cognato do castelhano *calle* (“rua”) (Ciorănescu, 1966).

Tabela 9 – Cores

| N.º | Inglês | Português | Romeno | Etimologia |
|-----|---------------|--------------|---------------|------------------------------------|
| 87. | <i>red</i> | vermelho | <i>roșu</i> | < lat. <i>rosĕum</i> |
| 88. | <i>green</i> | verde | <i>verde</i> | < lat. <i>virĕm</i> < <i>vīrĕm</i> |
| 89. | <i>yellow</i> | amarelo | <i>galben</i> | < lat. <i>galbĭnum</i> |
| 90. | <i>white</i> | branco | <i>alb</i> | < lat. <i>album</i> |
| 91. | <i>black</i> | preto, negro | <i>negru</i> | < lat. <i>nĭgrum</i> |

Fonte: Elaboração própria

Observações:

- Em (87), o adjetivo *roșu* (“vermelho”) vem do latim *rosĕum*, étimo do português *róseo*. Na transição de *rosĕum* para

¹⁵ Os cumanos eram um povo seminômade de origem turca que migrou da Ásia Central para o Leste Europeu na Idade Média. Na segunda metade do século XI, eles se assenhorearam de algumas terras pertencentes à atual Romênia, de onde foram desterrados devido à invasão mongol, no século XIII (Niculescu, 1983, p. 63-64).

roșu, ocorreu a típica palatalização da sibilante [s] diante de semivogal [j] ou vogal anterior: [s] > [ʃ] (Dobrinescu, 1992, p. 228);

- Em (91), o adjetivo *negru* vem do latim *nīgrum*, que é um cognato bem evidente da forma portuguesa *negro*. Em ambas as formas, romena e portuguesa, ocorreu o abaixamento da vogal breve *ī* para [e]. O romeno tem ainda o adjetivo *lai* (“preto”), que é forma popular (Buescu, 2003, p. 248).

Tabela 10 – Palavras diversas

| N.º | Inglês | Português | Romeno | Etimologia |
|------|--------------|-----------|---------------|---|
| 92. | <i>night</i> | noite | <i>noapte</i> | < lat. <i>noctem</i> |
| 93. | <i>hot</i> | quente | <i>cald</i> | < lat. <i>caldum</i> < <i>calīdum</i> |
| 94. | <i>cold</i> | frio | <i>frig</i> | < lat. <i>frīgus</i> ¹⁶ |
| | | | <i>rece</i> | < lat. <i>recens</i> |
| 95. | <i>full</i> | cheio | <i>plin</i> | < lat. <i>plēnum</i> |
| 96. | <i>new</i> | novo | <i>nou</i> | < lat. <i>novum</i> |
| 97. | <i>good</i> | bom | <i>bun</i> | < lat. <i>bonum</i> |
| 98. | <i>round</i> | redondo | <i>rotund</i> | < lat. <i>rotundum</i> |
| 99. | <i>dry</i> | seco | <i>uscat</i> | ← <i>a usca</i> < lat. <i>*usticare</i> |
| | | | <i>sec</i> | < lat. <i>siccum</i> |
| 100. | <i>name</i> | nome | <i>nume</i> | < lat. <i>nomen</i> |

Fonte: Elaboração própria

¹⁶ *frigus*, -ōris (substantivo neutro da terceira declinação).

Observações:

- Em (94), há duas palavras para “frio”: *frig* e *rece*. A primeira é da mesma família lexical do adjetivo português *frio*, oriundo de *frigĭdum* (Cunha, 2010, p. 302). A segunda é cognata do português *recente*, mas note-se que a forma romena veio do nominativo *recens* (“fresco”, “novo”), reduzido a uma forma intermediária **reces* e depois a *rece*, que também pode significar “fresco” em romeno, como em *apă rece* (“água fresca”) (Ciorănescu, 1966; Buescu, 2003, p. 372).
- Em (95), o adjetivo *plin* vem do latim *plenum*, de que se originaram as formas divergentes *cheio* e *pleno* em português. A palatalização do encontro consonantal [pl] em [ʃ] obscurece, à primeira vista, os elos etimológicos entre a palavra romena *plin* e a sua cognata portuguesa *cheio*, de que também são cognatas as palavras *lleno* (castelhano), *plein* (francês) e *pieno* (italiano) (Bassetto, 2010, p. 86-87);
- Em (98), o adjetivo *rotund* (“redondo”) possui uma forma variante: *rătund*. A semelhança física entre a forma *rotund* e o étimo latino *rotundum* sugere tratar-se de um latinismo ou de uma reconstituição erudita, mas Ciorănescu (1966) o nega alegando que a mudança *ră* > *ro* é comum na região da Muntênia, como em: *răcodelie* > *rocodelii* (“trabalho manual”), *răpotini* > *ropotini*¹⁷.
- Em (99), o adjetivo *uscat* (“seco”) é o particípio do verbo *a usca* (“secar”), que viria da forma hipotética latina **usticare* ← *ustum* ← *urere* (“queimar”). O semantismo *queimar* >> *secar* é compreensível. Também se supõe o étimo hipotético **exsucare* (“secar”) para o verbo *a usca* (Coteanu, 2009);

¹⁷ *Răpotin*: feriado que se comemora costumeiramente nas três terças-feiras seguintes à Páscoa (cf. <<https://dexonline.ro/intrare/r%C4%83potin/243931/definitii>>).

3 Resultados e conclusões

Findas as observações etimológicas, é chegada a hora de contabilizar as contribuições lexicais das línguas que ajudaram a formar o vocabulário básico do romeno. Devido à necessidade de utilizar mais de um item lexical para traduzir certas palavras (*e.g. árvore: arbor e copac*)¹⁸, a contagem inicial de 100 itens arrolados na *Lista de Swadesh* subiu para um total de 113. Abaixo, o número absoluto de palavras com que cada língua contribuiu em cada campo semântico:

Tabela 11 – Distribuição dos étimos por campo semântico

| | Latim | Latim Culto | Francês | Eslavo | Búlgaro | Desconhecido/ Incerto | Total |
|--|-------|-------------|---------|--------|---------|--------------------------|-------|
| Palavras de classe fechada (1-12) | 12 | - | - | - | - | - | 12 |
| Adjetivos de tamanho (13-15) | 3 | - | - | - | - | - | 3 |
| Seres vivos (16-22) | 7 | - | 1 | - | - | - | 8 |
| Plantas e suas partes (23-27) | 5 | 1 | - | - | - | 1 | 7 |
| Partes do corpo dos seres vivos (28-53) | 25 | - | - | - | - | 3 | 28 |

¹⁸ Número dos itens para os quais se deu mais de uma tradução nas tabelas: 16, 23, 25, 36, 49, 59, 62, 68, 71, 83, 85, 94 e 99.

| | | | | | | | |
|---------------------------------------|----|---|---|---|---|---|----|
| Verbos de ação, perc. e decl. (54-71) | 20 | - | - | 1 | - | 1 | 22 |
| Elementos do meio ambiente (72-86) | 14 | - | - | 1 | 1 | 1 | 17 |
| Cores (87-91) | 5 | - | - | - | - | - | 5 |
| Palavras diversas (92-100) | 11 | | - | - | - | - | 11 |

Fonte: Elaboração própria

Com base nos dados apresentados na tabela 11, chega-se aos seguintes números e às seguintes percentagens, que representam quantitativamente as contribuições lexicais de cada língua para a formação do vocabulário básico romeno. Como o búlgaro pertence à família eslava, convém agrupar as suas contribuições lexicais àquelas das línguas eslavas em geral (lembrando: 113 itens lexicais = 100%):

Tabela 12 – Contribuição de cada língua para a formação do vocabulário básico romeno (em números absolutos e percentuais)

| | Latim | Latim Culto | Francês | Línguas eslavas | Desconhecido/ Incerto | Total |
|---------------------|--------|-------------|---------|-----------------|-----------------------|-------|
| Números absolutos | 102 | 1 | 1 | 3 | 6 | 113 |
| Números percentuais | 90,26% | 0,88% | 0,88% | 2,65% | 5,31% | 100% |

Fonte: Elaboração própria

A constatação de que 90,26% do vocabulário básico da língua romena tem origem latina é uma prova peremptória e incontestada de sua latinidade. Afinal, 102 das 113 palavras estudadas vieram do latim e foram transmitidas de geração em geração desde épocas remotas até a Idade Contemporânea – motivo pelo qual se optou por deixar em separado os termos oriundos do latim culto e do francês, que só forneceram vocábulos ao romeno nos últimos séculos, noutros contextos históricos.

A título de comparação, Dana Cojocar (2003, p. 11) traz estas percentagens sobre o léxico romeno geral:

- (1) Latim: 20%;
- (2) Línguas eslavas: 14%;
- (3) Grego: 2,37%;
- (4) Francês: 38,4%;
- (5) Latim culto: 2,4%;
- (6) Italiano: 1,7%;
- (7) Turco: 3,7%;
- (8) Húngaro: 2,17%;
- (9) Línguas germânicas: 2,3%;
- (10) Outras origens / incertos: 12,96%.

O visível contraste entre os 90,26% de vocabulário latino básico, obtidos por meio da aplicação da *Lista de Swadesh*, e os 20% de léxico latino geral é muito elucidativo. Os números ratificam a latinidade do léxico romeno primitivo e mostram que os seus falantes recorreram

sobremaneira ao francês para expandir o seu vocabulário técnico e literário, a ponto de os vocábulos de origem francesa constituírem hoje a maior porção lexical: 38,4%.

O léxico romeno ora se assemelha às línguas românicas ocidentais, ora apresenta matizes próprios. As palavras de classe fechada são bem inteligíveis a um falante de português: *eu, tu, el, noi, voi, ei, cu, cine, acest, acel, etc.* Em certos casos, o consonantismo romeno fez com que certos cognatos ficassem simplesmente irreconhecíveis (e.g. *a zbură* < **exvolare* < *volare*, cognato de *voar*). Noutros casos, ele herdou do latim termos não herdados pelas línguas românicas ocidentais. Por exemplo, o verbo romeno *a ști* (“saber”) vem da forma clássica *scīre*, enquanto os seus correspondentes homossemânticos *saber* (português), *saber* (castelhano), *savoir* (francês) e *sapere* (italiano) vêm de *sapĕre* (“ter sabor de”). Os numerosos comentários etimológicos ajudam a explicar a fragmentação do latim e mostram que o romeno, longe de ser um idioma exótico, tem mais semelhanças com os membros da família românica do que aparenta.

Referências

ALGEO, J. *The origins and development of the English language*. Boston: Wadsworth, 2010.

BASSETTO, B. F. *Elementos de filologia românica: história externa das línguas românicas*, v. 1. São Paulo: Editora da Universidade de São Paulo, 2013.

BASSETTO, B. F. *Elementos de Filologia Românica: história interna das línguas românicas*, v. 2. São Paulo: Editora da Universidade de São Paulo, 2010.

BUESCU, V. (coord.). *Dicionário de romeno-português*. Porto: Porto Editora, 2003.

CIORĂNESCU, Alexandru. *Dicționarul etimologic român*. Tenerife: Universidad de la Laguna, 1966.

COJOCARU, D. *Romanian grammar*. Durham, NC: SEERLRC/Duke University, 2003. Disponível em: <www.seelrc.org:8080/grammar/pdf/compgrammar_romanian.pdf>. Acesso em: 15 jan. 2018.

COTEANU, I. (ed.). *Dicționarul explicativ al limbii române* (ediția a II-a revăzută și adăugită). Bucharest, Romania: Academia Română, Institutul de Lingvistică. Editura Univers Enciclopedic Gold, 2009.

CUNHA, A. G. *Dicionário etimológico da língua portuguesa*. 4. ed. Rio de Janeiro: Lexicon, 2010.

DOBRINESCU, G. *Gramática da língua romena*. Rio de Janeiro: Presença, 1992.

FOX, A. *Linguistic reconstruction: an introduction to theory and method*. New York: Oxford University Press, 1995, 389 p.

GAFFIOT, F. *Dictionnaire Latin Français*. Nouvelle édition revue et augmentée. Version V. M. Komarov.[S. l.]: Gérard Gréco, 2016.

GARCÍA, M. I. M.; SOUZA, A. M. B. de. Lexical similarity level between English and Portuguese. *ELIA*, Sevilla, v. 14, n. 14, 2014, p. 145-163. DOI: <http://dx.doi.org/10.12795/elia.2014.i14.06>

GRANT, A. P. Swadesh's life and place in linguistics. *Diachronica*, v. 27, n. 2, 2010. Disponível em: <https://www.academia.edu/10649391/Morris_Swadeshs_life_and_place_in_linguistics>. Acesso em: jan. 2020.

HUBER, J. *Gramática do Português Antigo*. Tradução de Maria Manuela Gouveia Delille. Coimbra: Fundação Calouste Gulbekian, 1986.

HUNTLEY, D. The Old Church Slavonic. In: COMRIE, B.; CORBETT, G. (ed.). *The Slavonic languages*. New York: Routledge, 1993. p. 125-187.

KLUGE, F. *Etymologisches Wörterbuch des deutschen Sprache*. Strasburg: Karl Trübner, 1899.

MAURER Jr., Theodoro Henriques. *A unidade da România Ocidental*. São Paulo: USP, 1951.

NICULESCU, A. *História breve da língua romena*. Rio de Janeiro: Presença Edições; São Paulo: Ed. da Universidade de São Paulo, 1983.

ȘĂINEANU, L. *Dicționar universal al limbei române*. ed. VI. Craiova: Scrisul Romanesc, 1929.

SCRIBAN, A. *Dicționarul limbii românești*. [S. l.]: Institutu de Arte Grafice Presa Bună, 1939.

SWADESH, M. Towards Greater Accuracy in Lexicostatistic Dating, *International Journal of American Linguistics*, v. 21, n. 2, 1955. p. 121-137.

VIARO, M. E. *Etimologia*. São Paulo: Contexto, 2011.

Recebido em: 15 de maio de 2023.
Aprovado em: 08 de junho de 2023.