



Passado, presente e futuro dos dicionários de aprendizes

Learner's Dictionaries: past, present and future

Laura Campos de Borba*
Félix Bugueño Miranda**

RESUMO: Neste artigo, traçamos um panorama sobre o desenvolvimento da lexicografia de aprendizes, a qual produz uma classe específica de obras lexicográficas destinada a servir de auxílio no processo de ensino-aprendizagem de uma língua estrangeira. O objetivo do trabalho é apresentar uma síntese histórico-crítica desse gênero específico de dicionários, estabelecendo o estado da arte em que se encontra e procurando, ao mesmo tempo, determinar quais áreas do conhecimento oferecem subsídios para melhorar ainda mais esses dicionários. Como metodologia, aplicam-se princípios advindos da metalexicografia, tais como a necessidade de que toda informação do dicionário corresponda a um fato de norma real; além disso, emprega-se também o princípio de que o usuário deve poder tirar proveito de uma informação para resolver uma tarefa linguística. Esses princípios são testados, por sua vez, à luz de disciplinas como a Linguística de Corpus, a análise de erros e a cognição, e

ABSTRACT: In this paper we intend to establish an historical development of a special class of pedagogical lexicography, namely, learner's lexicography. Learner's dictionaries are specially designed as a support to foreign language-learning. We aim to present not only an historical panorama of this subject, but also we are especially interested in exploring which knowledge areas could help to improve learner's dictionaries. As a methodological framework, we use principles of metalexicography, such the fact that every piece of information contained in a dictionary must correspond to the language in use. A second metalexicographical principle is that the user must profit from the lexicographical information in order to use the language. Subsequently, both principles are tested together in the light of disciplines as corpus linguistics, error analysis and cognition; furthermore, we analyze these principles with the aid of data obtained from learner's corpora (in particular, errors performed by foreign-language

* Doutoranda em Estudos da Linguagem (Lexicografia) pelo Programa de Pós-Graduação em Letras da UFRGS. ORCID: <https://orcid.org/0000-0001-5316-9334>. lauraborba@hotmail.com. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

** Doutor em Filologia Românica (Universidade de Heidelberg). Instituto de Letras, Universidade Federal do Rio Grande do Sul. ORCID: <https://orcid.org/0000-0001-6234-101X>. felixv@uol.com.br.

ainda por meio de ferramentas como os *corpora* de aprendizes (particularmente no que se refere a erros). Como resultado da aplicação dessa metodologia, comprovamos que ditas áreas fornecem dados e perspectivas extremamente úteis para o desenho de dicionários de aprendizes do espanhol e do inglês. Para concluir, destacamos que a lexicografia de aprendizes deve ter um caráter claramente interdisciplinar.

PALAVRAS-CHAVE: Metalexicografia. Lexicografia de aprendizes. Interdisciplinaridade.

students). Our results prove that all these areas offer a very useful knowledge for the design of English learners' dictionaries and Spanish learners' dictionaries. We can conclude that the learner's lexicography must be an essentially interdisciplinary area.

KEYWORDS: Metalexicography. Learner's dictionaries. Interdisciplinary approach.

1. Introdução

Os livros didáticos e as gramáticas podem ser considerados como materiais primários no ensino de línguas, tanto de língua materna como de língua estrangeira ou adicional¹. No campo de investigação da lexicografia, há uma vertente específica que se preocupa com dicionários que auxiliem no ensino-aprendizagem de línguas: a lexicografia pedagógica². Na prática, todo instrumento lexicográfico apresenta um viés pedagógico justamente por estar relacionado à função de orientar o consulente a respeito de questões relativas ao uso da língua. Por analogia ao termo "pedagógica", poder-se-ia deduzir que os dicionários produzidos para fins pedagógicos se diferenciariam das demais vertentes da lexicografia por considerar dois fatores próprios do ensino-aprendizagem de uma dada língua estrangeira: o nível de aprendizagem da língua no qual se encontra o usuário visado (isto é, a proficiência adquirida em níveis anteriores) e as necessidades desse estudante (ou seja, o conteúdo da(s) etapa(s) pela(s) qual(is) o estudante deve seguir). A própria definição de

¹ Embora se reconheça que existe uma distinção conceitual entre ambos os termos, para efeitos deste trabalho, essa distinção não afeta os argumentos defendidos aqui.

² É pertinente observar que os dicionários de aprendizes de língua materna também se inserem nessa vertente. Porém, como não formam parte do escopo deste projeto, prescinde-se de maiores comentários.

dicionário pedagógico aponta para essa hipótese: “uma obra de referência desenhada especificamente para as necessidades didáticas práticas de professores e estudantes de uma língua”³ (HARTMANN; JAMES, 2001, s.v. *pedagogical dictionary*. Tradução nossa). Na prática, no entanto, não é possível garantir a priori que essa relação, *de facto*, exista.

No âmbito da Lexicografia Pedagógica, é possível distinguir duas vertentes. Por um lado, há dicionários desenhados para auxiliar no ensino-aprendizagem da língua materna, conhecidos no meio nacional como dicionários escolares. Por outro lado, há uma segunda vertente, menos conhecida no Brasil, e que se preocupa com dicionários monolíngues para aprendizes de uma língua estrangeira. Trata-se da Lexicografia de Aprendizes (HARTMANN; JAMES, 2001, s.v. *learner lexicography*).

O objetivo do presente trabalho é apresentar uma síntese crítica sobre o desenvolvimento da Lexicografia de Aprendizes visando determinar em que medida esses dicionários acompanham o estado da arte na pesquisa em Linguística.

2. Passado

A primeira das tradições lexicográficas que se preocupou (e ainda se preocupa) com dicionários de aprendizes é a de língua inglesa. Durante o século XX, observou-se o advento não apenas de uma nova classe de obras lexicográficas, mas também de novas propostas de descrição da língua com o auxílio de *corpora*⁴.

O primeiro marco na história da lexicografia de aprendizes de língua inglesa foi a publicação da obra que hoje é conhecida como *Oxford Advanced Learner's Dictionary* (OALD8, 2010). Albert Hornby, coordenador do projeto do dicionário à época, idealizou-o enquanto trabalhava como professor de inglês em uma universidade do

³ Cf. trecho original: “A reference work specifically designed for the practical didactic needs of teachers and learners of a language”.

⁴ Corpora tais como o British National Corpus ou o Word Banks (Collins), são empregados tanto para a descrição sistemática da língua inglesa na forma de gramáticas, como também para a compilação de dicionários.

Japão durante os anos 30 (JACKSON, 2002, p. 129). Incentivado por sua experiência docente, Hornby compilou uma obra pioneira principalmente no que concerne ao tratamento de padrões sintáticos, com foco na complementação sintática dos verbos e nas combinatórias léxicas do inglês. Outro aspecto pelo qual sua obra se destacou diz respeito ao fornecimento de numerosos exemplos e definições simplificadas (RANALLI; NURMUKHAMEDOV, 2014, p. 2). As primeiras edições do atual OALD8 (2010) refletem, assim, os pressupostos do método audiolingual na medida em que se dá importância fundamental à descrição sintática (HALL, 2016, p. 213).

Apesar de inovadora, Jackson (2002, p. 131) destaca que a proposta de Hornby para os padrões de complementação sintática na terceira edição do dicionário dificultava a consulta. Isto porque se empregavam códigos diferentes para indicar cada um dos 25 padrões básicos, aos quais se somavam 50 padrões secundários. Os consulentes da obra tinham problemas em memorizar tamanha quantidade de códigos, o que os forçava muitas vezes a retornar às páginas iniciais da obra para conferir de que padrão se tratava, onerando a consulta. Isto demonstra que, se o dicionário está desenhado para auxiliar no ensino-aprendizagem de uma língua estrangeira, uma descrição apurada da língua é, evidentemente, uma *conditio sine qua non*. No entanto, tão importante como o anterior é a apresentação de fatos linguísticos de uma forma que o aprendiz consiga interpretar as informações sem dificuldade.

Desde o final dos anos 1940 até o fim da década de 1970, o atual OALD8 (2010) gozou do status de ser o único dicionário de aprendizes de inglês como língua estrangeira. Em 1978, entretanto, foi publicada uma segunda obra com a mesma finalidade: o atual *Longman Dictionary of Contemporary English* (LDOCE, 2014), editado por Paul Proctor⁵. Conforme Jackson (2002, p. 130-131), o dicionário foi pioneiro na aplicação de um vocabulário de definidores a ser empregado na redação das definições. Tal vocabulário era composto das 2000 palavras mais comuns do inglês e

⁵ A 1ª edição foi publicada em 1978.

estava inserido em um movimento mais amplo em prol do controle de vocabulário (“vocabulary control”) que diferenciase os dicionários de aprendizes dos dicionários para falantes nativos. O objetivo da proposta era facilitar a compreensão de palavras possivelmente desconhecidas por meio de palavras que os estudantes já teriam aprendido em etapas anteriores de estudo do inglês (já que este e muitos dos demais dicionários para aprendizes visam estudantes de níveis mais avançados). Nota-se uma preocupação maior com o consulente ao se estabelecer um subconjunto léxico que permita a redação de definições de fácil compreensão. Contudo, as pesquisas posteriores demonstrarão que um vocabulário de definidores não garante, automaticamente, a compreensão das definições contidas no dicionário.

Além do vocabulário de definidores, Proctor propôs melhorias em dois outros aspectos. Por um lado, e segundo Rundell (2008, p. 223), o dicionário apresentava, pela primeira vez na lexicografia inglesa, uma ampla cobertura não apenas do inglês britânico, mas também do inglês americano. Por outro lado, empregou uma metodologia alternativa à de Hornby para tratar da questão dos padrões de complementação sintática. No entanto, tal metodologia manteve o uso de códigos para identificar cada padrão, o que resultou novamente em um ônus para o consulente. Jackson (2002, p. 131), por exemplo, observa que os consulentes optavam por recorrer aos exemplos para depreender os padrões sintáticos durante a consulta.

O terceiro marco na lexicografia de aprendizes de língua inglesa ocorreu ao final dos anos 80, com a publicação do atual *Collins Cobuild Advanced Dictionary* (CCAD, 2008)⁶, dirigido por John Sinclair. Trata-se do primeiro dicionário da lexicografia de língua inglesa a empregar um corpus computadorizado (*Bank of English / WordBanks*) (WB, 2005), fruto de um movimento conhecido como Revolução do Corpus (“Corpus Revolution”).

⁶ A 1ª edição foi publicada em 1987.

O emprego de um corpus computadorizado na redação do dicionário permitiu que se investigasse de maneira mais aprofundada a frequência de uso tanto de unidades léxicas como de seus significados. Tais informações eram incluídas no dicionário sob a forma de símbolos para indicar a frequência de uso de cada palavra-entrada e na ordenação das acepções da mais frequente para a menos frequente. Os exemplos, por sua vez, eram extraídos diretamente do corpus, recebiam poucas adaptações e estavam presentes junto a cada acepção. Segundo Lew (2013, p. 1), o emprego da frequência como critério de ordenação das acepções, a inclusão de exemplos extraídos diretamente do corpus e muitas vezes sem modificações, entre outras decisões, refletem uma metodologia específica: *corpus-driven*. Por meio dessa metodologia, são fornecidas informações considerando-se uma única variável: o corpus como princípio e fim da descrição da língua.

Além das inovações provenientes do corpus, outro aspecto que distingue a obra de Sinclair ainda hoje é a redação das definições por meio da técnica *whole-sentence definition* (JACKSON, 2002, p. 131). Conforme Farias (2009, p. 76), “a *whole-sentence definition* normalmente conforma um período composto por duas orações: a primeira coloca a palavra-entrada em um contexto, e a segunda apresenta a definição propriamente dita”. O verbete *sunrise*, por exemplo, em sua primeira acepção, contextualiza o período do dia no qual o fenômeno em questão ocorre: “Sunrise is the time in the morning when the sun first appears in the sky” (CCAD, 2008, s.v. *sunrise*). As *whole-sentence definitions* demonstram ser especialmente úteis para a definição de verbos. Isto acontece porque a definição de uma expressão de ação une muito bem intensão e extensão⁷. Não há, no entanto, como saber se a equipe do CCAD (2008)

⁷ Intensão, conforme o MLS (2010, s.v. Intension), é a “designação para a determinação de um conceito por meio do seu conteúdo” [Bezeichnung für die Bestimmung eines Begriffs durch seinen Begriffsinhalts]. Já a extensão é “o conjunto de objetos aos quais se estende uma mesma designação” [die Menge der Gegenstände, auf die der Ausdruck zutrifft] (MLS, 2010, Extension).

estava ciente desse fato. Assim, a discussão sobre a relação entre intensão e extensão é ainda pertinente.

CCAD (2008) constitui uma síntese entre oferecer uma imagem da língua real, que é a língua que o aprendiz estuda, e uma preocupação maior com o destinatário final do dicionário, que é, evidentemente, o próprio aprendiz. O emprego de *corpora* garante que toda informação presente no dicionário corresponda a um fato de norma real. No que diz respeito ao usuário, e na esteira dos vocabulários de definidores, a adoção da técnica de *whole-sentence definition* constitui um avanço. Porém, essa preocupação não se estende aos outros segmentos informativos. Em outras palavras, o viés pedagógico da obra parece restringir-se unicamente à definição.

As novas edições dos dicionários de aprendizes das editoras Longman e Oxford e as obras que surgiram após o irrompimento da Revolução do Corpus também passaram a buscar respaldo em bases de dados. O quarto representante que passou a formar parte da lexicografia de aprendizes de língua inglesa hoje é conhecido como *Cambridge Advanced Learner's Dictionary* (CALD4, 2013). De acordo com Jackson (2002, p. 132), as inovações da obra incluem um maior aproveitamento dos recursos advindos de *corpora* em ao menos três aspectos: a) a cobertura das variedades britânica, americana e também australiana do inglês⁸; b) o emprego de um *corpus* de aprendizes de inglês como língua estrangeira para fornecer tabelas de falsos cognatos em relação a 16 outras línguas; e c) a inclusão de colocações e fraseologias.

Outra proposta notável diz respeito ao tratamento dos verbetes polissêmicos. Cada acepção é introduzida por um distinguidor semântico [*semantic signpost*]⁹,

⁸ A opção de incluir a variedade do inglês australiano é uma decisão extremamente acertada. No entanto, no que diz respeito à diatopia, é evidente que, paulatinamente, outras variedades do inglês, tais como o inglês canadense e o inglês sul-africano vão cobrando cada vez mais relevância (cf. CRYSTAL, 2003, p. 36-46). Embora fuja do foco do presente trabalho, uma redefinição do escopo diatópico dos dicionários de aprendizes de inglês é uma tarefa ineludível.

⁹ Cf. Klotz e Herbst (2016, p. 62). No entanto, é necessário destacar que há uma variação terminológica. Empregam-se também termos como *semantic marker* e *index term*.

¹⁰ Cf. Seção 3.3 para mais detalhes.

recurso que ajuda a identificar as acepções para, assim, facilitar a busca do consulente pela acepção mais adequada, como, por exemplo, s.v. *follow* (*go, happen, understand, be interested*).

Ao total, os dicionários de aprendizes do inglês somam seis obras. Além das quatro já mencionadas, outros dois dicionários que ingressaram mais tardiamente (anos 2000) no conjunto são o *Macmillan English Dictionary for Advanced Learner's* (MEDAL, 2007) e o *Merriam-Webster Advanced Learner's English Dictionary* (MWALE, 2008).

À medida que uma determinada obra lexicográfica propunha novas soluções de inclusão e tratamento de informações linguísticas, os demais dicionários as incorporavam em certa medida. Essa sequência de inovações conduziu a lexicografia de aprendizes de língua inglesa a um patamar bastante avançado de desenvolvimento no tocante à descrição da língua e serviu de incentivo a outras tradições lexicográficas, como as de língua alemã e espanhola.

Em síntese, poder-se-ia dizer que há duas características que distinguem a Lexicografia de Aprendizes do inglês. A primeira característica, relacionada aos primórdios da Lexicografia de Aprendizes, diz respeito ao fato de que dicionário e método de ensino-aprendizagem convergiam. A segunda característica se refere à apresentação apurada de uma descrição da língua, traço marcante dos dicionários de aprendizes do inglês.

3. Presente

O legado da Lexicografia de Aprendizes do inglês é inegável. No entanto, o atual estado da arte da investigação em Lexicografia e Linguística de Corpus, particularmente, demonstra que há aspectos que requerem atenção prioritária.

Em primeiro lugar, pouco se questiona se os dicionários de aprendizes realmente abarcam dimensões tais como o nível de aprendizagem do estudante visado

e as necessidades do mesmo; infere-se que, se são materiais didáticos para estudantes de línguas estrangeiras, estariam preparados para oferecer orientação adequada e informações de qualidade a este público-alvo.

Em segundo lugar, não é possível deduzir se as informações presentes sempre correspondem realmente a fatos de norma (ou seja, se são discretas¹⁰), nem se conseguem auxiliar de maneira efetiva o estudante a quem se destina o dicionário (ou seja, se são discriminantes).

3.1 Perfil de usuário

A formulação de um perfil de usuário constitui um procedimento estratégico que contribui para que as orientações dadas pelo dicionário sejam mais efetivas (ou seja, discriminantes). Conforme Andersen e Nielsen (2009, p. 358) e Verlinde e Binon (2009, p. 86), o consulente pretense é alguém cujas necessidades deveriam guiar o processo de compilação de uma obra. Para outras classes de dicionários, o perfil de usuário constitui um aspecto nebuloso¹¹. A Lexicografia Pedagógica oferece, em tese, a possibilidade de associar um dicionário a um currículo de ensino-aprendizagem, do qual seria possível derivar o nível de proficiência atingido até então e as possíveis necessidades do consulente pretense. No caso do ensino-aprendizagem de línguas estrangeiras, tais necessidades advêm não somente do conteúdo relativo à etapa de aprendizagem na qual o estudante se encontra, mas também da língua materna do aprendiz.

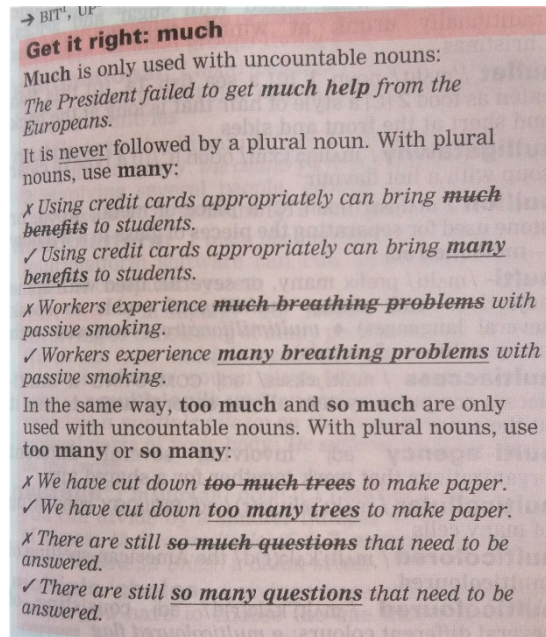
Geralmente, os dicionários de aprendizes se destinam a qualquer estudante da língua alvo como língua estrangeira. Tal metodologia, comum no campo da

¹⁰ Uma informação discreta é aquela que reflete a língua de maneira adequada e que sobrevive à contraprova empírica por meio da análise de registros da língua compilados em um corpus, por exemplo (BUGUEÑO MIRANDA; FARIAS, 2013).

¹¹ A classe dos dicionários gerais de língua é um exemplo disso. O perfil de usuário de qualquer dicionário da classe é sempre genérico.

Lexicografia de Aprendizes, resulta na produção de uma obra cujas informações não poderão satisfazer integralmente as necessidades de um estudante se se considera a sua língua materna. Tome-se, por exemplo, o verbete *much* no MEDAL (2007), o qual está reproduzido a seguir:

Figura 1 – *much*.



Fonte: MEDAL (2007, s. v. *much*).

As informações fornecidas no verbete são capazes de auxiliar um estudante brasileiro de inglês, pois salientam o fato de que há dois adjetivos intensificadores em inglês para expressar a grande quantidade de algo (*much* e *many*), enquanto em português há apenas um (*muito*). Entretanto, há outros casos de divergência entre o português e o inglês que o dicionário não enfatiza justamente porque não se propõe a satisfazer as necessidades de consultantes de uma língua materna específica. Um exemplo disso é o verbo *behave*, que em inglês pode ser intransitivo e transitivo; já o seu equivalente no português é pronominal (*comportar-se*). Somente um dicionário que se dirija a falantes de português como língua materna poderá incluir uma observação sobre essa diferença.

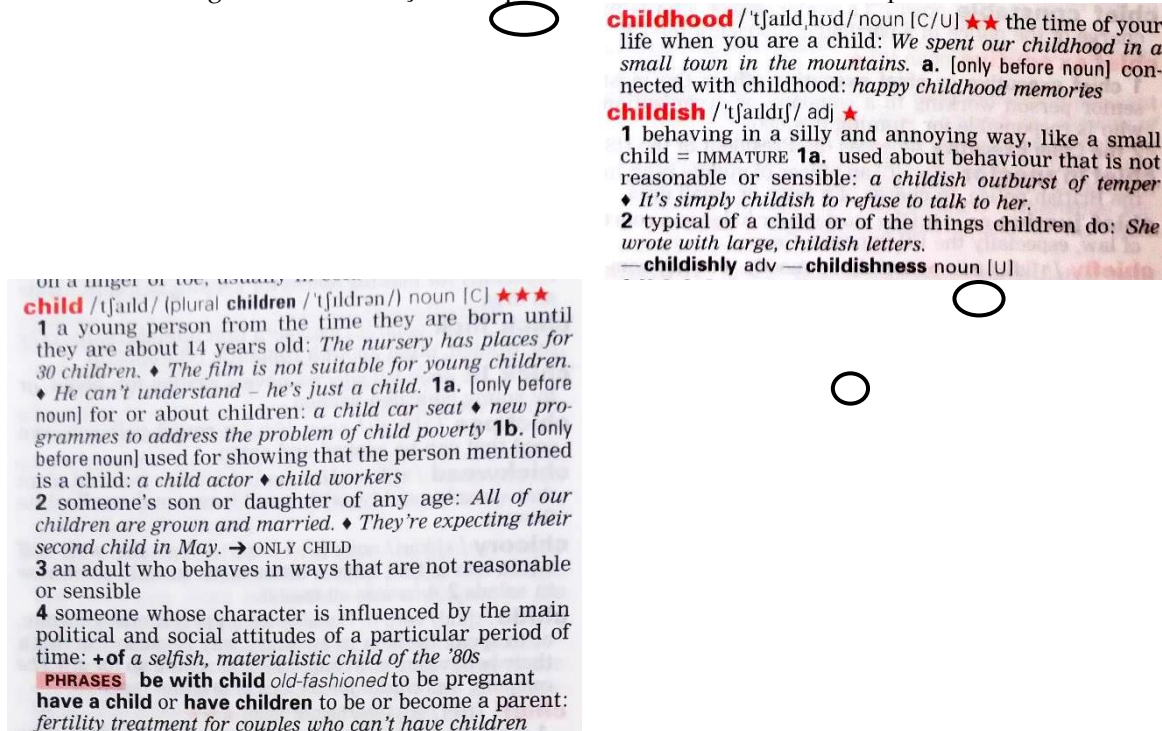
Em suma, a principal consequência da tomada de decisões com base em um perfil de usuário difuso (que, na verdade, é genérico) é a de que a obra lexicográfica oferecerá um auxílio talvez não tão eficaz. Na opinião de Rundell (2010 apud CHI, 2015), o modelo tradicional de dicionários de aprendizes direcionados para todo e qualquer estudante de uma dada língua estrangeira é uma questão ultrapassada, devendo-se ceder o lugar para dicionários “personalizados”.

3.2 Informações discretas e discriminantes

Informações de natureza discreta têm uma relevância dupla para os dicionários de aprendizes. Por um lado, essas obras deveriam refletir rigorosamente a norma contemporânea da língua; por outro lado, a norma necessária para que o estudante de língua estrangeira possa desempenhar-se adequadamente em situações de interação é justamente a contemporânea¹². Um exemplo de fornecimento de informações discretas em dicionários de aprendizes é o caso das marcações de frequência nas unidades léxicas incluídas no MEDAL (2007). As unidades léxicas mais comuns (“most common”) são marcadas com três estrelas; as muito comuns (“very common”), com duas; e as razoavelmente comuns (“fairly common”), com uma estrela, conforme os verbetes reproduzidos na Figura 2 (ver grifos nossos ao redor das estrelas). Ao estar baseado em um corpus de inglês contemporâneo (*World English Corpus*) (WEC, s.d.), não há dúvidas de que os índices de aferição de frequência são condizentes com a norma real do inglês.

¹² Cf. Schlatter *et al.* (2004, p. 350) a respeito do conceito de interação.

Figura 2 – informações de qualidade em um dicionário de aprendizes.



Fonte: MEDAL (2007, s.v. *child, childhood, childish*. Grifos nossos).

Entretanto, na Lexicografia de Aprendizes também há casos de informações não discretas. No dicionário *Collins Cobuild Learner's Dictionary* (CCLD, 2003), por exemplo, apesar de todas as unidades léxicas incluídas estarem documentadas no corpus que serve de base para a compilação da obra – o WB (2005) –, certas unidades léxicas presentes na nominata da obra não são, de fato, discretas. Ao se analisarem dois intervalos lexicográficos de onze páginas nas letras C e S, foram encontradas unidades léxicas com índice de frequência pouco significativo conforme o corpus no qual o dicionário se baseia. Trata-se das unidades *saddle-bag* (9), *carbolic acid* (18), *scandalize* (18), *calve* (21), *saddler* (27), *carouse* (28), *callus* (31), *capon* (31), *satsuma* (33), *scapula* (33), *sapper* (34), *cantonment* (36), *camembert* (38), *samaritan* (38) e *satirize* (39). Não foi possível compreender a razão dessa inclusão; tal decisão parece não condizer com as necessidades do usuário pretense. Poder-se-ia argumentar que essas unidades de baixa frequência correspondem a casos de disponibilidade léxica, ou seja, casos em que

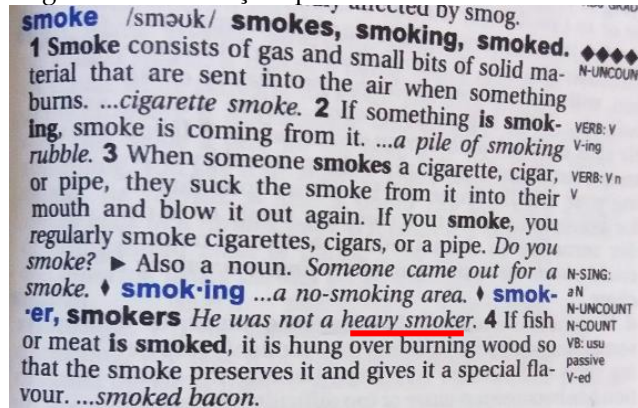
uma unidade é quantitativamente pouco relevante em termos estatísticos, mas, ainda assim, julgada pelo compilador da obra como potencialmente relevante para o consulente. Uma análise dessas palavras de baixa frequência não permitiu, porém, encontrar alguma razão plausível para pensar que se trata de unidades léxicas lematizadas pelo princípio da disponibilidade léxica.

Contudo, o exemplo do inglês não é tão grave quanto o de outras tradições lexicográficas, como a de língua espanhola, na qual é notório o fornecimento de informações não discretas. Um exemplo disto consiste na inclusão de palavras-entrada que não possuem respaldo em corpus (*i.e.*, *ghost-words*), como no caso do DiSal – *Diccionario Salamanca de la lengua española* (CUADRADO, 2006). Conforme Borba (2017a), de 158 verbetes analisados, 15 correspondiam a *ghost-words*. Ao incluir tais unidades léxicas, o dicionário não reflete a língua de maneira adequada¹³.

Já as informações capazes de auxiliar efetivamente o estudante visado pelo dicionário são consideradas discriminantes. Garantir apenas a natureza discreta das informações fornecidas não significa que o consulente vá obter um ganho real para satisfazer suas necessidades. Um exemplo disto pode ser encontrado em s.v. *smoke*, no CCLD (2003), reproduzido através da Figura 3 (ver grifo nosso sob *heavy smoker*):

¹³ Cf. 3.3 para mais detalhes.

Figura 3 – informações pouco efetivas ao consulente.

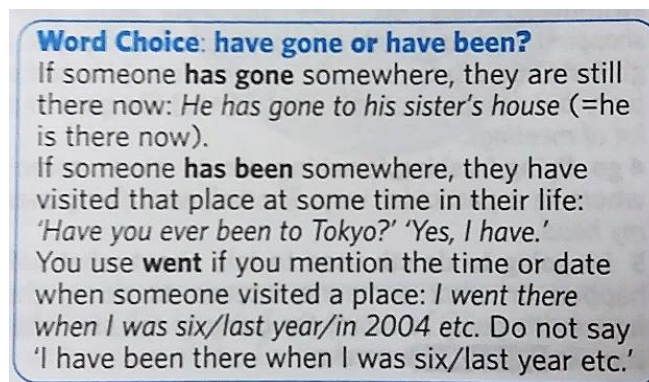


Fonte: CCLD (2003, s.v. *smoke*. Grifo nosso).

O dicionário é fiel à língua em uso ao incluir a combinação *heavy smoker* no exemplo. Contudo, não foi efetivo, porque não adverte o consulente de que se trata de uma combinação.

A importância de se incluir informações discriminantes também pode ser atestada em casos como *go* no *Longman Active Study Dictionary* (LAS, 2010, s.v.). Ao final do verbete, o dicionário comenta as diferenças entre os usos do presente perfeito *has gone* e *has been* e o pretérito simples *went*:

Figura 4 – informações de auxílio efetivo ao consulente em dicionários de aprendizes.



Fonte: LAS (2010, s.v. *go*)

Em suma, muito acertadamente, Ranalli e Nurmukhamedov (2012, p. 4) salientam que a maneira como as informações são apresentadas ao consulente nos dicionários de aprendizes carece de um viés mais didático. Em outras palavras, e como

já foi mencionado, o esforço por oferecer informações discretas pode acabar produzindo poucos resultados práticos se tais informações não forem também discriminantes.

As descobertas sobre a língua que o emprego de *corpora* possibilita contribuiriam ainda mais no que concerne à explicitação de informações discretas e discriminantes. Entretanto, esse aproveitamento parece, por vezes, restringir-se à comprovação da frequência de uso das unidades léxicas incluídas nas obras. Além de informações relativas à frequência, há outras aplicações da área com potencial de aproveitamento para a lexicografia, embora ainda pouco investigadas.

Primeiramente, o problema da indicação de combinatórias léxicas (Figura 3) faz-nos questionar se não seria possível tirar ainda mais proveito dos *corpora* disponíveis, tanto para o estudo desse fenômeno lexicológico como para o seu tratamento no âmbito lexicográfico. Vasiljevic (2014, p. 50) assinala, por exemplo, que casos de combinatórias léxicas como as colocações são “um dos maiores desafios para os estudantes de segunda língua, mesmo em estágios avançados de proficiência linguística”¹⁴ (Tradução nossa). Em vista disto, um dicionário que se dirija a estudantes de uma língua estrangeira deveria ser capaz de fornecer informações discretas e discriminantes acerca de combinatórias léxicas.

Em segundo lugar, o processo de inclusão de informações capazes de oferecer ao consulente um auxílio efetivo poderia considerar ferramentas como os *corpora* de aprendizes, relacionados à área da análise de erros. Conforme comentado já, normalmente estabelece-se um perfil genérico de consulente para os dicionários de aprendizes. Contudo, ao delimitar o perfil de usuário a estudantes que possuam uma determinada língua materna, poder-se-ia fornecer informações que fossem realmente discriminantes para esses estudantes por meio da análise dos erros presentes em

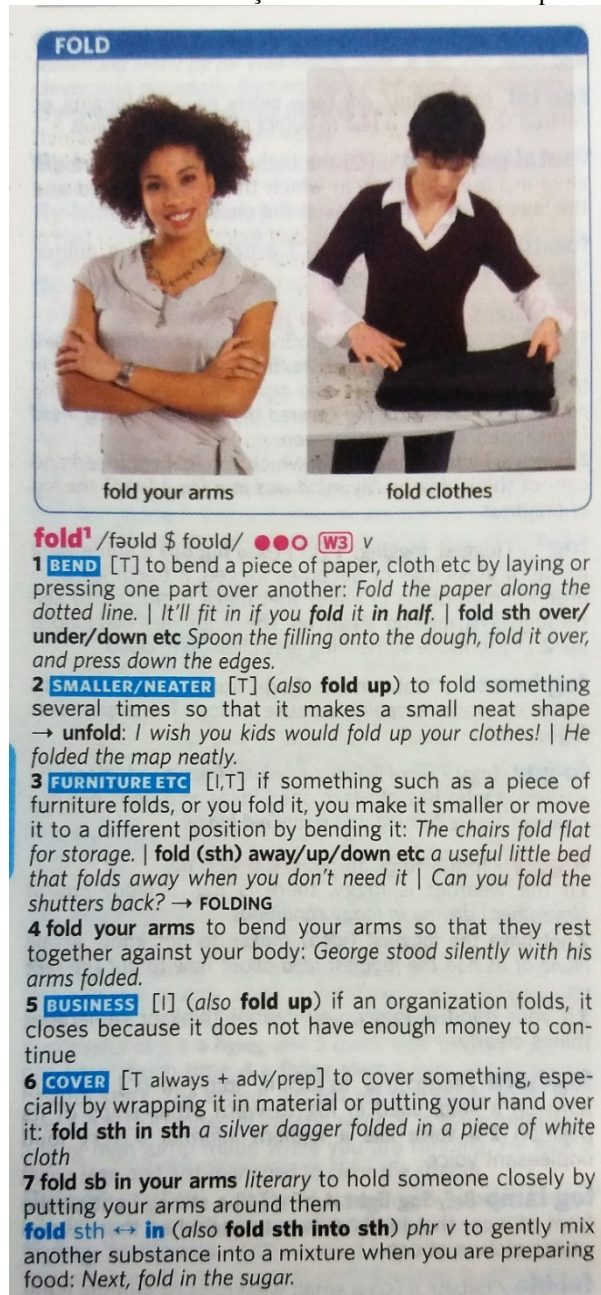
¹⁴ Cf. o trecho original: “one of the biggest challenges for second language learners even at very advanced stages of language proficiency”.

corpora de aprendizes. Dois exemplos de *corpora* compilados por produções de estudantes de língua estrangeira são o *International Corpus of Learner English* (ICLE, 2009) e o *Corpus de aprendices de español* (PARODI, 2015).

Em terceiro lugar, os dicionários de aprendizes de inglês citados até o momento empregam ilustrações em alguns dos seus verbetes. Entretanto, não há como saber, por exemplo, se foi considerado algum parâmetro de seleção de imagens; tampouco há estudos sobre o particular. Por um lado, um tipo de ferramenta que poderia auxiliar no estabelecimento desses parâmetros são os *corpora* de imagens, como o *Bank of Standardized Stimuli* (BOSS, 2010). Por outro lado, há duas perspectivas gerais acerca da percepção visual – *bottom-up* e *top-down* – e uma perspectiva acerca da percepção específica de objetos (STERNBERG R. J.; STERNBERG K., 2012, p. 85-86). Dessas três perspectivas, *top-down* é especialmente útil no tratamento de gravuras, particularmente no caso dos *realia*¹⁵ e nos casos em que uma língua distribui as designações de maneira diferente em relação à outra. Por exemplo: no tocante ao verbo *to fold*, nem para tudo o que é “dobrável” em inglês se pode empregar o equivalente *dobrar* do português. Uma gravura como a apresentada em LDOCE (2014, s.v. *fold*) é especialmente elucidativa para os casos de anisomorfismo linguístico.

¹⁵ Segundo Bugueño Miranda (2010, p. 67), *realia* são “manifestações extremas e absolutamente particulares da forma de organizar a realidade que uma língua possui”.

Figura 5 – Uso de ilustrações em dicionários de aprendizes.



Fonte: LDOCE (2014, s.v. *fold*).

3.3 Duas outras manifestações do presente: Lexicografia de Aprendizes do alemão e do espanhol

Nas últimas décadas, foram compilados também dicionários de aprendizes do alemão e do espanhol. É notável a disparidade existente entre os representantes das duas tradições.

No caso da língua alemã, segundo Bugueño Miranda e Jardim (2010, p. 52), os dicionários de aprendizes surgiram somente no início dos anos 1990. Dentre os aspectos positivos destacados pelos autores estão a marcação de regências e de casos, de maneira a oferecer informações discriminantes para o consulente. Soma-se a isto o fornecimento de exemplos igualmente discriminantes, pois auxiliam na compreensão das definições e do comportamento sintático das unidades léxicas. Por outro lado, há também pontos negativos, como a escassa indicação de pronúncia das unidades léxicas e a apresentação pouco discriminante das colocações, pois os dicionários não as indicam de forma clara. Cabe observar ainda que a lexicografia alemã não emprega dados de *corpora* com a mesma extensão que a lexicografia de aprendizes de língua inglesa. A consequência prática dessa decisão é que a indicação de frequência se restringe às pouco mais de 4000 unidades léxicas mais frequentes na língua. Curiosamente, tal indicação de frequência não ocorre nas versões integrais dos dicionários, mas sim apenas nas suas versões reduzidas.

Já a situação da lexicografia de aprendizes de língua espanhola é a mais alarmante. Há somente três obras disponíveis e todas auxiliam o consulente apenas de maneira parcial. Trata-se do já mencionado Cuadrado (2006), do *Diccionario Señas para la enseñanza de la lengua española para brasileños* (FONTES, 2001) e do *Diccionario de la lengua española para estudiantes de español* (GONZÁLEZ, 2002). Alguns dos principais problemas apontados em Borba (2017a) dizem respeito à inclusão de unidades léxicas com pouco respaldo em *corpora*. No caso específico do DiSal (CUADRADO, 2006), soma-se a isto o fato de incluir *ghost-words*. Observa-se ainda uma discrepância do tratamento de formas canônicas e variantes em relação à sua frequência de uso. No DiLEs (2002), a remissão de *cachemira*, com 119 registros no *Corpus de Referencia del Español Actual Anotado* (CREAa, 2015) para *cachemir*, com 27 registros no mesmo corpus é um exemplo disto¹⁶.

¹⁶ Cf. Borba (2017a) para mais exemplos.

Outro aspecto deficitário nas três obras da lexicografia de aprendizes hispânica é o tratamento assistemático dos padrões de complementação sintática dos verbos. Muitas vezes, os exemplos são o único meio a partir do qual é possível (mas nem sempre) depreender o comportamento sintático das unidades léxicas. O tratamento atribuído às combinatórias léxicas, por sua vez, também é pouco satisfatório. Segundo Borba (2017b), paradoxalmente, a única alternativa que o estudante de espanhol como língua estrangeira tem para resolver estas e outras questões é consultar dicionários compilados para falantes nativos de espanhol.

4. Futuro / considerações finais

É notória a necessidade de aperfeiçoamento dos dicionários de aprendizes. Há certos pontos de convergência da Lexicografia Pedagógica com outras áreas do conhecimento que poderiam ser melhor explorados em prol do fornecimento de informações discretas e discriminantes. Uma teoria da definição lexicográfica poderia auxiliar na redação de definições mais elucidativas. A Cognição, particularmente no que diz respeito à percepção de imagens, pode guiar na formulação de uma metodologia que permita um uso mais eficiente das ilustrações no dicionário.

Conforme Chi (2015, p. 181), a própria Lexicografia de Aprendizes é (em princípio) fruto da relação interdisciplinar entre lexicografia e ensino de línguas estrangeiras. Na mesma esteira, Andersen e Nielsen (2009, p. 356) defendem que, “como a lexicografia é, por natureza, interdisciplinar, desenvolve-se em muitos casos a partir da contribuição de um amplo número de disciplinas dentre as quais uma é a linguística”¹⁷ (Tradução nossa). O estabelecimento de relações interdisciplinares com outras áreas de conhecimento consiste, portanto, em uma demanda a ser suprida pelas pesquisas em metalexigrafia.

¹⁷ Cf. o trecho original: “since lexicography is in nature interdisciplinary, it thrives in many cases on input from a large number of disciplines of which one is linguistics”.

Referências

ANDERSEN, B.; NIELSEN, S. Ten Key Issues in Lexicography for the Future. *In*: BERGENHOLTZ, H.; NIELSEN, S.; TARP, S. (ed.). **Lexicography at a Crossroads**. Dictionaries and Encyclopedias today, lexicographical tools tomorrow. Bern: Peter Lang, 2009. p. 355-365. DOI <https://doi.org/10.1093/ijl/ecs009>

BORBA, L. C. Dicionários para aprendizes na lexicografia hispânica: Señas (2001), Salamanca (1996) e Espasa Calpe (2002). *In*: STURM, L.; VALÉRIO, P. (org.). **Anais do IX Seminário de Ensino de Línguas Estrangeiras (SELES), V Seminário de Ensino de Língua Materna (SELM), I Seminário Nacional Integrado da Área das Linguagens, Códigos e suas Tecnologias (SENI)**. V. 2. Passo Fundo: Editora Universidade de Passo Fundo, 2017a. p. 442-459. DOI <https://doi.org/10.29327/silicti2019.146891> Disponível em: http://selesselm.upf.br/images/pdf/anais_seles_e_selm_PARTE2.pdf. Acesso em: 30 abr. 2017.

BORBA, L. C. **Panorama da lexicografia hispânica**: subsídios para o professor de ELE. Saarbrücken: NEA - Novas Edições Acadêmicas, 2017b.

BRODEUR, M. **BOSS Bank of Standardized Stimuli**. Montreal: McGill University, 2010. Disponível em: <https://sites.google.com/site/bosstimuli/>. Acesso em: 01 fev. 2017.

BUGUEÑO MIRANDA, F. V. O dicionário bilíngue como problema linguístico e lexicográfico. *In*: HWANG, A. D.; NADIN, O. L. (Orgs.). **Linguagens em Interação III**: estudos do Léxico. Maringá: Clichetec, 2010. p. 65-91.

BUGUEÑO MIRANDA, F. V.; FARIAS, V. S. Proposta de um modelo de avaliação de dicionários escolares de Língua Portuguesa. *In*: XIV SEMINÁRIO NACIONAL DE LETRAS E LINGUÍSTICA E IV SEMINÁRIO INTERNACIONAL DE LETRAS E LINGUÍSTICA, 2013, Uberlândia. **Anais do SILEL**. Uberlândia: EDUFU, 2013. v. 3. p. 01-20. Disponível em: http://www.ileel.ufu.br/anaisdosilel/wp-content/uploads/2014/04/silel2013_1100.pdf. Acesso em: 10 jun. 2017.

BUGUEÑO MIRANDA, F. V.; JARDIM, C. R. Os learner's dictionaries do inglês e os Lernwörterbücher do alemão: uma simples transposição de nomes? **Contingentia**, Porto Alegre, v. 5, n. 1, p. 41-67, 2010. Disponível em: <http://seer.ufrgs.br/index.php/contingentia/issue/view/957>. Acesso em: 07 fev. 2017

CALD4. CAMBRIDGE. **Cambridge Advanced Learner's Dictionary**. Cambridge: Cambridge University Press, 2013.

CCAD. HAPERCOLLINS. **Collins Cobuild Advanced Dictionary**. Glasgow: HarperCollins, 2008.

CCLD. COLLINS. **Collins Cobuild Learner's Dictionary**. Concise edition in colour. 2nd edition. Glasgow: HarperCollins, 2003.

CHI, A. Researching Pedagogical Lexicography. In: JACKSON, H. (ed.). **The Bloomsbury Companion to Lexicography**. London: Bloomsbury Publishing, 2015. p. 165-187. DOI <https://doi.org/10.5040/9781472541871.ch-009>

CREAa. REAL ACADEMIA ESPAÑOLA. **Corpus de Referencia del Español Actual Anotado**. 2015. Disponível em: <http://www.rae.es/recursos/banco-de-datos/crea>. Acesso em: 30 nov. 2016.

CRYSTAL, D. **English as a Global Language**. 2nd edition. New York: Cambridge University Press, 2003.

GONZÁLEZ, C. (ed.). **Diccionario de la lengua española para estudiantes de español**. Espasa Calpe, 2002.

CUADRADO, J. Gutiérrez (dir.). **Diccionario Salamanca de la Lengua Española**. Madrid: Santillana, 2006.

FARIAS, V. S. *Whole-sentence definition versus definição por *genus proximum + differentiae specificae*: um contraste entre duas técnicas definitórias*. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 17, n. 1, p. 73-100, 2009. DOI <https://doi.org/10.17851/2237-2083.17.1.73-100> Disponível em: <http://periodicos.letras.ufmg.br/index.php/relin/issue/view/199>. Acesso em: 07 fev. 2017.

FONTES, M. SEÑAS: **Diccionario para la Enseñanza de la Lengua Española para Brasileños**/Universidad Alcalá de Henares. Tradução: Eduardo Brandão. São Paulo: Martins Fontes, 2001.

HALL, G. Method, Methods and Methodology: historical trends and current debates. In: HALL, G. **The Routledge Handbook of English Language Teaching**. London: Routledge, 2016, p. 209-223.

HARTMANN, R. R. K.; JAMES, G. **Dictionary of Lexicography**. Londres: Routledge, 2001.

ICLE. GRANGER, S. **International Corpus of Learner English**. Louvain-la-Neuve: Université Catholique de Louvain, 2009. Disponível em: <http://www.fltr.ucl.ac.be/fltr/germ/etan/cecl/Cecl-Projects/Icle/icle.htm>. Acesso em: 01 fev. 2017.

JACKSON, H. **Lexicography**. An Introduction. London: Routledge, 2002.

KLOTZ, M.; HERBST, T. **English Dictionaries**. A Linguistic Introduction. Berlin: Erich Schmidt Verlag, 2016.

LAS. PEARSON. **Longman Active Study Dictionary**. 5th edition. Harlow: Pearson Education, 2010.

LDOCE. PEARSON. **Longman Dictionary of Contemporary English**. 6th edition. Harlow: Pearson Education, 2014.

LEW, R. Dictionaries and Technology. In: CHAPELLE, C. (Ed.). **The Encyclopedia of Applied Linguistics**. New Jersey: Wiley-Blackwell, 2013. p. 1-08.

MEDAL. RUNDELL, M. (ed.). **Macmillan English Dictionary for Advanced Learners**. 2nd edition. London: Macmillan, 2007.

MLS. GLÜCK, H. (hrsg.). **Metzler Lexikon Sprache**. Stuttgart: J.B. Metzler, 2010.

MWALE. MERRIAM-WEBSTER. **Merriam-Webster's Advanced Learner's English Dictionary**. Springfield: Merriam-Webster, 2008. DOI <https://doi.org/10.2307/1223820>

OALD8. TURNBULL, J. (ed.). **Oxford Advanced Learner's Dictionary**. 8th edition. Oxford: Oxford University Press, 2010.

PARODI, G. Corpus de aprendices de español (CAES). **Journal of Spanish Language Teaching**, v. 2, n. 2, p. 194-200, 2015. DOI <https://doi.org/10.1080/23247797.2015.1084685>

RANALLI, J.; NURMUKHAMEDOV, U. Learner Dictionaries. In: CHAPELLE, C. (Ed.). **The Encyclopedia of Applied Linguistics**. New Jersey: Wiley-Blackwell, 2012. p. 1-06. DOI <https://doi.org/10.1002/9781405198431.wbeal1418>

RUNDELL, M. Recent Trends in English Pedagogical Lexicography. In: FONTENELLE, T. **Practical Lexicography: a reader**. New York: Oxford University Press, 2008.

SCHLATTER, M.; GARCEZ, P. M.; SCARAMUCCI, M. O papel da interação na pesquisa sobre aquisição e uso de língua estrangeira: implicações para o ensino e para a avaliação. **Letras de Hoje**, Porto Alegre, g. 39, n. 3, p. 345-378, 2004. DOI <https://doi.org/10.15448/1984-7726.2017.1.24925>

STERNBERG, R. J.; STERNBERG, K. **Cognitive Psychology**. 6th edition. Belmont: Wadsworth, 2012.

VASILJEVIC, Z. Teaching Collocations in a Second Language: Why, What and How? **ELTA Journal**, Belgrado, v. 2, n. 2, p. 48-73, 2014. Disponível em: <http://eltajournal.org.rs/elta-journal-volume-2-no-2-dec-2014/>. Acesso em: 15 fev. 2017.

VERLINDE, S.; BINON, J. Pedagogical Lexicography Revisited. In: BERGENHOLTZ, H.; NIELSEN, S.; TARP, S. (ed.). **Lexicography at a Crossroads**. Dictionaries and Encyclopedias today, lexicographical tools tomorrow. Bern: Peter Lang, 2009. p. 69-89. DOI <https://doi.org/10.1093/ijl/ecs009>

WB. COLLINS. **WordBanks**. [S.l.]: Collins, 2005. Disponível em: <https://wordbanks.harpercollins.co.uk/>. Acesso em: 02 fev. 2017.

WEC. MACMILLAN. **World English Corpus**. About the corpus used in Macmillan Dictionary. [S.l.]: Macmillan, [s.d.]. Disponível em: <http://www.macmillandictionary.com/corpus.html>. Acesso em: 02 fev. 2017.

Artigo recebido em: 20.04.2018

Artigo aprovado em: 05.11.2018