

A IMPORTÂNCIA DAS COLOCAÇÕES NO ENSINO DE VOCABULÁRIO DE INGLÊS COMO LÍNGUA ESTRANGEIRA

Leila Maria Taveira Monteiro é doutora em Ensino de Língua Inglesa pela Universidade Federal Fluminense. Trabalha no Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro como professora de inglês para fins específicos. Tem como áreas de interesse aquisição do léxico em língua estrangeira, multimodalidade e ensino de língua estrangeira e ensino instrumental de inglês.

E-mail: leilamonteiro@hotmail.com

Resumo

O conceito de colocabilidade entre palavras e sua relevância no processo de aquisição de língua estrangeira se destaca na literatura contemporânea. Este trabalho enfoca o tema e apresenta as mais recentes abordagens sobre a questão lexical com foco no fenômeno das colocações.

Abstract

The concept on collocation between words and its relevance in the acquisition process of a foreign language has been highlighted in current literature. This article addresses the topic and presents the newest approaches to the lexical issue, focusing the phenomenon of collocations.

*Lutar com palavras é a luta mais vã.
Entanto lutamos mal rompe a manhã.
São muitas, eu pouco.
Palavras, palavras, (digo exasperado)
Se me desafias, aceito o combate.
(Drummond, 1983)*

1) Um fenômeno e sua conceituação

Todas as línguas, de um modo geral, possuem frases e construções mais ou menos fixas, não necessariamente idiomáticas. Em inglês, essa constatação é particularmente verdadeira, talvez pelo fato de que a maioria das palavras abarca diversos significados, alguns bastante diferentes, outros, apenas tenuamente distintos. Na literatura, esses grupos de palavras ou coocorrências recebem variadas denominações, entre essas, *colocações*, termo adotado para o foco deste artigo.

A investigação da colocabilidade entre palavras pode se desenvolver a partir de uma abordagem *estatística*, ou *fraseológica*. A primeira aborda a distribuição numérica da coocorrência de itens lexicais, em um determinado contexto, como em Sinclair (2003). Esse tipo de levantamento está, preferencialmente, relacionado à Linguística de Corpus, disciplina que se tornou fortemente associada ao uso de computadores para fins de coleta e análise de dados linguísticos. A segunda vertente, a fraseologia, tem seu foco voltado para o estudo dos diferentes tipos de combinação entre palavras, sua tipificação e convencionalidade, como em Cowie (1998) e Howarth (1998).

Segundo Wray (2002, p.7), as primeiras observações registradas, a respeito de um possível grau de constância entre palavras, remonta a meados do século XIX, quando John Hughlings Jackson (1835-1911), neurologista britânico, pesquisou pacientes afásicos e sua aparente capacidade para produzir segmentos linguísticos fixos, como textos religiosos, expressões de cumprimento e rimas, habilidade que não demonstravam na produção de unidades novas da fala.



O conceito de colocabilidade e o termo *colocação* foram introduzidos pelo linguista inglês John Rupert Firth (1890-1960), em seu trabalho *Modes of Meaning* de 1957, no qual defende que parte do significado de uma palavra está ligado a outras que a acompanham (*apud*, Tagnin, 1999, p.26). Embora o pesquisador tenha influenciado sua geração e cunhado a frase: “Você deve julgar uma palavra por sua companhia.”,¹ sua contribuição, nessa área, foi negligenciada por várias décadas. Esse aparente desinteresse pode ter se originado da ausência de uma definição clara, por parte de Firth, que se limitou a ilustrar o conceito com exemplos, mas, do mesmo modo, pode se dever às restrições práticas, à época, para a realização de estudos textuais (Berry-Roghe, 1973, p. 103). Desde então, diversos pesquisadores têm buscado conceituar o fenômeno da coocorrência, à luz de modernas teorias da Linguística Aplicada, e o estudo ganhou mais espaço com a crescente facilidade que os recursos de informática trouxeram para a busca de palavras em seu ambiente.

Halliday, que foi aluno de Firth, recolocou a questão em pauta (1966) e, posteriormente, seus estudos em sociosemiótica (1978) tiveram um grande impacto, para a noção de *como* e *por que* realizamos determinadas escolhas discursivas. Para Halliday, o fenômeno da colocação é um elemento crucial para a análise do léxico, uma vez que traz informações probabilísticas importantes para o linguista, cujo interesse _ segundo o autor _ não está apenas no que sabemos sobre a língua, mas, igualmente, no que fazemos *com* e *através* dela.

O enfoque no aspecto da combinação entre palavras foi, do mesmo modo, privilegiado por Saussure (1916-1966), que se referiu ao sequenciamento de elementos que dispensam análise individual e passam a funcionar como unidades integradas de significação. À época, essa abordagem foi temporariamente abandonada, sob a influência do conceito de competência gramatical de Chomsky (1965), e retomada, quando novas teorias, baseadas na competência comunicativa de Hymes (1972), passaram a ser o mote das pesquisas na área.

No cenário atual, em oposição ao antigo enfoque descontextualizado, que deu origem às populares listas de vocabulário, advoga-se que o significado de uma palavra nova deve ser aprendido em seu contexto. A esse aspecto, alguns autores têm se referido como “ambiente/equilíbrio ecológico”, como em Sardinha (2000a, p. 51), para quem “[...] colocação é um tipo de padronização do léxico que diz respeito à associação entre palavras.” (*ibid*, p. 49).

Essa percepção de conjunto facilitaria seu reconhecimento e lembrança, quando há necessidade de recuperá-las na memória, já que “[...] palavras tendem a ter uma estrutura altamente colocacional” (Souza, 2004, p.20). É certo que uma palavra isolada não deve ser vista como uma unidade vazia de conteúdo, pois carrega em si um significado intrínseco, moldado por suas participações em textos anteriores. Entretanto, ao mesmo tempo em que transmite sentidos previamente adquiridos, recebe de um diferente contexto um renovado potencial de significação que emana do novo ambiente e nele se completa. Esse processo interativo é destacado em Leffa, (2000, p. 24):

[...] a palavra não só assume o significado imposto pelo texto, mas também determina seu significado. [...] Na verdade, o que se tem entre o texto e a palavra é um processo de interação baseado em algumas regras fundamentais, onde o texto, não necessariamente dá um significado à palavra, mas privilegia um de seus possíveis traços semânticos. (Leffa, 2000, p. 24)

Nesse continuado movimento de troca, o significado exato, de qualquer item lexical, é determinado pelo contexto em que se encontra, ou seja, pelas palavras que o cercam e com as quais combina. Para Leffa (*ibid*, p.17), a habilidade nessa delicada combinação é uma das competências mais difíceis de adquirir, principalmente nos aspectos produtivos da escrita e da fala, e um dos mais expressivos indicadores de proficiência por parte do aprendiz de língua estrangeira.

Por exemplo, um aluno que, inadvertidamente, use a expressão “*a great accident*”, considerando que “*great*” pode ser (entre outras opções) “grande, enorme, imenso”, para significar um “grande acidente” _ no sentido de grave, sério_ não forma um texto preciso: deveria usar “*dreadful, major, serious, terrible*”ⁱⁱ. Ou seja, “*great*” não se *coloca* com “*accident*” nesse contexto.

A ausência de critérios claros que permitam uma abordagem padronizada nas investigações tem resultado em uma ampla variedade terminológica, para rotular duas, ou mais palavras que são produzidas em sequência, sem (grande) esforço por parte do usuário.

Para Weinert (1995, p. 182), por exemplo, as diferentes nomenclaturas _fórmula, língua pré-fabricada, língua para pronto uso, pedaços, unidades multivocabulares, expressões fixas, ou semifixas, e outras (*formula, pre-fabricated language, ready-made language, chunks, multi-word units, fixed/semi-fixed units*) _ referem-se, essencialmente, ao mesmo fenômeno.

Já Schmitt e Underwood (2004) adotam a expressão “sequências formulaicas” (*formulaic sequences*) para rotular sequenciamentos que podem variar profundamente, de simples expressões para preencher o discurso (*fillers*), como “*kind of*”, ou exercer funções, “*I’m sorry*”, até provérbios, “*Barking dogs seldom bite*”, passando por colocações, “*make a mistake*”, e expressões idiomáticas, “*to pull someone’s leg*”.

Boers (2006, p.257) assume posição semelhante e nomeia, indistintamente, colocações e expressões idiomáticas, como “*formulaic sequences*”, sob o argumento de que ambas pertencem à categoria de “expressões multivocabulares padronizadas” (*standardized multiword expressions*) da língua (*ibid*, p. 246).

Sardinha (2000b, p. 50-1) opta por “*chunks*”, termo que traduz como “porções”. Para o autor, a denominação engloba de forma efetiva diversos padrões de colocabilidade léxico-gramaticais, além de ser uma nomenclatura amplamente utilizada e reconhecida na área de ensino/aprendizado de língua estrangeira.

Para Howarth (*op.cit.*, p. 25), a questão da colocabilidade se trata de um fenômeno que, examinado atentamente, revela-se muito mais complexo e diversificado do que, à primeira vista, pode se supor. O linguista argumenta que uma excessiva flexibilidade, no uso indistinto da terminologia, pode induzir a perigosos equívocos de conceituação e sugere que as expressões “*linguagem formulaica*”, ou “*fórmulas*”, não sejam usadas como termos guarda-chuva que englobem, em uma única categoria, todo tipo de combinações. O autor argumenta:

A sugestão de que ‘*linguagem formulaica*’, em oposição à língua gerada segundo regras, é uma categoria única, que abarca todas as características significativas de combinações de palavras, pode levar a uma impressão equivocada. Na falta de critérios formais para a identificação e categorização não está claro se ‘*linguagem formulaica*’ inclui ‘pedaços pré-fabricados’ tais como colocações e expressões idiomáticas.ⁱⁱⁱ (Howarth, 1998, p.25).

Assim como há divergências a respeito da nomenclatura mais apropriada, a natureza complexa e, ao mesmo tempo, fugidia, da combinalidade entre palavras dá origem a uma variedade de percepções teóricas sobre o fenômeno. Wray (*op.cit.*, p.9), por exemplo, acredita que nenhuma conceituação conseguiu, até o momento, descrever, de modo amplo e indiscutível, a essência da circunstância da coorrência, que define de forma abrangente como:

[...] uma sequência, contínua ou não, de palavras ou outros elementos, que é, ou aparenta ser, pré-fabricada: ou seja, armazenada e recobrada da memória em sua totalidade no momento em que é utilizada, ao invés de estar sujeita à produção, ou à análise pelo sistema gramatical da língua.^{iv} (Wray, 2002, p. 9)

Outros autores, entretanto, mantêm uma postura crítica a respeito dessa generalização, já que o grau de estabilidade entre as palavras pode ser percebido de forma distinta, por diferentes indivíduos. Essa diferença de percepção é visível especialmente se contrastarmos aprendizes de língua estrangeira que tendem a perceber unidades lexicais, palavra por palavra, com falantes nativos, para quem as fórmulas transmitem, mais facilmente, um todo de significado (Boers, *op.cit.*, p. 247).

A partir da análise das principais definições encontradas na literatura _ em uma tentativa de criar mecanismos que facilitassem a classificação desses itens vocabulares _ Partington (1998, p. 15-16), sugere sua categorização segundo três critérios, ou enfoques distintos:

i) Enfoque textual: nesse caso, o fenômeno é definido com base na linearidade da língua, pois decorre da existência de um item lexical junto a outro em um texto. Essa é a posição em Sinclair (1991, p.170) ^v: “Colocação é a ocorrência de duas ou mais palavras, distantes um pequeno espaço de texto umas das outras.”. Posteriormente, Hoey (2005, p.3) veio a criticar a definição como um conceito bastante limitado que, a seu ver, não reflete a abordagem feita pelo próprio Sinclair e propõe o termo “coocorrência lexical” (*lexical cooccurrence*), para se referir à ideia de proximidade física entre palavras.

ii) Enfoque psicológico, ou associativo: a definição se apoia na intuição linguística do falante, que, por meio de exposição, desenvolve a capacidade de escolher coorrências usuais na língua. Esse é o enfoque adotado por Aitchinson (2003, p.91) para quem: “Aprende-se o significado de uma palavra, provavelmente, ao se notar as palavras que vêm juntas.” ^{vi} O mesmo tratamento ocorre em Leech, (1974, p. 20) ^{vii}: “O sentido colocacional consiste das associações que uma palavra faz, por conta dos sentidos das outras palavras, que tendem a ocorrer em seu ambiente”.

(iii) Enfoque estatístico: essa abordagem é a preferida por aqueles que transitam na área da Linguística de Corpus, como em Moon (1998, p. 26): “O fenômeno da colocação denota coorrências que se repetem de forma frequente, ou são estatisticamente significativas, haja, ou não, laços semânticos especiais entre os itens colocacionais.” ^{viii} Igualmente em Hoey (1991, p. 6), para quem “Colocação tem sido o nome dado à relação que um item lexical tem com itens que aparecem com maior probabilidade do que o mero acaso em seu contexto (textual)” ^{ix}. Posteriormente, Hoey (2005, p. 4) avaliou o conceito unicamente estatístico como limitado e insatisfatório, por não

esclarecer sobre o fenômeno em si, ou o seu propósito, e passou a definir colocação como “[...] uma propriedade do léxico mental de um indivíduo”^x.

A variedade de critérios se origina da falta de regras que expliquem o mecanismo da colocabilidade. Benson (1989, p. 4) destaca a imprevisibilidade do fenômeno e defende que: “Colocações deveriam ser definidas não apenas como ‘combinações recorrentes de palavras’, mas como ‘combinações *arbitrárias* recorrentes de palavras.”^{xi} (itálico nosso). Na interpretação de Wanner (2004, p. 98), uma colocação é uma combinação de itens lexicais, de natureza binária e lexicalmente restrita, não regida por regras universais de restrição semântica. O autor ilustra a natureza idiossincrática do conceito, com a colocação inglesa usual para significar um chá forte, com muita cafeína: “*strong tea*” e não “**powerful tea*”, embora ambos os adjetivos (*strong/powerful*) possam ser traduzidos como “forte”.

Tagnin (1989, p. 82), para quem colocabilidade é “o aspecto que caracteriza a forma peculiar de expressão numa dada língua”, combina os critérios citados por Partington (*op.cit.*) e define colocações como “uma combinação lexical recorrente, não idiomática, coesa, cujos constituintes são contextualmente restritos e de coocorrência arbitrária.” (Tagnin, 1999, p. 41). Em obra posterior (2002, p.193-4), a autora distingue entre fórmulas ou colocações, as quais define como “palavras que coocorrem em frequência maior do que caso se tratasse de uma combinação aleatória” e expressões idiomáticas, que conceitua como “[...] expressão cujo significado total não corresponde à soma dos significados individuais de seus componentes.”.

Essa distinção é, igualmente, defendida por Bahns (1993, p. 57), que avalia colocações como palavras que refletem seu significado intrínseco, frequentemente utilizadas juntas, de forma natural, com diferentes graus de ligação e peso psicológico, como, por exemplo, “*blow a trumpet*” (tocar trompete). Já, expressões idiomáticas são conceituadas como relativamente imutáveis, cuja tradução não reflete os significados distintos de seus componentes, como a expressão “*to be on cloud nine*”, que significa “estar muito feliz”, e nada tem a ver com a tradução literal, algo como: “*estar na nuvem nove”.

Em contrapartida, o pesquisador defende a necessidade de um sistema de categorização após o estabelecimento de características distintas. Para tal, acredita ser preciso uma análise criteriosa, que considere a existência de tipos fraseológicos diversos. Essa categorização tipificaria os itens lexicais combinados segundo características intrínsecas à sua própria natureza, considerando-se a observação de diferenças entre a produção de nativos e não nativos.

Apesar da falta de parâmetros consensuais que facilitem o trabalho de educadores e professores, a observação desse tipo de padrões é, sem dúvida, de grande importância no ensino de língua estrangeira, pois sua utilização pode certamente concorrer para a sensação de naturalidade na fala ou na escrita do aprendiz.

Encerrando o debate sobre o tema da conceituação, acreditamos que, a despeito das diferentes possíveis abordagens, seja correto afirmar que o fenômeno da colocabilidade se refere à maneira como as palavras se combinam em uma língua, de forma recorrente e arbitrária. Quando produzida de modo apropriado, essa delicada combinação gera um texto que soa natural, ou seja, não causa estranheza ao falante nativo.

2) Um fenômeno e sua categorização

O uso da tecnologia aprofundou a discussão acerca do melhor caminho a se adotar para a análise e a classificação de coocorrências. Para Sinclair (2003), a falta de um critério consensual deve ser solucionada, através do levantamento de *corpora*, baseado em medidas estatísticas, que espelhem a frequência de uso, processo que, como já mencionado, tem sido imensamente facilitado pela popularização do computador.

Howarth (*op.cit.*, p.26) reconhece a contribuição da informática, mas faz importantes restrições a uma abordagem meramente quantitativa que, a seu ver, não considera as sutis gradações de significado e os aspectos psicológicos envolvidos na questão da colocabilidade. O pesquisador argumenta que esse tipo de análise não leva em conta o fator da competência e destaca que processos que se baseiam unicamente na frequência correm o risco de apontar uma relevância, equivocada, de combinações altamente transparentes (que exemplifica com *have children*), frequentes, apenas, em função do tópico específico da amostra analisada.

Como solução, o autor defende a necessidade de um sistema de categorização após o estabelecimento de propriedades distintas. Para tal, acredita ser preciso uma análise criteriosa, que considere a existência de tipos fraseológicos diversos. Essa categorização tipificaria os itens lexicais combinados segundo características intrínsecas à sua própria natureza, considerando-se a observação de diferenças entre a produção de nativos e não nativos. Em decorrência, Howarth (*ibid.*, p.27) cria uma proposta de classificação hierárquica de segmentos fraseológicos que leva em conta a forma interna das combinações, sua função externa, especificidade semântica, restrições sintáticas e o maior ou menor grau de comutabilidade dos itens lexicais envolvidos. Dessa forma, os elementos se colocam sobre um contínuo que se estende de combinações livres a idiomas puros, passando por combinações restritas e idiomas figurativos.

Segundo a conceituação sugerida pelo autor (*ibid.* p.28), as *combinações livres*, localizadas em um dos extremos do contínuo, consistem de elementos usados em seu sentido literal, que aceitam livre substituição. Já as *restritas* apresentam um dos componentes usados em sentido figurativo e têm uso mais limitado, como o nome sugere. *Idiomas figurativos* têm um sentido metafórico e uma interpretação literal, ao passo que *idiomas puros* se encontram no extremo oposto das combinações livres. Esses são os mais opacos, pois não se pode deduzir seu significado, considerando-se o sentido isolado de seus componentes, os quais não aceitam permuta.

Embora defenda a necessidade de uma tipificação, Howarth reconhece que dificuldades importantes estão envolvidas na aplicação das conceituações propostas. Não se pode delimitar, com segurança, a divisão entre uma e outra categoria, já que o critério de comunicabilidade é significativamente sujeito a interpretações pessoais e componentes psicológicos.

O posicionamento de uma colocação pode, ainda, alterar-se, adquirindo um maior, ou menor grau de convencionalidade. Essa transformação ocorre, quando, ao ser reiteradas vezes utilizada por uma comunidade linguística, a expressão se consagra pelo uso, ou o movimento contrário pode ocorrer, quando a combinação perde força e deixa de fazer parte do elenco de itens convencionais da língua. Essa instabilidade é ilustrada em Louro (2001, p. 24), pela expressão “vaca louca” (*mad cow*):

Essa colocação foi cunhada com a intenção de nomear popularmente a doença cujo nome científico em português é Encefalopatia Espongiforme Bovina [...]. A comunidade em geral adotou esse nome popular, e pode-se dizer que *mad cow* e *vaca louca* já são colocações fixas, consagradas ou convencionais. (Louro, 2001, p. 24)

Em um movimento oposto ao caso exemplificado por Louro, igualmente em português, a expressão “o samba do crioulo doido” foi, durante um período, muito usada para indicar confusão, falta de clareza. Hoje, já não é mais reconhecida pelas novas gerações, pois sua origem – um samba popular da década de 60 – perdeu-se na memória do povo ^{xii}. O mesmo ocorreu com a expressão inglesa, “*to rain cats and dogs*”, que surgiu na era medieval, quando gatos e cachorros se empoleiravam nas vigas dos tetos e de lá pulavam, quando chovia muito. Hoje, é mais usual que se diga “*It’s pouring!*”, para se fazer menção à chuva forte. Ou seja, o todo de significado pode, eventualmente, tornar-se mais ou menos claro, em função da própria natureza dinâmica da linguagem verbal.

A despeito da reconhecida dificuldade em se estabelecer um critério consistente de tipificação, a proposta de um construto evolutivo reflete a busca por um modelo de análise, aplicável aos diferentes níveis do fenômeno colocacional, e tem sido sugerida por outros autores. Benson, Benson e Ilson (1997) distinguem três diferentes graus de coesão: no menos acentuado estão as *combinações livres*, quando as palavras podem ser encontradas separadamente, combinadas com diversas outras, e não se constituem por si só em itens lexicais. Nesse caso, a gramaticalidade é a única restrição à diversidade de combinações. Por exemplo, na expressão “*buy + a car*” (comprar um carro), é possível usar o verbo “*buy*” associado a diversos outros complementos (*a house, a dress, a book, presents*), assim como muitos outros verbos (*sell, have, look at, love*) podem se unir ao substantivo “*car*”.

No extremo oposto, encontram-se as *expressões idiomáticas*, totalmente fixas, e que não espelham o significado das palavras individualmente, como por exemplo, “*to pull one’s leg*”, que literalmente significaria “puxar a perna de alguém”, mas cujo sentido real é “implicar com alguém”.

Entre esses extremos se situam as *colocações* propriamente ditas, os itens que acreditamos apresentar maior dificuldade para o aprendiz, por não serem totalmente livres nem, tampouco, completamente fixas. São sequências que funcionam segundo uma relação de parceria que permite a combinação com outras palavras. É comum que os alunos se ressentam, por exemplo, da falta de regras que regulem o uso dos verbos “*do*” e “*make*”, ambos traduzidos como “fazer”. Sem nenhum critério lógico aparente para o aprendiz, diz-se “*make a visit*”, mas “*do a favour*”. Para isso, embora possa se tentar certa padronização de conceitos, não há regras a seguir, é preciso que o aluno memorize, praticamente, caso a caso. As expressões idiomáticas equivalem aos *idiomas puros* sugeridos em Howarth (*op.cit.*), por apresentarem um alto grau de opacidade no seu sentido: “*kick the bucket*” não quer dizer “chutar o balde”, como a tradução literal poderia levar a crer, mas “abandonar, desistir, deixar de lado possíveis consequências ruins”, e “*take up a course*” não significa “apanhar, pegar um curso”, mas “frequentá-lo”.

Sinclair (1991, p.109-10) qualifica essa gradação com os termos, *princípio da livre escolha* e *princípio idiomático*. O primeiro determina que a parceria entre as palavras ocorre a partir de uma possibilidade quase ilimitada de escolhas por parte do usuário, restritas apenas pelas regras da gramaticalidade da língua. O segundo diz respeito à construção textual, iniciando-se pelo uso de unidades idiomáticas pré-fabricadas que, embora pareçam analisáveis em elementos distintos, constituem um todo de significado. Para o autor, ambos os princípios se articulam e se alternam, de forma dinâmica, no processo de construção de um texto coerente e coeso.

Lewis (2000, p. 8) propõe uma abordagem semelhante de gradação e classifica o fenômeno de coocorrência segundo um espectro que abrange três categorias distintas:

colocações livres, colocações relativamente fixas, que permitem um número limitado de combinações e as *totalmente fixas*, unidades lexicais consagradas e que, por não permitirem alteração em seus sequenciamentos, configuram-se como expressões idiomáticas. Para o autor, unidades muito fortes, como “*drug addict*” (viciado em drogas), podem ser, até mesmo, consideradas palavras compostas (Lewis, *ibid*, p. 31).

Segundo o modelo adotado em Tagnin (2002, p.195 e 1999, p.41-2) e Wanner (2004, p.98-9), as colocações podem ser ainda rotuladas em função da classe gramatical de seus elementos: para os autores, as sequências nessa categoria são formadas por dois itens lexicais denominados de *base* – geralmente um substantivo, é aquele que determina, tem caráter autônomo e possui maior valor semântico – e *colocado*, a unidade determinada, que acrescenta características semânticas ao valor da base e cuja categoria gramatical dá nome à colocação. Dessa forma, é possível montar os seguintes grupos:

- a- substantivo+substantivo= colocação nominal (*fire alarm*)
- b- adjetivo+substantivo= colocação adjetiva (*historic event*)
- c- verbo+substantivo= colocação verbal (*catch a cold*)
- d- advérbio+ adjetivo / verbo + advérbio= colocação adverbial (*wickedly funny / blankly gaze*)
- e- adjetivo + preposição = colocação preposicional (*similar to*)

Para os autores, o processo de categorização facilita sua percepção e posterior emprego por parte do educando, além de concorrer para que a construção do significado se faça a partir de blocos, ao invés de palavras isoladas, procedimento que, frequentemente, concorre para graves equívocos.

A despeito das diferentes nomenclaturas e critérios de classificação propostos, a noção de coocorrência entre unidades lexicais e o termo *colocação* ganharam, definitivamente, visibilidade nos estudos da área. O conceito de combinalidade entre palavras e sua importância para o aprendiz de língua estrangeira tem se refletido de modo decisivo na criação de materiais e na sugestão de abordagens instrucionais, resultando em uma crescente percepção do tema por parte de educadores e professores da área.

3) Um fenômeno e a sala de aula de língua inglesa

Mais recentemente, o significativo consenso sobre a conveniência de tratamentos instrucionais para o ensino e o aprendizado de vocabulário, a partir da noção de colocabilidade, assumiu uma terminologia particular e uma nova dimensão com os estudos de Hill (1999) e Celce-Murcia (2001). Os autores cunharam, respectivamente, as expressões *competência colocacional* e *competência formulaica*, para se referir à necessidade de uma proficiência no tocante à combinação de palavras. No cenário específico de aquisição de língua estrangeira, a necessidade desse tipo de competência é destacada por Sardinha (2000b, p. 51), para quem: “A observação de padrões é tida como de suma importância no ensino de língua estrangeira, pois a sensação de ‘naturalidade’ na fala ou na escrita depende em grande parte do emprego de padrões.”

Uma possível relação entre competência colocacional e fluência é, igualmente, vislumbrada por outros autores: Louro (*op. cit.*, p. 61), por exemplo, acredita que “[...] conhecer e usar as colocações resulta em fluência na comunicação.” e Hill (*op.cit.*, p.5) defende que a naturalidade e a velocidade da fala de falantes nativos se devem, em

grande parte, ao conhecimento de porções fixas da língua: “quando não conhecem as colocações, que expressam, exatamente, o que querem dizer, os alunos criam enunciados mais longos que aumentam a probabilidade de novos erros.”^{xiii}.

Para Porto (1998), as falhas da abordagem estrutural, que não encontrou um meio de permitir a utilização criativa das regras gramaticais, e as limitações da abordagem comunicativa, que disponibiliza um vasto repertório sem tornar o aprendiz gramaticalmente competente, podem ser solucionadas por um equilíbrio a ser alcançado com o uso de sequências formulaicas, termo que aqui utilizamos, de modo abrangente, como em Schmitt e Underwood (2004), e Boers (*op. cit.*). Para Porto (*op. cit.*), esses “blocos” de sentido, incluindo aspectos estruturais, como voz passiva e orações condicionais, permitem ao aprendiz maior fluência, mesmo nos níveis iniciantes, contribuindo para a sensação prazerosa de competência linguística, ao facilitar a progressiva sofisticação de recursos fraseológicos. Pensar a língua, a partir desse enfoque, significa priorizar o papel das unidades lexicais, principalmente daquelas que acontecem em sequência, como é o caso das colocações.

Na visão de Michael Lewis (1993, 1997), que preconiza a percepção, a memorização e a produção da língua-alvo, através de unidades lexicais em blocos significativos, consagrados pelo uso, sem necessidade de sua decomposição, a habilidade para perceber a língua em blocos de significado é especialmente positiva no processo de aquisição:

Uma sensibilidade para o processo da fala em blocos pode ajudar o aprendiz de diversas maneiras: ao identificar unidades para a aprendizagem, ao relacionar o vocabulário à gramática, ao processar a língua que escutam, planejar e produzir a língua fluentemente eles mesmos”.^{xiv} (Lewis, 1997, p. 158)

Pensar a língua, a partir desse enfoque, significa priorizar o papel das unidades lexicais, principalmente daquelas que acontecem em sequência, como é o caso das colocações. Vale ressaltar que o pesquisador não inovou, ao propor que o vocabulário seja percebido, ensinado e adquirido em blocos de significado: ao apresentar sua proposta (*The Lexical Approach*, 1993), Lewis faz menção a vários autores que já haviam escrito sobre o assunto, entre eles Bahns (*op.cit.*), McCarthy (1990), Nattinger e DeCarrico (1992) e Pawley e Syder (1983). Seu grande mérito, acreditamos, encontra-se na maneira como combinou teoria e prática, sistematizando e estruturando a dimensão do vocabulário no ensino de inglês como língua estrangeira.

Nesse contexto, subscrevemo-nos à proposta de Lewis (1997, 1993), por colocar o léxico em evidência e destacar a necessidade de que o aprendiz tenha sua atenção despertada para a idiomacidade da língua-alvo, aprendendo “a dizer como se diz”. Essa convencionalidade está profundamente relacionada à questão da coocorrência entre palavras, assunto que, devido à ampla diversidade de fontes na literatura, foi, sucintamente, aqui condensado.

4) Conclusão

Se é fato que “a língua não é só léxico, mas o léxico é o elemento que melhor a caracteriza e a distingue das outras.” (Leffa, 2000, p.19), o processo de aquisição de vocabulário se caracteriza pelo seu potencial de crescimento e aprofundamento continuado, seja em língua materna ou estrangeira. No segundo caso, o uso proficiente do repertório vocabular se torna ainda mais complexo e penoso. Curiosamente, apesar

de o senso comum, a tradição e a literatura valorizarem o conhecimento de um vocabulário rico e criativo, e a despeito de diversas abordagens instrucionais destacarem a importância do aprendizado de palavras novas, a questão lexical tem sido, em geral, estigmatizada como subproduto de tratamentos que priorizam a gramática, ou outras habilidades linguísticas.

Notadamente, aspectos relativos à estrutura do fenômeno linguístico são mais visíveis e suscetíveis à análise e à observação, do que fatores pertinentes à área de vocabulário. Essa peculiaridade explicaria, na concepção de Farghal e Obiedat (1995), a preferência da maioria dos linguistas por enfoques estruturais. Na concepção dos pesquisadores, contudo, a fronteira entre sintaxe e léxico é muito mais sutil e menos relevante do que, a princípio, se acredita. Em decorrência, os autores sugerem a existência de um contínuo dinâmico, que vai da linguagem pré-construída, moldada por padrões sintáticos, a um extremo oposto, onde o vocabulário leva a língua a modelos totalmente originais e o sentido potencial se realiza a partir do contexto.

Parece-nos que essa fluidez entre léxico e gramática estabelece uma posição conciliatória que reflete de modo realista a contínua interação entre ambas as dimensões linguísticas. Acreditamos, igualmente, que o enfoque em blocos significativos da língua, resultantes do fenômeno da colocabilidade entre palavras, oferece um caminho seguro e potencialmente efetivo na área de ensino de língua estrangeira por proporcionar uma ponte entre o domínio gramatical e a esfera do léxico.

Se, historicamente, no que tange a procedimentos instrucionais, a ênfase tem recaído em abordagens implícitas, especialmente por meio da prática de leitura, por um lado acreditamos que a validade desse enfoque pedagógico é inegável. Contudo, procedimentos instrucionais indiretos têm um limite para o que podem cobrir. Além disso, estudos têm apontado ser crucial que haja intenção consciente, por parte do aprendiz, de aprender palavras novas (Leffa, 2000; Laufer e Hill, 2000; Souza, 2004 e 2007; Souza e Braga, 2007), fator desconsiderado no tratamento indireto. Para Leffa (*op.cit.*, p. 35-6):

Para haver aprendizagem é necessário um esforço de atenção, não só para o significado da palavra, mas também para a sua forma. [...] A aprendizagem incidental não é inteiramente 'incidental' e, para ser bem sucedida, depende de vários fatores, nem sempre presentes nas tarefas executadas pelos aprendizes. (Leffa, 2000, p. 35-6)

Em consequência, avaliamos, que uma abordagem indireta para o ensino de vocabulário tem se provado insatisfatória, na medida em que há limites, para o que se pode esperar de técnicas fundamentadas em aprendizado incidental, particularmente no que diz respeito a colocações e expressões idiomáticas. Embora reconheçamos a validade de leituras extensivas para o aprimoramento do vocabulário, esse tratamento apresenta importantes restrições quanto ao seu efeito, dentre elas a dificuldade em se aferir, adequadamente, os resultados do processo. Vale igualmente considerar que, diferentemente do que ocorre no caso de língua materna, a aquisição de língua estrangeira se inicia com a apresentação direta de palavras novas e “[...] é só mais tarde, quando o aluno já possui um vocabulário de cerca de 3.000 a 5.000 palavras, que o processo de aprendizagem incidental tem início.” Leffa (*op.cit.*, p. 35).

Dessa forma, advogamos um tratamento intencional, continuado e sistematizado, para o ensino e o aprendizado do vocabulário, com foco especial no aspecto da coocorrência entre palavras, objetivando levar nossos alunos ao desenvolvimento de

uma proficiência linguística caracterizada pela naturalidade e pela efetividade no processo comunicativo da língua alvo.

ⁱ “You shall judge a word by the company it keeps”. Cf. FIRTH, John Rupert. *Modes of Meaning: Papers in Linguistics 1934-1951*. London: Oxford University Press. 1957. Pp. 190-215.

ⁱⁱ Sugestões em: **OXFORD Collocations Dictionary for Students of English**. New York: Oxford University Press. 2002. P.3.

ⁱⁱⁱ “It may give a misleading impression to suggest that ‘formulaic language’ is a single category, to be contrasted with language generated by rule, and encompassing all significant features of word combinations. In the absence of formal criteria for identification and categorization it is not clear whether ‘formulaic language’ includes such ‘prefabricated chunks’ as collocations and idioms.” (Howarth, 1998, p. 25)

^{iv} “[...] a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar.” (Wray, 2002, p. 9)

^v “Collocation is the occurrence of two or more words within a short space of each other in a text.” (Sinclair, 1991, p.170).

^{vi} “Word-meaning is probably learned by noting the words which come alongside”. (Aitchinson, 2003, p. 91)

^{vii} “Collocational meaning consists of the associations a word acquires on account of the meanings of words which tend to occur in its environment.” (Leech, 1974, p. 20).

^{viii} “Collocation typically denotes frequently repeated or statically significant co-occurrences, whether or not there are any special semantic bonds between collocating items.” Moon (1998, p. 26)

^{ix} “Collocation has been the name given to the relationship a lexical item has with items that appear with greater than random possibility in its (textual) context.” (Hoey, 1991, p. 6-7).

^x “[...] a property of the individual’s mental lexicon.” Hoey (2005, p. 4)

^{xi} “Collocations should be defined not just as ‘recurrent word combinations’, but as ‘arbitrary recurrent word combinations’” (Benson, 1989, p. 4)

^{xii} “O Samba do Crioulo Doido”. Letra de Sérgio Porto, conhecido como Stanislaw Ponte Preta, e gravado pelo Quarteto em Cy em 1968. Exemplo a critério da autora.

^{xiii} “When students do not know the collocations which express precisely what they want to say they create longer utterances which increase the likelihood of further errors.” (Hill, 1999, p. 5).

^{xiv} “Sensitivity to chunking in spoken discourse helps learners in many ways: identifying units for learning, relating vocabulary to grammar, processing language they are listening to, planning and producing language fluently themselves.” (Lewis, 1997, p.158).

Referências bibliográficas

AITCHISON, Jean. **Words in the Mind: An Introduction to the Mental Lexicon**. Oxford: Blackwell. 3rd Ed. 2003. 332 p.

BAHNS, Jens. Lexical Collocations: a Contrastive View. **ELT Journal**, v. 47, n.1, p.56-63. 1993.

BENSON, Morton. The Structure of the Collocational Dictionary. **International Journal of Lexicography**, v. 2, n.1, p. 1–13.1989.

_____.; BENSON, Evelyn; ILSON, Robert. **The BBI Combinatory Dictionary of English: a guide to word combinations**. Philadelphia, PA: John Benjamins. Revised edition. 1997. 386 p.

BERRY-ROGHE, Godelieve. “The Computation of Collocations and their Relevance in Lexical Studies.”. In: AITKEN, A. J.; BAILEY, R. W.; HAMILTON-SMITH, N.

(Eds.). **The Computer and Literary Studies**. Edinburgh: Edinburgh University Press. Pp 103-112. 1973. 369 P.

BOERS, Frank, et al. Formulaic sequences and perceived oral proficiency: putting a Lexical Approach to the test. **Language Teaching Research**, v. 10, n. 3, p. 245-61. 2006.

CELCE-MURCIA, Marianne. **Teaching English as a Second or Foreign Language**. 2nd Ed. Boston: Heinle & Heinle. 2001. 567 p.

CHOMSKY, Noam. **Aspects of the theory of syntax**. Cambridge: MIT Press. 1965. 276 p.

COWIE, Anthony Paul. **Phraseology: Theory, Analysis and Applications**. Oxford: Oxford University Press. 1998. 245 p.

DRUMMOND, Carlos Andrade de. **Antologia Poética**. São Paulo: Editora Record. 1983.

FARGHAL, Mohammed; OBIEDAT, Hussein. Collocations: a neglected variable in EFL. **IRAL**, v. 33, n.4, p. 315-31. Nov, 1995.

FIRTH, John Rupert. **Modes of Meaning: Papers in Linguistics 1934-1951**. London: Oxford University Press. 1957. Pp. 190-215.

HALLIDAY, Michael A. K. **Language as Social Semiotic**. London: Edward Arnold. 1978.

_____. "Lexis as a Linguistic Level." In: BAZELL, Charles Ernest; et al. (Eds.). **In Memory of J.R. Firth**. London: Longman. 1966.

HILL, Jimmie. Collocational Competence. **English Teaching Professional**, v. 11, p. 3-6. 1999.

HOEY, Michael. **Lexical Priming: a new theory of words and language**. London, New York: Routledge. 2005. 230 p.

_____. **Patterns of lexis in text**. London: Oxford. 1991. 293 p.

HOWARTH, Peter Andrew. Phraseology and Second Language Proficiency. **Applied Linguistics**, v.19, n. 1, p. 22-44. 1998.

HYMES, Dell. **On Communicative Competence**. Philadelphia, PA: University of Pennsylvania Press. 1972.

LAUFER, Batia; HILL, Monica. What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention?. **Language Learning & Technology**, v. 3, n.2, p. 58-76. 2000.



LEECH, Geoffrey N. **Semantics: The Study of Meaning**. Harmondsworth: Penguin. 1974. 400 p.

LEFFA, Vilson J. “Aspectos externos e internos da aquisição lexical.”. In: _____. (Org.). **As Palavras e Sua Companhia: O léxico na aprendizagem das línguas**. Pelotas: EDUCAT. 2000. Pp. 15- 44. 255 p.

LEWIS, Michael. (Ed.). **Teaching Collocations: Further Developments in the Lexical Approach**. Hove, UK: Language Teaching Publications. 2000. 248 p.

_____. **Implementing the Lexical Approach: Putting Theory into Practice**. Hove, UK: Language Teaching Publications. 1997.

_____. **The Lexical Approach**. Hove, UK: Language Teaching Publications. 1993.

LOURO, Inês da Conceição dos Anjos. **“Enxergando” as colocações: para ajudar a vencer o medo de um texto autêntico**. Tese de Doutorado. Faculdade de Filosofia Letras e Ciências Humanas da Universidade de São Paulo. 2001.

McCARTHY, Michael J. **Vocabulary**. Oxford: Oxford University Press. 1990. 173 p.

MOON, Rosamund. **Fixed Expressions and Idioms in English: A corpus-based approach**. Oxford: Clarendon Press, 1998. 358 p.

NATTINGER, James R; DeCARRICO, Jeanette S. **Lexical Phrases and Language Teaching**. Oxford: Oxford University Press. 1992. 240 p.

PARTINGTON, Alan. **Patterns and Meaning: Using Corpora for English Language Research and Teaching**. Amsterdam/Philadelphia: John Benjamins. 1998. 182 p.

PAWLEY, Andrew; SYDER, Frances. “Two puzzles for linguistic theory: nativelike selection and nativelike fluency.”. In: RICHARDS, Jack; SCHMIDT, Richard. **Language and Communication**. (Eds.) London: Longman. Pp. 191–226. 1983. 288 p.

PORTO, Melina. Lexical Phrases and Language Teaching. **FORUM**, v.36, n. 3, p: 22. 1998.

SARDINHA, Antônio Paulo Berber. Lingüística de Corpus: Histórico e Problemática. **DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada**, v.16, n.2, p: 323-67. São Paulo. 2000a.

_____. “Computador, corpus e concordância no ensino da léxico-gramática de língua estrangeira.”. In: LEFFA, Vilson J. (Org.). **As Palavras e Sua Companhia: O léxico na aprendizagem das línguas**. Pelotas: EDUCAT. 2000b. Pp. 47-65. 255 p.

SCHMITT, Norbert; UNDERWOOD, Geoffery. "Exploring the processing of formulaic sequences through a self-paced reading task.". In: _____. (Ed.). **Formulaic sentences: acquisition, processing and use**. London: John Benjamins Co. 2004. Pp. 173-90. 303 p.

SINCLAIR, McHardy John. **Reading Concordances: an introduction**. Softcover: Prentice Hall. 2003. 200 p.

_____. **Corpus, Concordance, Collocation**. Oxford: Oxford University Press. 4th Ed. 1991. 197 p.

SOUZA, Patrícia Nora de. O Uso da Hipermídia na Aprendizagem Implícita de Vocabulário. **The Specialist**, v. 28, n. 1, p. 59-85. 2007.

_____. **O uso da hipermídia no ensino e aquisição lexical no contexto da leitura em língua estrangeira**. Tese de Doutorado. Campinas, São Paulo: UNICAMP. 2004.

_____.; BRAGA, Denise Bértoli. "Os Efeitos da Hipermídia no Aprendizado.". In: SCARAMUCCI, Matilde V. Ricardi GATTOLIN, Sandra R. B. **Pesquisas sobre vocabulário em língua estrangeira**. Campinas, S.P.: Mercado das Letras. 2007. Pp. 155-74. 162 p.

TAGNIN, Ortweiler Esther Stella. **Os Corpora: Instrumentos de Auto-Ajuda para o Tradutor**. *Caderno de Tradução*, v. 1, n. 9, p.191 -219. 2002.

_____. **Convencionalidade e Produção de Texto: um dicionário de colocações verbais inglês/português - português/inglês**. Tese de Livre-Docência. São Paulo: Universidade de São Paulo. 1999.

_____. **Expressões Idiomáticas e Convencionais**. São Paulo: Ática. 1989. 88 p.

WANNER, Leo. Towards Automatic Fine-Grained Semantic Classification of Verb-Noun Collocations. **Natural Language Engineering**, v. 10, n.2, p. 95-143. Jun. 2004.

WEINERT, Regina. The Role of Formulaic Language in Second Language Acquisition: a review. **Applied Linguistics**, v.16, p: 180 - 205. 1995.

WRAY, Alison. **Formulaic language and the lexicon**. Cambridge: Cambridge University Press. 2002. 356 p.