

Caracterização formal das expressões idiomáticas em português

Lúcia Fulgêncio
Universidade Federal de Minas Gerais
luciafulgencio@hotmail.com

Resumo

A idiossincrasia semântica das expressões idiomáticas é fato bem conhecido, mas ainda falta uma descrição detalhada de suas idiossincrasias no plano morfossintático. Este estudo, baseado em dados do português, concentra-se nas seguintes marcas formais, encontradas apenas em expressões fixas idiomáticas: (a) ocorrência preferencial de formas marcadas, em especial do feminino e do plural; (b) a ocorrência de nomes como núcleo do SN que não funcionam como o *locus* morfossintático, que governa a concordância; e (c) a impossibilidade de participação em diversos processos gramaticais, como a inversão de termos em estruturas coordenadas. Esses traços conferem a base para a hipótese de que as expressões fixas devem ser analisadas como itens lexicais, não sendo sujeitas às regularidades que governam a estrutura interna dos sintagmas em geral. Além disso, chama-se atenção para um importante traço semântico, a saber, (d) a ocorrência de estruturas anafóricas sem referência. Os dados do presente estudo provêm da elaboração do *Dicionário de expressões fixas e convencionais do português brasileiro*, que revelou o número surpreendente de mais de 8.500 expressões de uso normal e corrente na língua contemporânea.

Palavras chave: expressões idiomáticas, expressões fixas, *markedness*, referência, anáfora

1. Expressões fixas e suas características

Uma das construções presentes nas línguas – assim como tratadas por autores como Goldberg (1995, 2006), Tagnin (1989, 2013) e Fulgêncio (2008) – são as construções fixas ou idiomáticas. Trata-se de sequências de palavras que são repetidas sempre com o mesmo formato, com o mesmo léxico e a mesma ordem, como por exemplo *de vez em quando, quem dera, sem mais aquela, de mais a mais, dar as caras, como se não bastasse, lua de mel e não estar nem aí*. Em espanhol temos casos como *estar entre dos fuegos, de la boca para afuera, lo que no mata engorda, cara o cruz, caso serio, a oscuras, por supuesto, en cuanto a, casi casi, mover cielo y tierra*. Eventualmente há variantes ou pequenas modificações. Por exemplo, são possíveis grupos com pequenas variações como *à proporção que / à proporção em que; a qualquer custo / a todo custo; fincar o pé / bater o pé*. No entanto, pode-se afirmar que essas alternâncias são limitadas e estão muito longe da variabilidade possível em enunciados montados pelo falante no momento da produção. Um levantamento recente para a confecção de um dicionário de expressões fixas e convencionais (Fulgêncio, a sair) revelou que o português contém aproximadamente 8.500 expressões fixas – contando-se apenas as de uso corrente na linguagem atual. Esse número mostra a grande importância das expressões fixas, que constituem uma parte significativa dos itens lexicais da língua.

A composição interna das expressões fixas¹ é idiossincrática, tanto do ponto de vista formal quanto lexical ou semântico – ou seja, não há motivação identificável. Por esse motivo são consideradas “convenções linguísticas” (Tagnin, 1989, 2013): têm tal formato, tal léxico, tal estrutura e tal significado por mera convenção. Isso faz com que

¹ Também chamadas fraseologismos.

se comportem como itens lexicais, a serem aprendidos individualmente, já que as regras sintáticas e semânticas não as descrevem.

Alguns linguistas mencionam uma suposta orientação metafórica quanto ao significado, mas essa posição é inadequada, como se mostra em Fulgêncio e Ciriaco (a sair). Quando em espanhol se usa a expressão *tener la mosca detrás de la oreja* o ouvinte não chega ao significado da expressão através de um processamento metafórico a partir das palavras *mosca* ou *oreja*; assim também em português, quando se diz que algo foi compreendido *ao pé da letra*, o significado do grupo não é extraído por meio de um processamento metafórico construído no momento da decodificação, envolvendo os itens *pé* ou *letra*. Idem para *de cabo a rabo*, *espírito de porco*, *da pá virada*, *afinal de contas*, *à beça*, *mandar às favas* e tantos outros dados: é bem evidente a impossibilidade de relação metafórica nesses casos. Mesmo em expressões como *céu da boca*, onde se pode imaginar alguma relação entre o significado dos itens e o significado da expressão, pode-se afirmar que o falante proficiente não processa os itens que compõem o grupo idiomático para chegar ao significado „palato“, uma vez que o significado da expressão se memoriza como um bloco já pronto. O falante e o ouvinte já têm armazenados os grupos fixos, e não precisam montá-los nem fazer qualquer processamento metafórico para chegar ao significado. As expressões fixas são agrupamentos de mais de uma palavra, cristalizados e acessíveis na memória lexical. Apesar de graficamente plurilexical, uma expressão fixa constitui uma unidade linguística coesa que compõe um único item.

Geralmente o foco de estudo dos fraseologismos se orienta para o âmbito semântico, já que os idiomatismos são caracterizados pelo fato de o significado de cada item individual não levar ao significado do grupo. As expressões fixas geralmente não são tratadas do ponto de vista formal, e são pouco estudadas as possíveis restrições de caráter sintático. Costuma-se pensar que as expressões fixas representam um grupo de significado não literal com léxico cristalizado, mas, no mais, supostamente se conformariam à estrutura da língua. Alguns linguistas inclusive afirmam que as expressões idiomáticas seriam bem formadas do ponto de vista sintático: “Sustento que as expressões idiomáticas são bem formadas, e que as regras de boa-formação são regras „parametrizadas“ simples.” (Williams, 1994:10).²

Outros autores chegam a perceber a existência de divergências sintáticas, mas nota-se uma tendência a vê-las como marginais:

Expressões com mais de uma palavra precisam se conformar às regras gramaticais da língua [...], embora idiosincrasias gramaticais realmente apareçam em algumas expressões bem conhecidas: [...]. (Fernando, 1996: 34)

Contrariamente à crença desses autores, a análise dos dados comprova que as expressões fixas funcionam como blocos que nem sempre se sujeitam às condições morfossintáticas da língua. É possível verificar que as locuções e expressões fixas, além de terem um significado particular, muitas vezes possuem restrições sintáticas que violam as regras de boa formação. Nem sempre as expressões fixas são bem comportadas, e nem sempre obedecem às regras que valem para a formação de sintagmas novos construídos pelo falante: em muitas delas encontramos idiosincrasias formais inesperadas. Paralelamente às questões de significado, as expressões fixas apresentam características formais muito especiais, que as distinguem das demais construções livres da língua.

² Tradução minha, aqui e nas outras citações de obras não em português.

Alguns desses aspectos formais foram observados por autores como Gross (1996: 12), que evidencia o “bloqueio de propriedades transformacionais” de algumas expressões. São mencionadas como idiossincrasias formais a impossibilidade de passivização, de pronominalização, de deslocamento, de relativização e de ocorrência em estruturas clivadas. Gross aponta que, para uma sentença construída pelo falante como *A polícia prendeu o assassino*, são possíveis estruturas como:

- passiva: *O assassino foi preso pela polícia.*
- pronominalização: *A polícia o prendeu.*
- topicalização: *O assassino, a polícia prendeu.*
- *cleft-sentence*: *Foi o assassino que a polícia prendeu.*
- relativização: *O assassino que a polícia prendeu [...]*

Porém, no caso de sintagmas verbais idiomáticos o mesmo não ocorre. Por exemplo, tanto no caso do português *pagar o pato* quanto no espanhol *pagar el pato* são inaceitáveis estruturas como as mencionadas acima. A partir de uma sentença como *Pedro pagou o pato* não são possíveis estruturas como as seguintes:

- passiva: **O pato foi pago por Pedro.*
- pronominalização: **Pedro o pagou.*
- topicalização: **O pato, Pedro pagou.*
- *cleft-sentence*: **Foi o pato que Pedro pagou.*
- relativização: **O pato que Pedro pagou [...]*

Isso mostra que *Pedro pagou o pato* não é um exemplo da construção transitiva: sujeito agente, verbo, objeto paciente. Há ainda vários outros casos em que as expressões fixas não respeitam a contraparte sintática da língua. No que se segue serão focalizadas somente quatro outras idiossincrasias formais, a saber: a presença de negativa obrigatória, a irreversibilidade da ordem sintática, a não evocação de referentes de nomes e o uso de anáforas ou nomes na forma marcada.

2. Negativa obrigatória

Observa-se em alguns fraseologismos a presença obrigatória da negativa – ou, em outras palavras, a impossibilidade da forma positiva, que é sempre possível em construções livres – como por exemplo: *não ter a ver*; *não caber em si*; *como se não bastasse* – ou em espanhol *no tener pelos en la lengua*.

Espera-se que, para qualquer construção negativa, possa corresponder uma estrutura positiva relacionada. Assim, da sentença *Não faça isso*, pode-se inferir a possibilidade da forma positiva correspondente, *Faça isso*. Nas expressões fixas, às vezes essa correspondência é rompida. Várias expressões só aparecem na forma negativa e recusam a estrutura positiva correspondente, como *não diga!* e também *não brinca!* Essas fórmulas exprimem espanto diante de um relato inesperado, e não admitem as formas positivas **diga!* / **brinca!* (com significado paralelo).

Outra observação sobre esses casos é que, apesar de em *não diga!* ou *não brinca!* aparecer a palavra *não*, as sentenças não são semanticamente negativas; isto é, com essas expressões não se pede para o interlocutor não falar ou não brincar, nem tampouco

se nega nada, mas somente indica-se espanto pelo fato relatado. Note-se, ainda, que *não diga* é usado mesmo por falantes que não usam *diga* como imperativo negativo – na fala coloquial brasileira do Sudeste o imperativo é *diz*, negativo *não diz*.

Outros exemplos de expressões que não admitem a forma positiva são os seguintes:

- *não ata nem desata* (**ata e desata*)
- *não falar coisa com coisa* (**falar coisa com coisa*)
- *não medir esforços* (**medir esforços*)
- *não há de ser nada* (**há de ser nada* / **há de ser alguma coisa*)
- *não fede nem cheira* (**fede e cheira*)

Também em espanhol há fraseologismos que só admitem a forma negativa, como por exemplo: *no caber ni un alfiler, no tener pelos en la lengua, no dar el brazo a torcer, ¡no digas!*

3. Ordem sintática: irreversibilidade e imobilismo

Em estruturas construídas (que não são idiomáticas), a conjunção *e* coordena palavras ou sintagmas de mesma classe, e normalmente esses dois elementos são permutáveis: [A e B] ou [B e A]. Na língua em geral, em sintagmas coordenados pode-se alternar a ordem, de forma que o primeiro elemento passa para a segunda posição e vice-versa. A reversibilidade em construções coordenadas construídas (não fixas) é permitida pela sintaxe: posso dizer *pera e laranja* ou *laranja e pera*, tanto faz. A ordem pode ser determinada dentro do texto, em função do tópico discursivo, por exemplo, mas em termos sintáticos a escolha é livre.

Por outro lado, numa expressão como *par ou ímpar, vivo ou morto, cara ou coroa, arco e flecha*, a ordem não pode ser alterada. No caso de expressões fixas, nem sempre a ordem dos elementos coordenados é livre: existe o conjunto *de forno e fogão*, mas não se diz **de fogão e forno*; o mesmo acontece em *firme e forte* ou *ser unha e carne*, onde os elementos coordenados não admitem a permuta de ordem (**forte e firme*, **ser carne e unha*). Vemos nesses casos que nem sempre as estruturas idiomáticas respeitam propriedades estruturais correntes na língua.

Os **qualificativos** podem ocorrer em ordem fixa nos fraseologismos, mesmo tendo ordem livre com relação ao núcleo do SN em sintagmas construídos, como se vê nos seguintes exemplos:

- *ter em alta conta* (**ter em conta alta*)
- *em altos brados* (**em brados altos*)
- *de largo espectro* (**de espectro largo*)

Isso vale igualmente para outros tipos de expressão fixa:

- *por assim dizer* (**por dizer assim*)
- *pôr uma pedra em cima* (**pôr em cima uma pedra*)
- *fazer uma tempestade em copo d'água* (**em copo d'água, fazer uma tempestade*)

4. Relação simbólica irregular

Um aspecto das expressões fixas que parece especialmente interessante é que, mesmo sem considerar possibilidades de modificações ou correspondências formais, observa-se que uma parte delas são mesmo agramaticais, assim como se apresentam na superfície. Ou seja, há expressões cuja composição interna infringe as regras estruturais da língua. Se esses sintagmas fossem montados pelo falante, seriam necessariamente excluídos, pelo fato de serem mal formados na sua configuração superficial, não obedecendo aos parâmetros sintático-semânticos da língua.

Uma questão relativa às expressões idiomáticas nominais diz respeito ao **referente**. Em sintagmas componenciais, o referente do SN é sempre um elemento mais restringido do que o que é evocado pelo núcleo do sintagma. Por exemplo: no SN *o livro de linguística em cima da mesa*, temos um núcleo genérico *livro* que é restringido pelos demais elementos do sintagma, de modo que o referente do SN integral é um ser mais específico („aquele livro em especial que é de linguística e que está em cima da mesa“). Em sintagmas montados pelo falante, o referente do SN inteiro é sempre um membro do conjunto evocado pelo núcleo.

Em um SN, o nome é o núcleo e o centro de referência (Perini *et al.*, 1996: 83; Liberato, 1997). Esse centro de referência pode se igualar à referência de todo o SN (por exemplo, quando o SN é só o nome *Pedro*) ou pode ser um hiperônimo do referente do sintagma, já que os demais integrantes do SN (determinantes, modificadores etc) operam de modo a restringir o referente evocado pelo núcleo.

O que acontece em expressões fixas é diferente: se decomposmos as expressões nominais vemos que em muitos casos o referente do grupo inteiro não é nem mesmo semelhante ao referente do nome que integra essa expressão. Por exemplo, em expressões como *água-marinha*, *ar-condicionado* ou *meio-fio*, o referente é totalmente diferente do hipotético referente dos nomes *água*, *ar* ou *fio*. Digo “hipotético” porque esse referente do nome **não** entra na composição do referente do idiomatismo. Por exemplo: em *água-marinha* o referente é „um tipo de pedra preciosa azul“, enquanto o referente do núcleo, caso fosse processado segundo as regras regulares, seria „água“ (que não é acionado na mente do ouvinte). No caso de *água-marinha*, a pedra preciosa não é uma restrição de *água*, e o modificador não restringe o núcleo. Também em *ar-condicionado* ocorre o mesmo: o aparelho que tem esse nome não é um tipo de *ar*.

Ou seja, os componentes do SN (determinantes, modificadores, etc), que normalmente servem para restringir o referente evocado pelo núcleo do SN (formado pelo nome), no caso das expressões idiomáticas não servem para restringir ou delimitar o referente do nome. Nesse caso só há um único referente evocado: o da expressão completa. Portanto, todos os elementos componentes da expressão ajudam igualmente na definição desse referente. Todos os elementos (tanto o nome como os seus satélites – determinantes e modificadores) têm o mesmo papel na composição do referente de compostos nominais, e o referente não é acionado prioritariamente pelo núcleo. Esse tipo de delimitação do referente é totalmente diferente do que acontece em sintagmas componenciais, montados pelo falante no momento do enunciado.

Resumindo: com SNs composicionais (de formação regular) acontece o seguinte:

- O núcleo evoca um referente mais genérico que é restringido pelos demais componentes do sintagma, de forma que o referente do sintagma inteiro é mais específico do que o referente do núcleo; um SN evoca, portanto, no mínimo dois referentes: o do sintagma e o do núcleo.

- Todos os nomes que compõem o SN têm o seu referente evocado.
- O núcleo é não só o controlador da referência, como também o controlador morfossintático (segundo Zwicky, 1993, o “*morphosyntactic locus*”). Isso quer dizer que se o núcleo é masculino, o SN inteiro também é; se é feminino, o SN também é.

Em alguns SNs que incluem expressões idiomáticas a estrutura interna é diferente daquela que se vê em SNs composicionais:

- O núcleo, por si só, não evoca referente nenhum.
- Se não há um referente para o núcleo, só há um referente acionado: o da expressão completa. Isto é, não é acionado um referente do nome e outro mais especificado para o grupo nominal idiomático, mas somente um único referente.
- Os itens que integram o grupo nominal, como os qualificativos, não servem para restringir o referente evocado pelo nome, como acontece em sintagmas construídos; ao contrário, servem para definir o próprio referente idiossincrático, que pode ser totalmente diferente do referente do nome isoladamente. Os termos integrantes do grupo idiomático (determinantes, modificadores etc) participam da composição conceitual e da identificação do referente tanto quanto o núcleo.
- O núcleo não é o controlador morfossintático. Nas expressões idiomáticas é possível que o nome mencionado seja de um gênero e o SN como um todo seja de outro gênero. Temos como exemplo “**um curta metragem**”: o SN é masculino, mas o nome *metragem* é feminino, o que fica evidenciado inclusive no modificador *curta* que aparece no feminino. Outros exemplos em que o gênero do SN é diferente do gênero do nome são: **o pele vermelha** (mas: *a pele*), **o testa de ferro** (mas: *a testa*), **o bóia fria** (mas: *a bóia*), **o regra-três** (mas *a regra*).
- Nenhum nome que integra o composto idiomático pode ser retomado por meio de clítico ou de anáfora pronominal. Só é possível retomar anaforicamente o referente de toda a expressão fixa.

Também os compostos que apresentam a forma [V – Nome] – como por exemplo **guarda-volumes**, **beija-flor**, **pica-pau**, **marca-passos**, **porta-estandarte**, **quebra-cabeça**, **salva-vidas**, **tira-gosto** ou **vira-lata** – apresentam características inesperadas, bem diferentes das regras que operam em sintagmas construídos pelo falante:

- Do ponto de vista morfológico, esses compostos têm o verbo congelado na 3ª pessoa do singular do presente do indicativo. Em alguns casos (onde a formação é diferente de [V – Nome]) o verbo pode aparecer em outro tempo verbal que não o presente, como em *bem-te-vi*, *disse-me-disse*, *deus-dará*, *deus-nos-acuda* e *venha a nós*, mas sempre aparece congelado (isto é, não pode ser conjugado em outros tempos ou pessoas).
- Do ponto de vista sintático, os compostos [V – Nome] comportam-se como um nome e não como um sintagma verbal (Basilio, 2000:16), tendo as seguintes características:
 1. Não é possível coordenar elementos internos ao composto, como se vê pela impossibilidade de
 - (1) O beija (*e cheira) – flor é pequeno.
O beija-flor (*e frutos) é pequeno.

2. Não é possível intercalar elementos entre o verbo e o nome, como se vê pela impossibilidade de

(2) O beija (*qualquer tipo de) – flor é pequeno.

O beija (*a linda) – flor é pequeno.

- A relação semântica entre a expressão completa e seus componentes formais seria, se levada em conta, idiossincrática, quase chegando ao ponto em que cada caso tenha que ser descrito separadamente: *beija-flor* tem algo a ver com flor, embora o esquema FLOR não seja evocado; *bem-te-vi* não tem nada a ver com nenhum de seus componentes, se os consideramos como palavras separadas da língua.

5. Retomada anafórica

No caso de expressões fixas e idiomáticas é também possível o uso de itens ou lacunas anafóricas **sem referente**, o que nunca ocorre em construções livres. Se isso ocorresse numa construção nova (não fixa), não seria possível determinar o significado, e a comunicação seria afetada negativamente. Mas numa expressão idiomática, como o significado é fixo, já convencionado e já determinado para o bloco inteiro, o fato de haver anáforas com referente não identificável não interfere na compreensão, uma vez que o significado das partes não é computado. Em situação composicional, não idiomática, o ouvinte precisa recuperar o referente das anáforas para montar o significado; mas quando se trata de uma expressão fixa, o significado já está pronto na memória lexical. Não interessa qual é o referente de um item anafórico que possa aparecer dentro de uma expressão – só interessa o significado do grupo idiomático inteiro, que já está determinado *a priori*. Exemplos de anáfora sem referente são *estar na Ø dele*, *dar uma Ø de*, *não estar nem aí*, *ora essa*, *sem mais aquela*, *sem mais Ø nem menos Ø* e *ficar elas por elas*.

No grupo dos idiomatismos encontram-se expressões como *cair nessa*, *corta essa*, *ficar elas por elas* e *estar na mesma*. Observe-se que a anáfora aparece no feminino – nesses e em muitos outros casos. Se não há antecedente de gênero feminino nem referente de sexo feminino, seria de se esperar que a anáfora migrasse para a forma não-marcada, que é a do masculino singular (Martin, 1975). Estranhamente, a anáfora aparece no feminino, que é uma forma marcada no português. Trata-se de uma situação inesperada.

O mesmo fenômeno de uso de formas no feminino ou no plural sem relacionamento com um referente feminino ou plural parece ocorrer também em outras línguas: em espanhol, nas expressões *de una* (= sem rodeios, diretamente), *de ultima* (em último caso), *a oscuras* (no escuro), *a las perdidas* (de tanto em tanto) ou *hacerla corta* (ser breve); em italiano, nas formas *parcela* (conseguir) ou *prendersela* (irritar-se), onde também aparecem anáforas femininas sem referente.

A conclusão que se impõe é que as expressões fixas devem ser analisadas como itens lexicais, sem partes gramaticalmente articuladas, tal como não se tentaria explicitar na gramática as relações formais e semânticas de uma palavra simples como *mesa*. E, tal como os itens lexicais em geral, as expressões fixas devem ser aprendidas individualmente, uma por uma – o que, dado o seu grande número na língua, nos diz algo sobre a capacidade que têm os humanos de armazenar informação idiossincrática.

Referências

- Basilio, M. (2000). “Em torno da palavra como unidade lexical: palavras e composições”, *Veredas – revista de estudos lingüísticos*, Juiz de Fora: Editora UFJF, v. 4, n. 2, 9- 18.
- Fernando, C. (1996). *Idioms and idiomaticity*. Hong Kong: Oxford University Press.
- Fulgêncio, L. (a sair). *DEIC – Dicionário de expressões idiomáticas e convencionais*.
- Fulgêncio, L. (2008). *Expressões fixas e idiomatismos do português brasileiro*. Tese de doutorado, Belo Horizonte (Brasil): PUC Minas.
- Fulgêncio, L. e Ciríaco, L. (a sair). *The non-metaphoric nature of idioms*.
- Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- Goldberg, A. (2006). *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.
- Gross, G. (1996). *Les expressions figées en français*. Paris: Ophrys.
- Martin, J. W. (1975) “Concordância?” *Revista Brasileira de Linguística*, vol. 2, n. 2.
- Perini, M. A. *et al.* (1996). *O sintagma nominal em português: estrutura, significado e função*. Belo Horizonte: Número especial da *Revista de estudos da linguagem* – Faculdade de Letras da UFMG.
- Tagnin, S. E. O. (1989). *Expressões idiomáticas e convencionais*. São Paulo: Editora Ática.
- Tagnin, S. E. O. (2013). *O jeito que a gente diz: combinações consagradas em inglês e português*. São Paulo: Disal Editora.
- Williams, E. (1994). “Remarks on lexical knowledge.” In Gleitman, L. & Landau, B. (eds.), *The acquisition of the lexicon*. Cambridge, Mass.: The MIT Press.
- Zwicky, A. (1993). “Heads, bases and functors.” In Corbett *et al.* (eds), *Heads in grammatical theory*. Cambridge: Cambridge University Press.