

## Idiomaticidade, familiaridade e informação prévia no processamento de expressões idiomáticas do PB

*Idiomaticity, familiarity and preceding information in BP idioms processing*

Sara de Oliveira Gomes Barreto

Universidade Salgado de Oliveira, Juiz de Fora, MG, Brasil

Mercedes Marcilese

Ágata Jéssica Avelar de Oliveira

Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brasil



**Resumo:** Expressões idiomáticas são comumente caracterizadas como combinações fixas e não-composicionais nas quais o sentido figurado não constitui uma função da soma das partes; seu grau de composicionalidade, entretanto, é variável. Este artigo investiga a compreensão de expressões idiomáticas do tipo verbo-complemento no português brasileiro (ex. “pintar o sete”, “passar a bola”). Buscamos avaliar em que medida propriedades internas e externas das expressões podem influenciar no processamento das mesmas. Reportamos os resultados de um experimento de leitura automonitorada (conduzido por meio de uma *maze task*), no qual o segmento crítico corresponde à complementação do verbo: idiomático vs. literal (ex. abrir *o coração/ a porta*). Os resultados revelaram efeitos estatisticamente significativos das três variáveis investigadas, sugerindo que o *grau de idiomaticidade*, a *familiaridade* do falante com a expressão e a *informação prévia* afetam o processamento das expressões e, portanto, devem ser considerados num modelo que busque explicitar a representação, o acesso e a compreensão dessas estruturas.

**Palavras-chave:** Expressões idiomáticas; Idiomaticidade; Familiaridade; Informação prévia; Processamento linguístico

**Abstract:** Idioms are typically characterized as non-compositional strings of words in which figurative meaning cannot be derived from the interpretation of its parts; the degree of compositionality, however, is variable. This paper investigates the comprehension of verb + complement idioms in Brazilian Portuguese (e.g. “passar a bola”, “pintar o sete”). Our research intends to explore to what extent internal and external factors affect the comprehension of idioms. We report the results of a maze task (self-paced reading experiment) in which the critical segment was related to the verb complement: idiomatic vs. literal complement (e.g. open the heart/ the door). The results revealed significant effects of the three investigated variables, suggesting that the degree of idiomaticity, the speaker's familiarity with the idiom and the previous context affect the processing of the idioms and, therefore, should be considered in a model that aims to explain the representation, access and comprehension of these structures.

**Keywords:** Idioms; Idiomaticity; Familiarity; Preceding information; Language processing

## Introdução

Na abordagem tradicional, associada à visão aristotélica sobre a linguagem figurada, o sentido literal é assumido como convencional e não geraria um esforço especial para sua compreensão (no contexto de propostas de cunho lógico-filosófico e/ou linguístico, cf. SEARLE, 1993/1979; GRICE, 1975; no âmbito de estudos sobre o processamento da linguagem, cf. JANUS; BEVER, 1985; BOBROW; BELL, 1973; para uma caracterização geral de perspectivas de cunho tradicional, cf. GLUCKSBERG, 2003; GLUCKSBERG; KEYSAR, 1993; BÖRJESSON, 2011). No que tange ao processamento linguístico, de acordo com essa perspectiva, o sentido literal de expressões simples se encontraria armazenado nas entradas lexicais; já o sentido literal de expressões complexas seria o resultado de uma combinação dos sentidos literais das partes. Em ambos os casos, a interpretação seria, em boa medida, independente de contexto. O sentido não literal, por sua vez, é tido na abordagem tradicional como não convencional. Por esse motivo, demandaria um esforço interpretativo especial para ser alcançado e apenas seria processado se a interpretação literal não se mostrar adequada ao contexto. De acordo com essa visão sobre a linguagem figurada, a compreensão do sentido não literal ocorreria, portanto, de forma indireta.

Pesquisas desenvolvidas no âmbito da Semântica Cognitiva têm questionado a visão tradicional do sentido não literal ou figurado/metafórico como um fenômeno estritamente linguístico propondo uma caracterização do sentido metafórico sob os termos de processos conceituais (LAKOFF; JOHNSON, 1980; GIBBS, 1994; LAKOFF; TURNER, 1989, dentre outros). De acordo com essa abordagem, o sentido (linguístico ou não) é construído a partir de estruturas conceituais convencionalizadas, e as categorias mentais formam-se a partir da experiência de desenvolvimento e de atuação no mundo dos falantes. Lakoff e Johnson (1980) consideram que a metáfora seria um elemento essencial para a categorização do mundo e para a formação de processos mentais.

De modo geral, estudos contemporâneos sobre os usos não literais da linguagem têm contribuído para destacar a relevância da linguagem figurada na cognição humana, uma vez que ela não mais é vinculada apenas ao discurso poético, mas reconhecida como instrumento da linguagem cotidiana. Não existe, no entanto, uma abordagem única para dar conta do aparente contraste que, em termos de processamento, existiria entre linguagem literal e não literal. Como veremos mais à frente, há propostas que defendem o processamento direto do sentido figurado (cf. GLUCKSBERG, 2003; dentre outros), mas há também quem postule que só se chega à interpretação metafórica a partir da derivação

inicial do significado literal da expressão, seguida pela sua rejeição dentro de um dado contexto e a subsequente busca por uma alternativa apropriada, o que envolveria um processamento indireto em três estágios (JANUS; BEVER, 1985; BOBROW; BELL, 1973; dentre outros).

Há ainda propostas que podem ser vistas como “unificadoras” ou “conciliadoras”, já que postulam que ambos os tipos de processamento – direto ou em estágios – podem de fato acontecer a depender de um conjunto de fatores (GIORA, 1997, 1999; BOWDLE; GENTNER, 2005).

De acordo com Saeed (2003, p. 244), expressões idiomáticas são estruturas formadas por mais de uma palavra, não composicionais e seriam aprendidas pelos falantes como unidades semânticas. As expressões surgiriam a partir de um processo de fossilização pelo qual as palavras individuais não preservariam mais seus sentidos independentes quando combinadas no interior da expressão. Nesse sentido, as expressões parecem desafiar a visão segundo a qual, para que se proceda à compreensão de uma sentença, é necessário que sejam reconhecidas as palavras que a compõem, que seus sentidos individuais sejam acessados no léxico mental e que esses elementos sejam combinados conforme determinadas restrições gramaticais. Expressões idiomáticas constituem assim um desafio para as teorias do processamento da linguagem, até mesmo pelo fato de serem bastante comuns nas línguas, isto é, não constituem exceções, mas fazem parte do funcionamento “normal” dos sistemas linguísticos.

Apesar da definição mais difundida e apresentada anteriormente, que salienta a propriedade da não composicionalidade, muitas expressões idiomáticas são decomponíveis em algum grau, uma vez que o sentido de suas partes parece contribuir de forma independente para a compreensão do sentido figurado da expressão como um todo (GIBBS; NAYAK, 1989; GIBBS, NAYAK et al., 1989; NUNBERG, 1978). O grau de decomponibilidade das mesmas seria, portanto, variável (GIBBS, 1989) e as expressões poderiam ser classificadas em função de um “espectro de idiomaticidade”: em alguns casos, as partes colaboram de forma mais ou menos explícita na construção do sentido figurado, em outras, essa contribuição não é transparente. Por exemplo, em expressões do português brasileiro (doravante, PB) como “passar a bola”, percebe-se uma maior decomponibilidade do que em “dar zebra” ou em “pintar o sete”.

Certas expressões admitem ainda a possibilidade de serem interpretadas também como sentenças literais (cf. “passar a bola”, “arregaçar as mangas”, “lavar as mãos” etc.). Além disso, expressões diferem em termos de seu nível de cristalização (MUELLER; GIBBS, 1987), do tipo de operações sintáticas às quais podem ser submetidas (FRASER, 1970), e também quanto a outros aspectos

que podem se mostrar importantes relativamente ao modo como as mesmas são processadas e compreendidas pelos falantes. Mesmo expressões “muito idiomáticas” diferem-se umas em relação às outras em termos de quão familiares ou convencionais são para os falantes e, ainda, há de se considerar a possibilidade de influência da informação que preceda a leitura ou a escuta da expressão idiomática e que possa, ou não, favorecer a compreensão metafórica da expressão. Nesse sentido, o quanto o contexto prévio pode facilitar o acesso aos sentidos literal ou figurado é uma questão relevante para a pesquisa no que tange ao denominado “problema do reconhecimento” (Stern, 2000): como os falantes identificam que uma dada estrutura linguística veicula ou não um sentido metafórico?

Além dos fatores mencionados, a própria estrutura interna das expressões idiomáticas apresenta uma considerável variação, o que torna difícil comparar, em termos da sua natureza nominal ou verbal, expressões como: “levado da breca” (pessoa que faz coisas impensadas), “barra pesada” (situação difícil, pessoa grosseira ou violenta), “cutucar a onça com vara curta” (arrumar problema, provocar); “mão de vaca” (indivíduo que não gosta de gastar, mesquinho) etc.

Diante da variabilidade de fatores, tanto intrínsecos às próprias expressões quanto externos a serem considerados, as seguintes questões de pesquisa podem ser levantadas: (i) Em que medida os aspectos levantados se mostram relevantes no processamento de expressões idiomáticas pelos falantes no que diz respeito à sua compreensão? (ii) Tais aspectos atuam de forma isolada ou a compreensão de expressões depende da interação de um conjunto de fatores? (iii) Levando em consideração o curso temporal do processamento, os fatores levantados podem influenciar em alguma medida qual sentido é acessado em primeiro lugar (idiomático ou literal)? O presente artigo tem como objetivo principal discutir esses pontos e reportar evidências experimentais que podem contribuir para iluminar as questões colocadas.

Dada a impossibilidade de se investigar simultaneamente todos os tipos de expressões que podem ser identificados no PB, optamos por selecionar um único tipo formado por verbo (V) acompanhado de complemento. Escolhemos assim expressões nas quais o complemento equivale a um sintagma formado por um, dois ou até três elementos. Selecionamos para nossa pesquisa empírica expressões idiomáticas verbais nas quais o complemento poderia aparecer na forma de um sintagma nominal (1), de um sintagma determinante (2) ou de um sintagma preposicionado (3):

- (1) V+NP (*Noun Phrase*)  
– como em  $_{VP}$ [dar  $_{NP}$ [zebra]].

- (2) V+DP (*Determiner Phrase*)  
– como em  $_{VP}$ [chutar  $_{DP}$ [o  $_{NP}$ [balde]]].
- (3) V+PP (*Prepositional Phrase*)  
– como em  $_{VP}$ [carregar  $_{PP}$ [em  $_{DP}$ [as  $_{NP}$ [tintas]]]]].

Com base nos resultados experimentais obtidos em pesquisas sobre o processamento de expressões idiomáticas, bem como de outras ocorrências de linguagem não literal, partimos da ideia de que a compreensão de expressões idiomáticas envolve não apenas informação puramente sintática e lexical, mas também elementos vinculados ao conhecimento de mundo, incluindo aspectos culturais e pragmáticos. Nesse sentido, assumimos que vários fatores podem influenciar o processamento desse tipo de expressão (facilitando ou dificultando sua compreensão), dentre os quais focalizaremos aqui: o grau de idiomaticidade da expressão (GIBBS; NAYAK, 1989; GIBBS, NAYAK; CUTTING, 1989), a familiaridade do falante com a mesma (em termos da maior ou menor experiência consciente do falante com a expressão) (BLASKO; CONNINE, 1993), além da informação prévia fornecida pelo contexto discursivo no qual a expressão se encontra inserida (i.e. o contexto pode fornecer pistas mais ou menos explícitas quanto a uma eventual interpretação figurada das expressões) (CACCIARI; TABOSSI, 1988; McGLONE et al., 1994).

## 1 O processamento da linguagem não literal

Até algumas décadas atrás, a maioria dos modelos de compreensão do discurso, chamados de modelos pragmáticos padrão (SPM – *Standard Pragmatic Models*), preconizavam a prioridade do sentido literal sobre o não literal (cf. GLUCKSBERG, 2003; GLUCKSBERG; KEYSAR, 1993 para uma caracterização da perspectiva tradicional nesses moldes). São exemplos desta perspectiva as propostas Janus e Bever (1985) para a compreensão de metáforas e de Bobrow e Bell (1973) no que tange às expressões idiomáticas, dentre outras. De acordo com essa visão, a ativação do sentido literal seria compulsória: os falantes/ ouvintes computariam, primeiramente, o sentido literal e, na sequência, apenas se houvesse alguma incompatibilidade entre o sentido literal e o contexto, seria computado o sentido não literal, sendo este opcional, eventual e indiretamente processado (nunca como primeira opção).

Diferentemente de abordagens que assumem um processamento indireto, há propostas que não atrelam a compreensão do sentido figurado – seja no caso das metáforas, das expressões idiomáticas ou de outras formas de linguagem não literal – à inviabilidade de interpretação literal dentro de um dado contexto.

Propostas mais recentes se distanciam dos modelos pragmáticos antes mencionados e assumem que, como qualquer outra forma de compreensão da linguagem, a compreensão da metáfora e do sentido figurado em geral não seria opcional, mas mandatória e automática. Assim, o acesso ao sentido metafórico não seria mediado por uma rejeição do sentido literal e nem modulado a partir de uma comparação – já que o sentido metafórico não seria construído por uma comparação indireta ou implícita –, mas aconteceria em virtude de uma “rota” direta para interpretação figurada (GLUCKSBERG, 2003; GILDEA; GLUCKSBERG, 1983; GLUCKSBERG; NEWSOME; GOLDVARG, 2011; dentre outros; para um trabalho sobre o processamento de metáforas nominais do PB com resultados compatíveis nessa direção, ver RICCI, 2016).

No contexto da hipótese do processamento direto da linguagem não literal, Glucksberg (2003) formula um modelo para compreensão de metáforas baseado na ideia de “inclusão de classe”. Glucksberg argumenta que metáforas não constituem comparações indiretas, ou implícitas, mas declarações categóricas de inclusão de classe. De acordo com o autor, para além da similaridade, metáforas são compreendidas diretamente, já que os falantes interpretam expressões do tipo *my lawyer is a shark* (“meu advogado é um tubarão”) como literalmente falsas, ao passo que, as símiles correspondentes, como *my lawyer is like a shark* (“meu advogado é como um tubarão”) são interpretadas como verdadeiras. Glucksberg (2003) considera que a categoria *lawyer-shark* (“advogado-tubarão”) só pode ser considerada literalmente falsa ao se pensar na criatura marinha, mas não se se levar em conta um maior grau de abstração para se referir à categoria de predadores em geral. Veículos metafóricos como *shark* admitiriam uma referência dupla que, na maioria das metáforas, encontra-se implícita. Em que medida uma análise semelhante pode ser estendida para outras formas de linguagem não literal é uma questão a ser investigada (pensemos aqui em expressões idiomáticas que também admitem uma referência dupla – literal e não-literal – como “descascar o abacaxi”, por exemplo).

Além do modelo pragmático de acesso indireto e da proposta de acesso direto ao sentido figurado, na literatura podem ser encontradas abordagens que postulam que ambos os tipos de processamento – indireto e direto – podem de fato acontecer. A hipótese da “trajetória da metáfora” (*Career of metaphor hypothesis*), formulada por Bowdle e Gentner (2005), diz respeito a um espectro de progressão das expressões metafóricas: desde quando são “novas” – sendo interpretadas indiretamente como comparações implícitas – até o ponto em que se conclui o processo de convencionalização e passam a ser processadas preferencialmente por categorização (i.e. processamento direto). Segundo os autores: “a hipótese

da carreira da metáfora fornece uma estrutura teórica unificada que pode resolver o debate entre modelos de comparação e de categorização da metáfora” (tradução nossa) (BOWDLE; GENTNER, 2005, p. 193). A hipótese de Bowdle e Gentner para as metáforas pode ser facilmente estendida no caso das expressões idiomáticas que também passam por processos de convencionalização e cristalização.

Por sua vez, a Hipótese do Grau (ou gradiente) de Saliência (GIORA, 1997; GIORA; FEIN, 1999) defende que a distinção entre significado literal e não literal seja substituída pela ideia de significados mais e menos salientes, de modo que um princípio geral de saliência dos significados é que seria relevante para dar conta da compreensão da linguagem. Significados mais salientes ou convencionalizados, frequentes, familiares e autorizados por contexto prévio seriam processados de forma direta e prioritária, ao passo que os sentidos menos salientes o seriam de forma indireta; já no caso de sentidos igualmente salientes, estes seriam processados em paralelo.

Giora (1999, 2003) defende que a maioria das palavras apresenta diversos significados que variam em termos de sua saliência relativa, de forma que, quando um significado metafórico se mostra altamente saliente, ele será acessado primeiramente e a metáfora, rapidamente compreendida. Já um sentido metafórico menos saliente será compreendido de maneira mais lenta. O sentido de uma palavra, ou de uma expressão seria saliente basicamente em função de sua convencionalidade e de sua familiaridade. Se uma palavra ou expressão apresenta dois sentidos recuperáveis diretamente do léxico, aquele que for mais “popular”, mais prototípico, ou ainda, aquele que for usado com mais frequência em uma dada comunidade será o mais saliente. Outro fator a ser considerado, conforme a autora, é o contexto: o sentido ativado pelo contexto prévio, ou ainda, aquele que for mais previsível no contexto prévio será o mais saliente. Todos esses fatores teriam um papel no grau de saliência associado ao sentido atribuído a uma dada expressão. Vários dos pontos aventados por Giora (1997, 1999, 2003) serão retomados ao apresentarmos nossa pesquisa empírica.

## 2 As expressões idiomáticas: representação, acesso e processamento

Expressões idiomáticas são caracterizadas como cadeias de palavras cuja interpretação não se deriva de forma composicional a partir da interpretação de suas partes (CACCIARI; TABOSSI, 1988). Certas expressões são, no entanto, decomponíveis em algum grau. No que diz respeito ao papel da composicionalidade na compreensão



de expressões idiomáticas, não existe um consenso entre as diversas análises disponíveis. Na literatura, podem ser identificadas três perspectivas principais nesse sentido: (i) abordagens que rejeitam o papel da composicionalidade e, segundo as quais, expressões idiomáticas seriam representadas no léxico mental como itens listados num léxico idiomático específico ou como “palavras longas” (BOBROW; BELL, 1973; SWINNEY; CUTLER, 1979; dentre outros); (ii) teorias de cunho composicional, que defendem que a estrutura semântica interna da expressão é relevante para a compreensão do sentido idiomático (GIBBS; NAYAK, 1989; GIBBS, NAYAK; CUTTING, 1989; dentre outros) e (iii) propostas “híbridas” que incorporam características tanto de abordagens composicionais quanto não composicionais; i.e. defendem que a compreensão das expressões idiomáticas pode ocorrer em função de análise composicional e/ou de acesso direto, quando as palavras que compõem a expressão constituem uma configuração familiar e facilmente reconhecível pelo falante (CACCIARI; TABOSSI, 1988; TITONE; CONNINE, 1999).

Dentre as propostas que rejeitam o papel da composicionalidade na compreensão das expressões, destacamos o modelo de processamento serial baseado na hipótese de uma lista de expressões idiomáticas (BOBROW; BELL, 1973; CHAFE, 1968; FRASER, 1970; WEINREICH, 1968). Nessa perspectiva, os falantes, em um primeiro momento, realizariam uma análise literal da expressão idiomática, descobrindo o sentido metafórico pretendido apenas em um segundo momento. Para Bobrow e Bell (1973), expressões idiomáticas são mentalmente representadas e processadas como itens lexicais. As expressões não seriam listadas no léxico simplesmente como “palavras longas”, mas armazenadas de forma independente em um “léxico idiomático” específico. De acordo com essa visão, a compreensão de expressões idiomáticas ocorreria de acordo com um modelo em três estágios: o falante, primeiramente, processaria o sentido literal, em seguida essa interpretação inicial seria rejeitada e finalmente o léxico idiomático seria consultado e forneceria a interpretação adequada. Assim sendo, os falantes precisariam de um tempo maior para alcançar a compreensão de expressões idiomáticas do que para processar usos literais das mesmas expressões, ou sentenças cujos sentidos literais fossem equivalentes às expressões.

Outra abordagem para o processamento de expressões idiomáticas, frequentemente chamada modelo de processamento duplo (ESTILL; KEMPER, 1982; SWINNEY; CUTLER, 1979), defende uma análise do sentido literal da expressão idiomática em paralelo com a recuperação de sua interpretação metafórica. Swinney e Cutler (1979) também assumem uma representação

mental diferenciada para as expressões idiomáticas (Hipótese da representação lexical). Neste caso, no entanto, embora seja defendido que as expressões seriam mentalmente representadas e processadas como unidades lexicais, as mesmas seriam armazenadas no léxico “normal” e não num léxico idiomático especial, como proposto por Bobrow e Bell (1973). O processo de recuperação começaria assim que a primeira palavra da expressão fosse reconhecida e aconteceria em paralelo com a computação do sentido literal da expressão.

A hipótese do acesso direto (GIBBS, 1980, 1986), por sua vez, defende que o sentido figurado de uma expressão idiomática pode ser diretamente acessado sem que ocorra a análise de seu sentido literal. Nessa perspectiva, a identificação do sentido figurado de uma expressão não se daria em paralelo à computação de seu sentido literal. Em vez disso, ela precede a derivação do sentido literal, a qual só ocorreria nos casos em que a interpretação idiomática da cadeia de palavras não fosse bem sucedida em sua integração com o contexto.

Gibbs et al. (1989) investigaram em que medida o grau de decomponibilidade semântica das expressões influenciaria o processamento das mesmas. Inicialmente, os autores levantaram a hipótese de que a análise que os falantes realizam ao ouvir ou ler cada uma das partes de expressões semanticamente decomponíveis tornaria mais lento o processamento das mesmas, em comparação com o processamento de expressões não decomponíveis – sendo, estas, processadas como unidades semânticas cristalizadas. Os resultados reportados sugerem, no entanto, que os falantes compreendem expressões decomponíveis mais rapidamente do que as semanticamente não decomponíveis. Diante de uma expressão decomponível, os participantes da pesquisa conseguiram atrelar significados individuais a suas partes e, rapidamente, reconhecer a forma significativa segundo a qual essas partes individuais combinam-se para formar o sentido figurado geral da expressão idiomática. Assim, contrariando as previsões iniciais, os resultados obtidos por Gibbs e Nayak (1989) e Gibbs et al. (1989) parecem indicar que a compreensão de expressões semanticamente não decomponíveis torna-se mais difícil, exatamente porque o seu sentido geral não pode ser determinado por suas partes individuais. É importante destacar que os autores questionam em que medida a análise literal de uma expressão idiomática pode ser considerada como equivalente à sua análise composicional, e ainda questionam qual seria o papel do sentido literal na compreensão das expressões idiomáticas.

Gibbs et al. (1989) consideram que a hipótese da decomponibilidade das expressões idiomáticas, por eles formulada, é neutra relativamente à possibilidade de os

falantes ativarem o sentido literal das palavras constituintes da expressão, ou não, durante o processamento das expressões idiomáticas. Segundo esta hipótese, a princípio, os falantes reconheceriam as partes individuais das expressões idiomáticas enquanto possuidoras de sentidos independentes e figurados, os quais contribuem para o sentido figurado geral da expressão. Os autores consideram que, em geral, os falantes tendem a realizar análises composicionais das cadeias idiomáticas ao longo da compreensão, nas quais as partes das expressões são tidas como independentes, mas não necessariamente literais, e cujos significados contribuem para os sentidos figurados globais dessas expressões. Dessa forma, os processos requeridos para a compreensão de expressões idiomáticas não difeririam daqueles envolvidos na compreensão da linguagem literal. De modo semelhante ao que ocorre na compreensão da linguagem literal, as expressões idiomáticas parecem se submeter aos mesmos processos em termos de computação sintática que aquela (PETERSON et al., 1989, 2001).

Os achados experimentais obtidos por Lage (2005) em pesquisa sobre expressões idiomáticas verbais do PB como “chutar o balde” se mostram compatíveis com a ideia de que a estrutura sintática das expressões é de fato computada durante o processamento. Utilizando a técnica de extração de potenciais bioelétricos relacionados a eventos, Lage investigou a computação de sentenças contendo VPs idiomáticos e não idiomáticos (Ex. *O menino chutou o balde / O menino chutou a bola*) e concluiu que na concatenação do objeto ocorre um esforço cognitivo maior no caso das expressões idiomáticas. De fato, ancorada no arcabouço teórico da Morfologia Distribuída, a autora defende que, em decorrência das fases da derivação sintática, o significado literal seja acessado primeiro, precedendo o idiomático, que seria atingido após uma tarefa dupla.

### 3 Idiomaticidade, familiaridade e informação prévia na compreensão de expressões idiomáticas

Como vimos, embora tipicamente expressões idiomáticas sejam caracterizadas como cadeias de palavras cuja interpretação não se deriva de forma composicional, certas expressões admitem algum grau de decomposição já que é possível verificar nelas uma distância menor entre os sentidos literal e figurado. Por outro lado, muitas expressões – embora não necessariamente sejam decomponíveis – apresentam contrapartida literal, como ocorre por exemplo em expressões como “quebrar o gelo”, “chutar o balde” ou “pintar o sete”. Além disso, expressões diferem em termos de seu nível de cristalização e relativamente a outros aspectos como o

nível de familiaridade dos falantes e o contexto discursivo no qual aparecem inseridas, por exemplo.

Diante da diversidade de fatores levantados, optamos por explorar a possível influência de três desses aspectos na compreensão de expressões idiomáticas: segundo Giora (1997, 1999), Cacciari e Tabossi (1988), dentre outros, o grau de idiomaticidade – relacionado a uma maior ou menor distância entre os sentidos literal e figurado de uma expressão (cf. NUNBERG; SAG; WASOW, 1994) –, o grau de familiaridade do falante com a expressão – vinculado à experiência consciente do falante com a mesma – e a informação prévia disponível, que pode ou não favorecer uma interpretação idiomática da expressão. Para investigar a atuação desses fatores, elaboramos um experimento de leitura automonitorada por meio de uma tarefa de *maze*.

## 4 Método

Em função de estudos anteriores apresentados nas seções prévias do presente artigo (em particular, CACCIARI; TABOSSI, 1988; GIORA, 1997, 1999, 2003; GIORA; FEIN, 1999), nossa hipótese é que o processamento de expressões idiomáticas é influenciado por fatores diversos que envolvem tanto características intrínsecas da própria expressão (como o grau de idiomaticidade) quanto o nível de familiaridade do falante com a mesma e a informação previamente fornecida pelo contexto discursivo no qual a expressão aparece. O experimento que reportamos a seguir teve como objetivo geral investigar o processamento de expressões idiomáticas no PB e avaliar se os fatores antes mencionados podem favorecer ou dificultar seu processamento. Buscando testar a hipótese formulada, delimitamos as seguintes variáveis independentes:

- *Familiaridade* da expressão: comparando expressões familiares *versus* raras – variável *within-subjects*;
- *Idiomaticidade* da expressão: comparando expressões mais transparentes *versus* mais opacas – variável *within-subjects*;
- *Contexto sentencial*: comparando contextos prévios mais ou menos informativos quanto ao possível sentido idiomático da expressão; esta terceira variável foi um fator grupal.

Como variáveis dependentes selecionamos: o tempo de leitura/escolha do complemento do verbo núcleo da expressão e o número de respostas-alvo – equivalente ao número de vezes em que o participante optou pelo sintagma que completava a expressão idiomática.

As condições experimentais em função das duas primeiras variáveis independentes (*idiomaticidade* e *familiaridade*) foram as seguintes:

- Condição 1: expressões mais transparentes e familiares. Ex. *passar a bola; arregaçar as mangas*.
- Condição 2: expressões mais transparentes e raras. Ex. *esconder o leite; ensacar fumaça*.
- Condição 3: expressões opacas e familiares. Ex. *dar zebra; pintar o sete*.
- Condição 4: expressões opacas e raras. Ex. *enfeitar o pavão; sentar na graxa*.

Cada expressão foi inserida em uma sentença que fornecia um contexto mais favorável para a interpretação figurada e em uma outra sentença que fornecia um contexto mais “neutro”, sem pistas explícitas.

Foram formuladas as seguintes previsões:

- Expressões familiares em contextos mais informativos devem registrar os menores tempos de reação e um número significativamente maior de respostas-alvo quando comparadas com as restantes condições;
- Expressões opacas e pouco familiares em contextos menos informativos devem registrar os maiores TRs, quando comparadas com as restantes condições, além de um número significativamente menor de respostas-alvo (i.e. opção pela complementação metafórica);
- Caso uma maior decomponibilidade associada às expressões tenha um efeito “facilitador” no processamento (cf. GIBBS; GONZALES, 1985; GIBBS et al., 1989; GIBBS; NAYAK, 1989; PETERSON et al., 2001, dentre outros), espera-se que as expressões mais transparentes registrem TRs significativamente menores quando comparadas às expressões opacas;

#### 4.1 Participantes

Participaram do experimento 40 informantes divididos aleatoriamente em dois grupos em função do tipo de material apresentado: sentenças contendo expressões inseridas em contextos mais e menos informativos quanto ao possível valor idiomático das mesmas. Todos os participantes eram estudantes universitários de cursos diversos, com idades variando entre 17 e 36 anos (21,5 – idade média), dos quais 9 homens e 31 mulheres.

#### 4.2 Materiais e procedimento

Como passo prévio à elaboração das sentenças experimentais, selecionamos as expressões idiomáticas a serem utilizadas com base nos resultados de três atividades de pré-teste: uma tarefa de *cloze*, uma de paráfrase e um questionário via formulário do *Google*

no qual era solicitado para o participante informar se conhecia ou não as expressões e fornecer uma paráfrase aproximada das mesmas. Os voluntários que realizaram essas tarefas não participaram do experimento em si. As duas primeiras atividades foram realizadas por 29 participantes (20 mulheres e 9 homens; idade média de 24 anos), enquanto o formulário foi respondido por 202 voluntários (137 mulheres e 65 homens, idade média de 30 anos). A partir dos dados coletados nessas tarefas, selecionamos e classificamos as 24 expressões que foram utilizadas no experimento de leitura. Foram consideradas expressões mais familiares aquelas que registraram acima de 80% de respostas afirmativas nas tarefas de paráfrase e reconhecimento; foram consideradas expressões raras aquelas com uma taxa de reconhecimento inferior a 40%.

Assumindo que expressões idiomáticas podem ser classificadas num *continuum*, optamos por agrupar as expressões em dois grandes grupos: mais transparentes e mais opacas.

Cada uma das expressões selecionadas foi inserida em duas sentenças distintas: uma fornecia um contexto mais informativo quanto ao sentido idiomático, enquanto a outra fornecia um contexto mais neutro. Cada participante foi exposto apenas a um dos tipos de contexto:

- Contexto mais informativo: Depois de um mês de descanso, a Ana viu que tinha muito trabalho acumulado e percebeu que era hora de **arregaçar as mangas** novamente.
- Contexto menos informativo: Após beber um café na cozinha, a Ana decidiu que já era hora de **arregaçar as mangas** efetivamente.

Foram utilizadas 24 sentenças experimentais e 24 distratoras. O número de distratoras foi menor do que a proporção utilizada normalmente em tarefas de leitura convencionais pelas próprias características da tarefa de *maze*, na qual o material “extra” nas frases funciona também como distrator. Todas as sentenças foram divididas em oito fragmentos, sendo o sétimo segmento, o crítico, nas frases experimentais.

Foi utilizando o *software Linger*<sup>1</sup>, que permite gravar o tempo de reação e as escolhas do participante. Durante a execução da tarefa, as frases apareciam em segmentos formados por duas alternativas apresentadas lado a lado. Os participantes deveriam decidir qual das alternativas constituía uma continuação coerente da sentença e pressionar o botão do teclado correspondente à sua decisão (direita/esquerda).

<sup>1</sup> Disponível em: <<http://tedlab.mit.edu/~dr/Linger/>>.



**Figura 1.** Ilustração esquemática do procedimento experimental.

## 5 Resultados e discussão

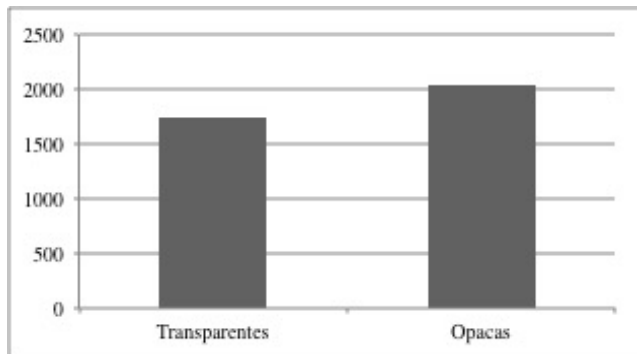
Os dados relativos à primeira variável dependente (tempo de reação – TR) foram submetidos a uma análise da variância (ANOVA) com *design* fatorial 2X2X2 (*idiomaticidade* × *familiaridade da expressão* × *contexto*), sendo a última variável um fator grupal e as restantes medidas repetidas. Essa análise revelou

um efeito principal de *idiomaticidade* ( $F(1,38)=11,6$   $p=0,001$ ) com tempos significativamente maiores para as expressões classificadas como mais opacas (Médias: 1748 ms para expressões mais transparentes e 2042,62 ms para as expressões mais opacas) (**Gráfico 1**).

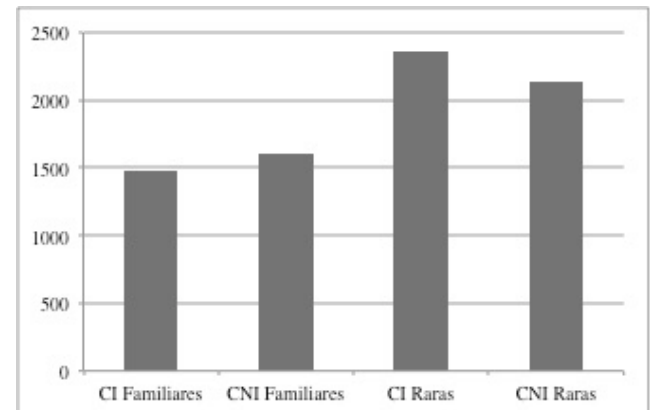
Ainda com relação aos TRs, foi obtido um efeito principal de *familiaridade* ( $F(1,38)=56,6$   $p=0,000001$ ) com médias significativamente maiores para as expressões raras (2248,95 ms) em relação às expressões familiares (1541,65 ms) (**Gráfico 2**).

O efeito da interação entre *familiaridade* e *contexto* aproximou-se do nível de significância ( $F(1,38)=3,56$   $p=0,06$ ). As médias foram: 1477,27 ms para expressões familiares em contextos informativos; 1606,03ms para expressões familiares em contextos não informativos; 2362 ms para expressões raras em contextos informativos e 2136 ms para expressões raras em contextos não informativos. Os maiores tempos de reação foram registrados na condição de contexto mais informativo associado a expressões raras.

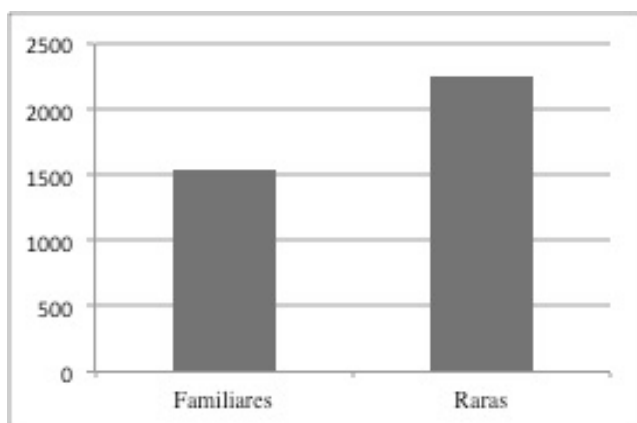
O **Gráfico 4** apresenta as médias de tempo de reação em milissegundos em todas as condições.



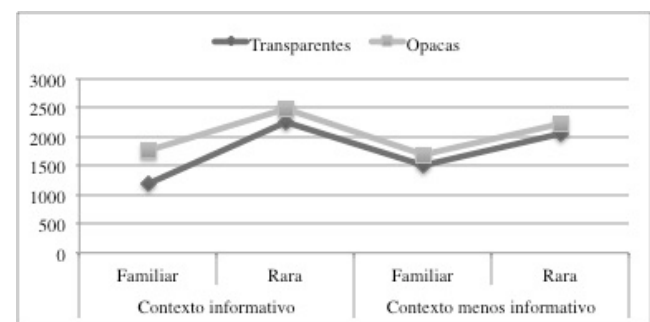
**Gráfico 1.** Tempo médio de reação em milissegundos em função de *idiomaticidade*.



**Gráfico 3.** Tempo médio de reação em milissegundos em função da interação entre *familiaridade* e *contexto*.



**Gráfico 2.** Tempo médio de reação em milissegundos em função de *familiaridade*.



**Gráfico 4.** Tempo médio de reação por condição experimental e por grupo.



Os dados relativos à segunda variável dependente (média de respostas-alvo) foram analisados a partir do Teste de Friedman e do Teste de Wilcoxon. A **Tabela 1** a seguir apresenta as médias de resposta por condição.

O teste de Friedman foi aplicado a cada um dos grupos individualmente e revelou um efeito significativo quando consideradas as variáveis analisadas (*idiomaticidade* e *familiaridade*). Os resultados do teste de Friedman são apresentados na **Tabela 2**.

O teste de Wilcoxon permitiu realizar uma comparação entre pares de condições no interior de cada grupo. Com relação aos resultados desse teste (cf. **Tabela 3**), destacamos o fato de, em todas as condições experimentais, ao compararmos ambos os grupos, os resultados terem se apresentado de forma semelhante no que diz respeito a serem, ou não, significativos. Essa mesma tendência só não foi observada no contraste entre expressões opacas raras e expressões transparentes raras (OR-TR, realçado na **Tabela 3**), em que, apenas no grupo que se deparou com contextos não informativos (G2), obtivemos uma diferença estatisticamente significativa. Esses resultados sugerem que, na ausência de um mínimo apoio contextual, a transparência tenha sido fator fundamental para a complementação idiomática das expressões quando as mesmas eram raras.

Nossos resultados reforçam a ideia – já aventada na literatura – de que uma maior composicionalidade das expressões pode facilitar seu processamento. As expressões transparentes, de forma geral, registraram TRs significativamente menores e mais escolhas do complemento alvo, independente do *contexto* prévio e do grau de *familiaridade* das mesmas.

No que diz respeito ao nível de *familiaridade* do falante com a expressão, nossos resultados indicam que expressões mais familiares foram processadas mais rapidamente e completadas pelo seu sentido figurado, na maioria das vezes. Tais resultados foram independentes do tipo de *contexto* e do grau de *idiomaticidade*.

Não foram registrados efeitos principais de *contexto*, i.e. quando considerado isoladamente, este não parece influenciar significativamente a compreensão das expressões. No entanto, quando associado com as outras variáveis, esse fator se mostrou relevante: expressões opacas receberam um número maior de respostas-alvo quando inseridas em contextos mais informativos. O tipo de contexto teve efeito no caso das expressões raras, mas não das mais familiares.

No que tange aos TRs, vimos ainda que o efeito de interação entre *familiaridade* e *contexto* ficou próximo do nível de significância. Uma ampliação da amostra original

**Tabela 1.** Médias de respostas-alvo por condição experimental.

Grau de Idiomaticidade	Contexto mais informativo		Contexto menos informativo	
	Familiar	Rara	Familiar	Rara
Transparente	5,7	2,7	5,25	2,6
Opaca	5,5	1,95	5,05	1,35

**Tabela 2.** Resultados da estatística a partir do teste de Friedman.

Variável Intergruppal (Tipo de contexto)	N	$\chi^2$	Graus de Liberdade	p-valor
Grupo 1 (Contexto mais informativo)	20	52,337	3	<0,001
Grupo 2 (Contexto menos informativo)	20	54,293	3	<0,001

**Tabela 3.** Resultados da estatística a partir do teste de Wilcoxon.

Condições Experimentais	G1		G2	
	Z	p-valor	Z	p-valor
TR-TF	-3,976 <sup>a</sup>	<0,001	-3,990 <sup>a</sup>	<0,001
OF-TF	-1,155 <sup>a</sup>	0,248	-1,155 <sup>a</sup>	0,248
OR-TF	-3,941 <sup>a</sup>	<0,001	-3,983 <sup>a</sup>	<0,001
OF-TR	-4,018 <sup>b</sup>	<0,001	-3,960 <sup>b</sup>	<0,001
OR-TR	-0,238 <sup>a</sup>	0,812	-3,054 <sup>a</sup>	0,002
OR-OF	-3,953 <sup>a</sup>	<0,001	-3,951 <sup>a</sup>	<0,001

seria importante para verificar se tal resultado se mantém e atinge o nível de significância estatística. Em linhas gerais, identificou-se um processamento mais rápido para expressões familiares em contextos informativos e um processamento mais lento para expressões raras, também em contextos mais informativos, o que pode sugerir a informação disponível, relacionada a expressões desconhecidas para o falante, pode ter gerado dúvidas na interpretação e, conseqüentemente, um aumento no tempo de processamento.

Quanto às repostas-alvo, observamos que expressões transparentes receberam um número semelhante de complementos idiomáticos, independente do tipo de *contexto*. Já, quando as expressões eram opacas, um contexto mais informativo favoreceu a complementação idiomática. Expressões opacas em contextos menos informativos registraram o menor número de repostas-alvo.

Ainda no que diz respeito às médias de resposta, tanto a transparência quanto familiaridade das expressões idiomáticas se revelaram fatores relevantes para facilitar o processamento idiomático das mesmas.

Em conjunto, os resultados obtidos são consistentes com as previsões levantadas e se mostram compatíveis com a hipótese inicial de que o processamento de expressões idiomáticas é influenciado por fatores diversos que envolvem tanto características intrínsecas da própria expressão (como o grau de *idiomaticidade*), quanto características extrínsecas das mesmas (como o nível de *familiaridade* do falante em relação à expressão), além das informações/pistas que o *contexto* sentencial prévio possa fornecer.

## Considerações finais

Os resultados reportados sugerem que expressões transparentes são processadas mais rapidamente do que as opacas, fato que se mostra compatível com a ideia de que as expressões não seriam tratadas como “blocos” ou “unidades prontas” durante a compreensão. Tais achados são consistentes com a proposta segundo a qual os itens que compõem a expressão seriam acessados no léxico de forma individual durante o processamento (GIBBS et al., 1989). Salientamos que nossos resultados não indicam necessariamente que a expressão em si seja primeiramente interpretada como literal e, apenas depois, como idiomática. A decomposição que parece ocorrer estaria vinculada ao acesso lexical dos itens individuais e sua posterior computação sintática e não a uma computação semântica da expressão completa propriamente dita.

Os modelos para o processamento de expressões que assumem uma visão “única” ou categórica – seja de acesso direto ou indireto ao sentido figurado –

parecem não dar conta do alto grau de variabilidade envolvido no processamento de tais estruturas. Nossos achados sugerem que os falantes tendem a realizar análises composicionais das cadeias idiomáticas durante a compreensão, nas quais as partes são tidas como independentes, mas não necessariamente como literais, e cujos significados contribuem para os sentidos figurados globais dessas expressões (GIBBS; GONZALES, 1985). Fatores intrínsecos à própria expressão (como o *grau de idiomaticidade*), externos (como o *contexto* prévio) e “mistos” como o nível de *familiaridade* (determinado pela experiência consciente do falante com a expressão em função da sua frequência de uso numa dada comunidade), parecem ser relevantes e precisam ser levados em consideração para uma melhor compreensão do modo como o processamento das mesmas ocorre.

## Referências

- BOBROW, Samuel A.; BELL, Susan M. On catching on to idiomatic expressions. *Memory and cognition*, n. 1, p. 343-346, mar. 1973.
- BÖRJESSON, Kristin. *The notions of literal and non-literal meaning in semantics and pragmatics*. 277f. Dissertação (Mestrado) – Faculdade de Filologia da Universidade de Leipzig, Alemanha. 2011.
- BOWDLE, Brian F.; GENTNER, Dedre. The career of metaphor. *Psychological review*, v. 112, n. 1, p. 193-216, fev. 2005.
- CACCIARI, Cristina; TABOSSI, Patrizia. The comprehension of idioms. *Journal of memory and language*, v. 27, p. 668-683, 1988.
- ESTILL, Robert B.; KEMPER, Susan. Interpreting idioms. *Journal of psycholinguistic research*, v. 11, p. 559-568, maio 1982.
- FRASER, Bruce. Idioms within a transformational grammar: *Foundations of language*, v. 6, p. 22-42, fev. 1970.
- GIBBS, Raymond. Skating on thin ice: literal meaning and understanding idioms in conversation. *Discourse Processes*, v. 9, p. 17-30, jan. 1986.
- GIBBS, Raymond. Spilling the beans on understanding and memory for idioms. *Context memory and cognition*, v. 8, p. 149-156, mar. 1980.
- GIBBS, Raymond. *The poetics of mind: figurative thought, language and understanding*. Cambridge: Cambridge University Press, 1994.
- GIBBS, Raymond; GONZALES, Gayle. Syntactic frozenness in processing and remembering idioms. *Cognition*, v. 20, p. 243-259, set. 1985.
- GIBBS, Raymond; NAYAK, Nandini P. Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive psychology*, v. 21, p. 100-138, jan. 1989.
- GIBBS, Raymond; NAYAK, Nandini P.; BOLTON, John L.; KEPPEL, Melissa E. Speakers' assumptions about the lexical flexibility of idioms. *Memory and cognition*, v. 17, p. 58-68, mar. 1989.

- GIBBS, Raymond; NAYAK, Nandini P.; CUTTING, Cooper. How to kick the bucket and not decompose: analyzability and idiom processing. *Journal of memory and language*, v. 28, p. 576-593, out. 1989.
- GILDEA, Patricia; GLUCKSBERG, Sam. On understanding metaphor: the role of context. *Journal of verbal learning and verbal behavior*, v. 22, p. 577-590, out. 1983.
- GIORA, Rachel. Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, v. 7, p. 183-206, 1997.
- GIORA, Rachel. On the priority of salient meanings: studies of literal and figurative language. *Journal of pragmatics*, v. 31, p. 919-929, jul. 1999.
- GIORA, Rachel. *On Our Mind. Salience, Context, and Figurative Language*. Oxford, New York: Oxford University Press, 2003.
- GIORA, Rachel; FEIN, Ofer. On understanding familiar and less familiar figurative language. *Journal of Pragmatics*, v. 12, p. 1601-1618, nov. 1999.
- GLUCKSBERG, Sam. The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, v. 7, n. 2, p. 92-96, mar. 2003.
- GLUCKSBERG, Sam; KEYSAR, Boaz. How metaphors work. In: ORTONY, Andrew (Ed.). *Metaphor and Thought*. Cambridge: Cambridge University Press, 1993. p. 401-424.
- GLUCKSBERG, Sam; NEWSOME, Mary R.; GOLDBERG, Yevgeniya. Inhibition of the Literal: Filtering Metaphor-Irrelevant Information During Metaphor Comprehension. *Metaphor and Symbol*, v. 16, n. 3-4, p. 277-298, jun. 2011.
- GRICE, Herbert P. Logic and conversation. In: COLE, Peter.; MORGAN, Jerry (Ed.). *Syntax and Semantics: speech*. Berkeley: Academic Press, 1975. p. 41-58.
- JANUS, Raizi A.; BEVER, Thomas G. Processing of metaphoric language: an investigation of the three-stage model of metaphor comprehension. *Journal of psycholinguistic research*, v. 14, n.5, p. 473-487, set. 1985.
- JOHNSON, Mark. *The body in the mind: the bodily basis of meaning, imagination and reason*. Chicago: University of Chicago Press, 1987.
- LAGE, Aléria. Aspectos neurofisiológicos de concatenação e Idiomaticidade em português do Brasil: um estudo de potenciais bioelétricos relacionados a eventos linguísticos (ERPs). Tese (Doutorado) – Programa de Pós-graduação em Linguística. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- MUELLER, Rachel A. G.; GIBBS, Raymond. Processing idioms with multiple meanings. *Journal of psycholinguistic research*, v. 16, n. 1, p. 63-81, jan. 1987.
- NUNBERG, Geoffrey. *The pragmatics of reference*. Indiana University Linguistics Club, 1978.
- NUNBERG, Geoffrey; SAG, Ivan; WASOW, Thomas. Idioms. *Language*, v. 70, n. 3, p. 491-538, 1994.
- PETERSON, Robert R.; BURGESS, Curt; DELL, Gary S.; EBERHARD, Kathleen. *Dissociation of syntactic and semantic analyses during idiom processing*. In: SECOND ANNUAL CUNY CONFERENCE ON HUMAN SENTENCE PROCESSING, New York, 1989.
- RICCI, Adiel Q. O processamento psicolinguístico da metáfora: um estudo experimental no PB. Dissertação (Mestrado em Linguística) – Universidade Federal Fluminense, Rio de Janeiro, 2016.
- SEARLE, John R. Metaphor. In: ORTONY, Andrew (Ed.). *Metaphor and Thought*. Cambridge: Cambridge University Press, 1979. p. 92-123.
- SWINNEY, David A.; CUTLER, Anne. The access and processing of idiomatic expressions. *Journal of verbal learning and verbal behavior*, v. 18, p. 523-534, out. 1979.

Recebido: 30/09/2017

Aprovado: 13/04/2018

**Contato:**

Sara de Oliveira Gomes Barreto <sarinhapsi@yahoo.com.br>