

# TOMARA-QUE-CAIA OU TOMARA QUE CAIA?

## UMA ANÁLISE CRÍTICA DO PROCESSO DE HIFENIZAÇÃO EM ITENS LEXICAIS COMPOSTOS À LUZ DO NOVO ACORDO ORTOGRÁFICO

**Miguel Pais Moreira Lopes** é mestrando em Estudos da Linguagem pela PUC-RJ e bolsista da CAPES, atuando na linha de pesquisa “Português brasileiro: descrições e aplicações específicas”. Está iniciando sua dissertação em estratégias e habilidades de escrita. **E-mail:** [mpaismlopes@yahoo.com](mailto:mpaismlopes@yahoo.com).

### RESUMO

O novo Acordo Ortográfico da Língua Portuguesa tem gerado muita polêmica quanto a se irá ou não facilitar o uso da língua escrita. O objetivo deste artigo é investigar uma diretriz proposta pelo novo Acordo Ortográfico no que tange ao emprego do hífen em palavras compostas e locuções. Temos como alvo determinar a eficácia da supressão deste diacrítico em palavras como *tomara que caia*, *bico de papagaio*, *dia a dia*, *maria vai com as outras* etc, bem como o impacto causado pela mudança no processo de leitura e percepção gráfica de itens lexicais.

### ABSTRACT

The new Orthographic Agreement of the Portuguese Language has generated much discussion on whether it will make it easier for people to use written language. The purpose of this article is to investigate a new rule in the Orthographic Agreement that has to do with the use of the hyphen signal in compounding. The aim is to determine to what extent the elimination of the hyphen in Portuguese words such as *tomara que caia*, *bico de papagaio*, *dia a dia*, *maria vai com as outras* etc will improve the use of language and what impact it will have in reading and in the graphic perception of words.

“Vou continuar a escrever como escrevo hoje.

[...] Os revisores encarregam-se disso”.

– José Saramago –  
(hipertexto, 2009)

## 1. Introdução

O novo Acordo Ortográfico da Língua Portuguesa tem gerado muita polêmica quanto a se irá ou não facilitar o uso da língua escrita. A frase de José Saramago contida na epígrafe parece ilustrar, de maneira sucinta, o sentimento de muitos quanto ao Acordo: um misto de apreensão e resistência quanto às novas regras e ao impacto que causarão.

Grande parte da comoção gerada pelo advento do novo Acordo, em vigor desde janeiro de 2009, tem a ver com uma visão equivocada sobre o que vem a ser *ortografia*. O conceito de *ortografia* é freqüentemente confundido por leigos com o conceito de *língua*. Essa confusão, não raro, tem gerado afirmações completamente infundadas, como a de que a queda do trema irá alterar a pronúncia das palavras ou de que modificações nas normas de escrita são uma afronta ao bem falar e acabam por conduzir a uma degradação do nosso idioma. Neste respeito, é importante ressaltar que

[...] a ortografia é uma norma, uma convenção social. Embora muitas vezes existam regras por trás da forma como se convencionou escrever as correspondências letra-som que usamos hoje, essas regras não deixam de ser convenções que, em sua gênese, não têm em si um sentido de obrigatoriedade, de necessidade. Tudo em ortografia é fruto de um acordo social, isto é, tudo foi arbitrado, mesmo quando existem regras que justificam por que em determinados casos temos que usar uma letra e não outra. (MORAIS, 2003, p.23)

Como forma cristalizada, ou produto de uma convenção social, a ortografia tem um propósito específico, a saber, promover um parâmetro gráfico que possa corresponder às diferentes variações lingüísticas, de modo que todos os falantes possam se entender no plano da escrita. Assim, desfaz-se por completo a idéia negativa que se cria em torno de uma norma ortográfica.

A língua, por sua vez, constitui um “sistema abstrato reconhecível nos muitos usos, orais ou escritos, que seus falantes fazem dela” (AZEREDO, 2008, p.63). Tendo por base essa definição, poderíamos nos perguntar: a escrita, com suas normas ortográficas e padrão de formas consagradas, constituiria um limite ao uso da língua? Ou seria um estímulo à expressão? A resposta a essa questão parece residir na forma como os falantes fazem uso das normas, tanto gramaticais quanto ortográficas. O domínio e a aplicação de regras de uso da língua é um fator importante para que um indivíduo possa expressar-se ou, a bem dizer, fazer-se entender, conforme coloca Azeredo (2008, p.63):

[...] o uso da língua é, em princípio, um **ato individual**. Mesmo individuais, porém, esses atos são normalmente acontecimentos intersubjetivos, visto que se realizam na e para a comunicação entre indivíduos ou sujeitos, que precisam, para compreender-se, estar ‘de acordo’ sobre o que significam os sinais que estão usando. (grifo do autor)

Sendo assim, parece ser razoável afirmar que um bom conhecimento das normas gramaticais, aliado ao conhecimento relativo a seu uso, pode contribuir para uma visão de escrita como porta de entrada para uma expressividade ilimitada e uma comunicação bem-sucedida. Por outro lado, dificuldades quanto ao emprego de normas, ou sua compreensão, poderiam resultar num fechamento por parte do indivíduo. A escrita passaria a ser vista como um emaranhado de imposições desconexas, o que contribuiria para a inibição da capacidade expressiva do falante. A ortografia, em todo o seu conjunto de preceitos, muitas vezes arbitrários e desprovidos de uma lógica unificadora, tem sido alvo dessas questões. Particularmente, o uso do hífen, com todas as suas exceções e diferentes critérios para emprego, tem sido um ponto alto de discussões, tanto por parte de gramáticos quanto de estudantes e professores.

O objetivo deste artigo é investigar uma diretriz proposta pelo novo Acordo Ortográfico no que tange ao emprego do hífen em palavras compostas e locuções. Temos como alvo determinar a eficácia da supressão deste diacrítico em palavras como *tomara que caía*, *bico de papagaio*, *dia a dia*, *Maria vai com as outras* etc, bem como fazer algumas conjeturas sobre possíveis impactos causados no processo de leitura e percepção gráfica de itens lexicais.

Este artigo está dividido em quatro seções: na primeira, fazemos um contraponto entre o uso do hífen antes do Acordo e as modificações relacionadas aos compostos e locuções, conforme estipuladas na Base XV. Tomamos por base a descrição de Azeredo (2004) e Bechara (2004) em confronto com a de Azeredo (2008) e a do Instituto Antônio Houaiss (2008). Em seguida, comparamos as modificações apresentadas com alguns pressupostos teóricos relativos ao conceito de palavra, conforme expostos em Basílio (2003), Kehdi (2002) e Monteiro (2002). A terceira parte procura



relacionar os conceitos expostos com uma avaliação crítica das mudanças, calcada nas observações de Chérng (2008) e Rego & Buarque (2009). A última parte destina-se às considerações finais.

## 2. O hífen em compostos antes e após o Acordo

A preocupação com uma unificação ortográfica não é algo novo. Cagliari (1999, p.71) mostra que as normas ortográficas começaram a surgir em função de uma resistência por parte da escrita em relação às mudanças da língua oral. A escrita do período arcaico da língua, que vai do século XIII ao XVI (INSTITUTO ANTÔNIO HOUAISS, 2008, p.18) era de caráter predominantemente fonográfico e havia um grande descompasso entre as contínuas alterações na língua falada e o registro ortográfico vigente. Foi somente no ano de 1904 que um compromisso maior com a unificação começou a tomar forma. A fonética ganhara importância nos trabalhos lingüísticos, sobretudo na decifração de escritas antigas. A obra de Gonçalves Vianna, intitulada *Ortografia nacional – simplificação e uniformização das ortografias portuguesas* teve grande repercussão tanto no Brasil como em Portugal, dando início a uma série de tentativas de unificação das normas ortográficas dos dois países (CAGLIARI, 1999, p.88). Tal processo é motivado não só por preocupações de caráter lingüístico, mas de natureza pedagógica e política (INSTITUTO ANTÔNIO HOUAISS, 2008, p.20).

O que se tem observado, no entanto, é o surgimento de algumas questões de ordem lingüística, relativas ao texto do Acordo. Machado (2009, p.13) ilustra bem esse ponto ao fazer referência à imprecisão com que é tratado o uso do hífen em compostos. Expressões vagas e passíveis de interpretação subjetiva como “certos compostos”, “em certa medida”, “noção de composição” etc. têm originado dúvidas e alguma confusão na compreensão do texto do Acordo. Esse processo pode gerar uma dificuldade ainda maior de aplicação das normas quanto ao uso do hífen, já consideradas problemáticas por muitos usuários da língua mesmo antes do advento do novo Acordo, conforme ilustrado pelo comentário dos autores do *Manual de Redação da Folha de São Paulo* (2001, p.132): “Há tantas exceções e casos especiais que a regra se torna pouco útil”.

Para compreender melhor a dificuldade a que Machado (2009, p.13) faz alusão, precisamos analisar a nova proposta de uso do hífen, mais especificamente no caso de itens lexicais compostos. Bechara (2004, p.99), anterior à vigência do Acordo, descreve a regra geral para emprego do hífen em compostos da seguinte maneira:

Só se ligam por hífen os elementos das palavras compostas em que se mantém a noção da composição, isto é, os elementos das palavras compostas que mantêm a sua independência fonética, conservando cada um a sua própria acentuação, porém **formando o conjunto perfeita unidade de sentido**. (grifo nosso)

É interessante notar que, logo após essa regra geral, Bechara (2004, p.100) especifica que as palavras compostas formadas por elementos que, isoladamente, perdem sua significação, mas que, juntos, constituem um bloco semântico estável, são marcados com o diacrítico. Como exemplo, o

autor cita as palavras *água-marinha*, *arco-íris*, *galinha-d'água*, *couve-flor*, *guarda-pó*, *pé-de-meia*, *pára-choque*, *porta-chapéu* etc. Cabe aqui ressaltar que cada um desses compostos é formado por combinações de elementos pertencentes às mais variadas classes gramaticais, como *água-marinha* (substantivo + adjetivo), *pé-de-meia* (substantivo + preposição + substantivo) e *guarda-pó* (verbo + substantivo). Em seguida, faz algumas ressalvas, dentre as quais o fato de as locuções, de qualquer categoria gramatical, não levarem hífen, pelo fato de não manterem uma unidade de sentido entre seus elementos.

Comparemos, agora, essa definição com o conceito gramatical de composição elaborado por Azeredo (2004, p.98) antes do advento do Acordo: “Chama-se *composição a união de dois ou mais lexemas para a criação de uma nova unidade fixa*: bomba-relógio, guarda-roupa, pé-de-vento, corre-corre, azul-piscina [...]” (grifo do autor). Convém lembrar que a noção de *lexema* diz respeito às “unidades de base do léxico e pertencem a inventários ilimitados e abertos, uma vez que novos radicais podem ser criados” (MONTEIRO, 2002, p.18).

Percebemos que tanto Azeredo (2004, p.98) quanto Bechara (2004, p.100) demonstram que o hífen é efetivamente empregado em compostos e que um dos critérios para determinação dessa classe lexical é a constituição de um bloco semântico sólido formado pela união justaposta de lexemas. Esses lexemas, por sua vez, combinam sua significação própria, individual, para formarem um novo item lexical, cujo significado vai além da mera soma do valor semântico de cada lexema considerado separadamente no contexto da união. Em outras palavras, poderíamos afirmar que o item lexical *criado-mudo* é um composto, pois designa um conceito particular (móvel de madeira comumente encontrado nas laterais de uma cama), e sua significação *vai além* da mera união dos significados particulares de *criado* (serviçal) e *mudo* (desprovido de fala), levando, portanto, hífen. Por outro lado, a formação *fim de semana* não poderia ser considerada um composto, já que sua significação enquanto bloco (a parte final de uma sequência de sete dias) se atém à simples união dos significados de cada vocábulo componente.

Vejamos, por fim, o tratamento dispensado aos compostos no que tange ao uso do hífen segundo a nova ortografia. A Base XV do Acordo Ortográfico (INSTITUTO ANTÔNIO HOUAISS, 2008, p.42-3) estipula que o hífen deve ser empregado em:

- a) Palavras compostas por justaposição cujos elementos (substantivos, adjetivos, numerais ou verbos) constituem uma nova unidade morfológica e de sentido, mantendo o acento próprio, mesmo que o primeiro elemento esteja reduzido;
- b) Topônimos iniciados por *grão* ou *grã*, por verbo ou se houver artigo entre seus elementos;
- c) Palavras compostas que designem espécies botânicas ou zoológicas;
- d) Palavras cujo primeiro elemento for *bem* ou *mal* e o segundo elemento começar por vogal ou *h*;

e) Palavras compostas cujo primeiro elemento seja *além, aquém, recém, sem*;

f) Encadeamentos vocabulares ocasionais e combinações históricas ou topônimos ocasionais.

As únicas ressalvas feitas (que proíbem o uso do hífen) dizem respeito às locuções de quaisquer categorias, as quais o Acordo identifica como substantivas, adjetivas, pronominais, adverbiais, prepositivas e conjuncionais, e aos “compostos em que se perdeu a noção de composição”. (INSTITUTO ANTÔNIO HOUAISS, 2008, p.43)

A problematização da mudança no uso do hífen reside justamente na distinção feita entre o que é classificado como *palavra composta* e como *locução*. O Acordo deixa dúvida quando menciona que “fazem exceção a esta regra palavras cuja grafia está consagrada pelo uso”, citando exemplos como *cor-de-rosa* e *mais-que-perfeito*. Seriam essas as únicas locuções em que se admite hífen? Qual o critério para determinar o fator *consagrado pelo uso*, algo que nos parece essencialmente subjetivo?

É interessante notar que a descrição dada por Bechara (2004, p.100) identifica apenas quatro tipos de locução: pronominal, adverbial, prepositiva e conjuntiva, ao passo que o Acordo registra duas mais: substantiva e adjetiva, sendo que a primeira (substantiva) sequer é mencionada em Bechara (2004). Além disso, em sua *Gramática Houaiss*, Azeredo (2008, p.444) registra o item lexical *pé de vento*, outrora grafado com hífen em Azeredo (2004, p.98), agora sem o sinal diacrítico. Teria essa mudança de grafia sido motivada pelo texto do Acordo especificar que os elementos constituintes de um item lexical composto são apenas “substantivo, adjetivo, numeral e verbo” (INSTITUTO ANTÔNIO HOUAISS, 2008, p.42)? Ou seria ela motivada por uma classificação de *pé de vento* como locução substantiva? Essa segunda hipótese seria impossível, já que o item *pé de vento* é considerado por Azeredo (2008, p.444) como exemplo de palavra composta, além do fato de a diretriz com respeito ao tratamento de locuções não constituir uma inovação do Acordo.

Chegamos, assim, ao cerne da questão: a fronteira tênue que separa locuções e palavras compostas. Parece-nos que o grande problema reside no fato de o Acordo registrar uma nova categoria de locuções: as substantivas. É interessante notar que a Nomenclatura Gramatical Brasileira, de acordo com a Portaria nº 36, de 28 de janeiro de 1959, não prevê a classificação *locução substantiva*. Também é digno de nota que essa especificação não é listada entre as classes de locuções em nenhuma das gramáticas atuais, conforme se pode constatar em Azeredo (2004, pp.271-2), Azeredo (2008, p.577), Bechara (2004, p.664) e Cunha & Cintra (2008, p.756). Todos esses autores reconhecem a existência de locuções adjetivas, adverbiais, conjuntivas, interjetivas, preposicionais e verbais, mas não a de locuções substantivas. A criação de uma nova categoria de locuções parece criar um problema para o processo de hifenização, na medida em que arrola palavras comumente hifenizadas, como *pé de moleque* e *maria vai com as outras*, outrora reconhecidas intuitivamente como palavras compostas, dotadas de uma unidade semântica, sob uma categoria não prevista nos principais compêndios de descrição gramatical da atualidade. Essa questão encontra profunda repercussão no texto do Acordo e é de grande relevância, na medida em que causa um impacto na forma



como se concebe o uso do hífen. Para compreender o problema na íntegra, devemos nos reportar ao campo de discussões em Morfologia sobre a distinção entre compostos e locuções.

### 3. Confronto com alguns pressupostos teóricos

Vejamos, primeiramente, algumas definições relacionadas à noção de composição. Basílio (2000, p.14) define o processo de composição como “uma formação em que temos duas ou mais palavras ou radicais”. Já Monteiro (2002, p.183) classifica o composto como “o vocábulo formado pela união de dois ou mais semantemas”. O problema básico, conforme ambos os autores salientam, é distinguir a palavra composta de uma locução. Monteiro (2002, p. 185) fornece a seguinte solução:

Parece-nos que grande parte do problema resulta de se interpretar a composição como um mecanismo morfológico. Na realidade, na maioria das situações, tem-se um processo de natureza sintático-semântica. Nesse sentido, Ortega (1990) afirma que a composição se baseia em critérios sintáticos, havendo a suposição de que certos vocábulos têm uma estrutura interna que reflete a estrutura interna das orações e sintagmas.

Observemos que na formação de **amor-perfeito** houve a combinação de um substantivo e um adjetivo no plano sintático, a julgar pela manutenção do processo de concordância nominal. Mas essa articulação gera uma especificação semântica, uma unidade semântica. Por isso, até construções oracionais podem expressar uma unidade de significado, como em: **um maria-vai-com-as-outras, o mama-na-égua, o não-te-esqueças-de-mim** (miosótis). (grifo do autor)

Temos, portanto, como condição básica para a caracterização do composto a relação sintático-semântica que é estabelecida entre os elementos do bloco lexical. Basílio (2000, p.14), por outro lado, problematiza formações como *maria vai com as outras* e *sempre viva* pelo fato de “não corresponderem a um processo lexical produtivo de composição”. A autora faz uma distinção entre composição como “processo sistemático de expansão lexical que consiste na combinação semântica de duas palavras inseridas num modelo estrutural pré-determinado [...]” (BASÍLIO, 2000, p.14), o que resultaria na designação de seres, eventos ou propriedades, como em *guarda-roupa, sofá-cama, olho-de-sogra* e *luso-brasileiro*, e uma outra dimensão da noção de composição, que “corresponde à eventual lexicalização de conjuntos de palavras, por motivos históricos de diferentes ordens” (BASÍLIO, 2000, p. 14), definição que teria como exemplo os termos *sempre viva* e *maria vai com as outras*.

A partir dessas colocações, tanto Basílio (2000, p.14-7) quanto Monteiro (2002, p.183-5) determinam alguns critérios para a identificação de compostos. São eles: (a) impossibilidade de permutar a ordem dos elementos do composto; (b) impossibilidade de omitir um dos elementos; (c) impossibilidade de inserir um terceiro elemento em qualquer parte do bloco semântico. Para ilustrar, tomemos como exemplo o item lexical *guarda-chuva*. Se tentarmos infringir cada um dos três princípios expostos acima, obteremos frases impossíveis em língua portuguesa, conforme a seguir:

- (1) \*O meu *chuva-guarda* é azul.  
(2) \*Começou a chover e, na hora de abrir meu *guarda*, ele emperrou.  
(3) \*Ontem comprei um *guarda-pequeno-chuva*.

Comprovamos assim, por meio desses três testes de caráter sintático-semântico que *guarda-chuva* constitui uma palavra composta. E de fato o é, constando no dicionário como tal. Entretanto, esse processo nem sempre é tão simples, conforme aponta Kehdi (2002, p.12):

Assinalemos, contudo, que, se não podemos estabelecer correspondência automática entre palavra e lexia simples (...), o teste de não-separabilidade **permite-nos levantar um grande número de compostos não-dicionarizados**. Sirva-nos de exemplo a expressão *casa de detenção*. Qualquer adjetivo a ela acrescentado não pode ocupar posição interna relativamente ao grupo. (grifo nosso)

Kehdi (2002, p.12) reconhece *casa de detenção* como uma palavra composta. De fato, se aplicarmos os testes mencionados, chegaremos a essa definição, a qual poderia ser estendida para muitos outros blocos de vocábulos que não constam no dicionário. Entretanto, poderia prevalecer o argumento de que o bloco semântico, neste caso, não vai além da significação de cada elemento, já que *casa de detenção* resguarda, de fato, o sentido de “local onde pessoas ficam detidas”. Assim, temos um conflito entre os testes sintático-semânticos e a unidade lexical do termo em questão, o que poria em dúvida sua classificação como palavra composta e, conseqüentemente, sua hifenização. Afinal, pela nova regra, se *casa de detenção* fosse considerado substantivo composto, deveria receber hífen.

Um outro exemplo tem que ver com a palavra *pé de vento*, mencionada em Azeredo (2008, p.444). De acordo com a nova ortografia, o item em questão não levaria hífen, o que nos permitiria concluir tratar-se de uma locução. Contudo, aplicando os referidos testes sintático-semânticos, poderíamos argumentar que *pé de vento* apresenta características de um substantivo composto (seria impossível conceber formações como \**pé vento de*; \**pé forte de vento*; e \**pé vento* funcionando em um enunciado). Uma avaliação do item como bloco semântico também poderia induzir um falante a considerar a palavra como indo além da união dos significados individuais de seus elementos (afinal, parece ser improvável que o bloco *pé de vento* possa ser concebido como uma mera união de *pé* e *vento* como significações isoladas).

#### 4. Uma análise crítica do uso do hífen pós-Acordo

Consideremos as seguintes frases, já escritas em conformidade com as novas regras ortográficas:

- (4) Minha irmã pegou um bico de papagaio...  
(5) O bico-de-papagaio é uma espécie muito pesquisada pelos botânicos.  
(6) Você pode me emprestar seu tomara que caia vermelho?

- (7) Meu dia a dia é uma correria só.  
(8) Ele esperava pacientemente, dia a dia, a vinda de sua amada.  
(9) Se eu fosse você, deixaria de ser um maria vai com as outras.  
(10) Se ela vir você assim, vai virar uma fera, e aí vai ser um deus nos acuda.  
(11) A meteorologia anunciou a chegada de um pé de vento nesta região.  
(12) Levantou a cabeça e viu um letreiro onde se lia “Casa de detenção”.  
(13) Estava à-toa na vida.  
(14) Esse menino reclama à toa.

Todos esses enunciados, completamente possíveis em nossa língua, são compostos por blocos de itens lexicais. As questões que se instauram são: devem ser considerados palavras compostas ou locuções? Partindo dessa classificação, devem ou não levar hífen? E que efeito a queda do diacrítico em itens lexicais outrora hifenizados pode produzir nos falantes?

Conforme já verificamos na seção anterior, um dos problemas é que a dificuldade de se estabelecer uma distinção clara entre os conceitos de palavra composta e locução torna-se um obstáculo à intuição do falante quanto ao emprego da regra de uso do hífen. Para usar o sinal diacrítico corretamente segundo a nova regra, faz-se necessária uma reflexão de ordem gramatical quanto à classificação do termo em questão, o que nem sempre é claro, conforme vimos. Além disso, o registro de uma nova categoria classificatória para as locuções, as substantivas, não prevista nas gramáticas tradicionais, obscurece ainda mais a lógica de emprego do hífen em itens lexicais definidos como compostos por alguns autores, como é o caso de *pé de vento* (AZEREDO, 2008, p. 444) e *maria vai com as outras* (BASÍLIO, 2000, p. 14).

Um segundo problema que podemos apontar tem relação com a eliminação da distinção semântica entre locuções com valores gramaticais distintos, como *dia a dia*, citado nas frases (7) e (8) e *à toa*, citado em (13) e (14). No caso de *dia a dia*, tínhamos, antes do acordo, uma marca gráfica clara que separava *dia-a-dia* (locução substantiva) de *dia a dia* (locução adverbial), ao passo que em *à toa*, podíamos claramente diferenciar a locução adjetiva (*à-toa*) da locução adverbial (*à toa*). Com a supressão do hífen em todas as locuções, dependeremos exclusivamente do contexto e de uma reflexão sintática para determinar a condição de tais termos. Poderíamos ainda estender essa problemática para termos como *bico de papagaio*. Pelo novo Acordo, o termo só levará hífen se corresponder à espécie botânica. A forma que corresponde à deformação na coluna perde o sinal diacrítico, o que impede que, isoladamente, possa ser diferenciada do enunciado descritivo referente ao conceito “parte integrante da constituição física de um papagaio”. Novamente, o contexto deve ser utilizado para realizar a diferenciação.

Um terceiro problema que gostaríamos de levantar tem que ver com o papel do hífen como marca gráfica de união lexical. Estudos recentes sobre aspectos lingüísticos e psicolingüísticos envolvidos no processo de aquisição da linguagem ressaltam a importância do reconhecimento gráfico da palavra como unidade lexical no processo de leitura. As técnicas e estratégias específicas envolvidas nesse processo têm que ver com o domínio de elementos fonéticos e estruturais das palavras, dentre os quais podemos



citar as regras de acentuação, silabação e aquisição de um vocabulário visual amplo. Reconhecer uma palavra em sua forma escrita ou impressa envolve duas habilidades: a) determinar sua pronúncia e b) estabelecer algum significado. O reconhecimento de um item lexical ocorre mediante uma análise fônica e estrutural do enunciado e um dos elementos envolvidos nessa análise é o reconhecimento de grupamentos hifenizados, como *guarda-chuva*. Em outras palavras, o hífen enquanto marca gráfica em *guarda-chuva* seria um dos elementos responsáveis pela decodificação do grupamento como um item lexical.

Transpondo essas observações para o âmbito do novo Acordo, poderíamos nos perguntar: a supressão de hífen nas chamadas locuções substantivas como *maria vai com as outras*, *tomara que caia*, *deus nos acuda*, *mama na égua* etc. poderia contribuir para uma maior dificuldade por parte do leitor no reconhecimento de tais itens como unidades lexicais? Sendo o hífen uma marca gráfica vívida da união dos elementos, sua ausência poderia tornar difícil a visualização do leitor de tais grupamentos como unidades sólidas, o que poderia resultar em problemas de compreensão e fluência na leitura de um texto.

Essa afirmação pode ser comprovada com base no estudo conduzido por Cherng (2008). Melanie Cherng se propôs a investigar o papel da hifenização no processamento de itens lexicais compostos do inglês (CHERNG, 2008, p.2). Segundo ela, informações coletadas sobre a velocidade de processamento de itens lexicais podem auxiliar no aprimoramento da alfabetização, pois contribuem para uma leitura mais eficiente e fornecem dados vitais para educadores que trabalham com dificuldades de leitura. A pesquisadora tomou como base estruturas que, no inglês, podem aparecer com ou sem o sinal diacrítico, como *grown-up* e *grown up*. Metade dos compostos selecionados para o teste lingüístico aparecia mais freqüentemente em sua forma hifenizada, ao passo que, com a outra metade, ocorria o oposto. As palavras foram inseridas em frases, tanto na forma hifenizada quanto na forma livre. Os participantes foram orientados a ler um conjunto de frases selecionadas aleatoriamente enquanto uma máquina registrava seus movimentos oculares. Os resultados mostraram que a supressão do hífen em determinadas formações acarretou um tempo de processamento maior, indicando a importância do diacrítico enquanto marca gráfica responsável pela visualização da unidade lexical. Nas palavras da autora<sup>1</sup>:

Furthermore, though the hyphen represents spatial segmentation in some aspects similar to a space, it still functions as a symbolic connector between lexemes both orthographically and semantically. Deleting a space significantly benefited word processing for spaced compound words in Juhasz et al. (2005) because it visually unified two related lexemes and thus indicated a unit of meaning to the reader. Deleting a hyphen was not significantly beneficial in the present study perhaps because a hyphen already signified to the reader the connection between the two lexemes. Furthermore, a hyphenated compound word is more unified than a space compound word. (CHERNG, 2008, pp.18-9)

Para que esse fato possa ser comprovado em compostos da língua portuguesa, faz-se necessário conduzir um experimento similar, possivelmente no campo da Psicolingüística, em seleção lexical. Um exemplo desse tipo de pesquisa está em Rego & Buarque (2009):

O presente estudo se propôs a investigar a influência da consciência sintática e da consciência fonológica na aquisição de regras ortográficas de naturezas distintas. Os resultados obtidos com crianças falantes do português, expostas a uma metodologia tradicional de ensino da ortografia, deram suporte à hipótese levantada por Nunes et al. (1994, 1996). Ficou evidenciado, através do dado longitudinal, que a consciência sintática é, possivelmente, um fator importante na aquisição de aspectos da ortografia que envolvem o conhecimento da classe gramatical a que uma palavra pertence. Este conhecimento gramatical, porém, não necessita ser explícito, isto é, constituído de habilidades prévias da criança para classificar palavras de acordo com as respectivas classes gramaticais. Os nossos resultados, no entanto, sugerem que este tipo de conhecimento mais explícito de gramática, embora menos eficaz para prever o progresso em ortografia a longo prazo, estaria associado a esta aquisição.

Encontramos respaldo para essas colocações também em Gomes (2008, p.63-69): “Se nos perspectivarmos pedagogicamente, notaremos que o uso dos sinais gráficos ajuda indiscutivelmente a leitura, especialmente nos primeiros anos de aprendizado da língua materna, ou no ensino de língua a estrangeiros”.

## 5. Considerações finais

Tendo analisado as principais diretrizes impostas pelo novo Acordo no tocante ao uso do hífen em compostos e locuções, podemos levantar algumas questões de cunho lingüístico. Essas questões dizem respeito às dificuldades que um falante pode encontrar no emprego do sinal diacrítico em função de cinco pontos fundamentais:

- (a) O tratamento impreciso dado pelo Acordo a compostos;
- (b) A dificuldade de se estabelecer uma distinção clara entre o que vem a ser uma palavra composta e uma locução;
- (c) O advento de uma categoria descritiva de locuções (substantivas) não prevista na NGB e nas principais obras de referência;
- (d) A eliminação da distinção gráfica em locuções com valores gramaticais diferentes;
- (e) O papel do hífen enquanto marca gráfica de união lexical.

Uma investigação ainda mais aprofundada dessas questões faz-se necessária, se havemos de determinar até que ponto elas poderão colocar em risco a compreensão e o uso do sinal diacrítico em questão. Existe ainda um longo caminho a ser percorrido, conforme aponta Azeredo, em entrevista a Josué Machado, da *Revista Língua*: “Dúvidas surgirão mesmo após o novo vocabulário ortográfico, pois a língua se expande, novas construções se cristalizam com o uso, e as pessoas se perguntarão a cada vez: leva ou não hífen?” (MACHADO, 2009, p.13).

## Notas

1. *Tradução livre da citação*: Além disso, embora o hífen represente uma segmentação espacial em alguns aspectos similar a um espaço, ele ainda funciona como um conector simbólico entre lexemas, tanto ortográfica quanto semanticamente. A supressão de um espaço beneficiou significativamente o processamento de itens lexicais em palavras compostas com espaço em Juhasz et al. (2005), pois provocou uma unificação visual de dois lexemas relacionados, indicando, assim, uma unidade de significado para o leitor. A exclusão do hífen não foi significativamente benéfica no presente estudo talvez pelo fato de o hífen já sinalizar para o leitor a conexão entre os dois lexemas. Além disso, uma palavra composta hifenizada é mais unificada se comparada a uma palavra composta sem hífen.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AZEREDO, J. C. **Fundamentos de Gramática do Português**. Rio de Janeiro: Jorge Zahar Editor, 2004. pp. 98-99.
- \_\_\_\_\_. **Gramática Houaiss**. São Paulo: Publifolha, 2008. pp. 444-5.
- BASÍLIO, M. M. P. "Em torno da palavra como unidade lexical: Palavras e composições". Extraído de: <<http://www.revistaveredas.ufjf.br/volumes/7/artigo1.pdf>>. Acessado em: 09/04/09. pp. 14-7.
- BECHARA, E. **Moderna Gramática Portuguesa**. Rio de Janeiro: Lucerna, 2004. pp. 99-100.
- CAGLIARI, L. C. "A Ortografia na Escola e na Vida". In: MASSINI-CAGLIARI, G. & CAGLIARI, L. C. **Diante das Letras: A Escrita na Alfabetização**. São Paulo: Fapesp, 1999. pp. 61-96.
- CHERNG, M. H. "The Role of Hyphenation in English Compound Word Processing". Extraído de: <[http://wescholar.wesleyan.edu/etd\\_hon\\_theses/199/](http://wescholar.wesleyan.edu/etd_hon_theses/199/)>. Acessado em: 15/04/09. pp. 2-18.
- CUNHA, C. & CINTRA, L. **Nova Gramática do Português Contemporâneo**. Rio de Janeiro: Lexikon, 2008. p. 756.
- GOMES, F. A. **O Acordo Ortográfico**. Porto, Portugal: Porto Editora, 2008. pp. 63-9.
- INSTITUTO ANTÔNIO HOUAISS. **Escrevendo pela nova ortografia: como usar as regras do novo acordo ortográfico da língua portuguesa**. São Paulo: Publifolha, 2008.
- KEHDI, V. **Morfemas do Português**. São Paulo: Ática, 2002. pp. 10-12.
- MACHADO, J. "A gramática como meio". In: **Revista Língua**. Ano III, Número 41, Março de 2009. pp. 12-3.
- FOLHA DE SÃO PAULO. **Manual de Redação**. São Paulo: Publifolha, 2001. p. 132.
- MONTEIRO, J. L. **Morfologia Portuguesa**. Campinas: Pontes, 2002. pp. 18, 183-5.
- MORAIS, A. G. "Ortografia: o que é? para que serve? por que ensiná-la?". In: **Ortografia: ensinar e aprender**. São Paulo: Ática, 2003. pp. 17-25.
- REGO, L. L. B. & BUARQUE, L. L. "Consciência sintática, consciência fonológica e aquisição de regras ortográficas". Extraído de: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-79721997000200003](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-79721997000200003)>. Acessado em: 23/03/2009.
- "Saramago e o Acordo Ortográfico". Extraído de: <<http://ciberduvidas.sapo.pt/controversias.php?rid=1793>>. Acessado em: 10/04/09.