

DECOMPOSIÇÃO E RECOMPOSIÇÃO NO PROCESSAMENTO LEXICAL: UMA REVISÃO DOS ESTÁGIOS DO RECONHECIMENTO VISUAL DE PALAVRAS COMPLEXAS¹

Daniela Cid de GARCIA

Universidade Federal do Rio de Janeiro (UFRJ)

RESUMO

Estudos sobre o reconhecimento visual de palavras têm mostrado resultados convergentes para um estágio de decomposição inicial equivalente para palavras transparentes (SAILBOAT) e opacas (BOOTLEG). Considerando que tanto palavras transparentes quanto opacas são decompostas, deve haver um estágio posterior de recomposição que daria conta da diferença entre os mecanismos de ativação do sentido de SAILBOAT e de BOOTLEG. Este estudo tem como objetivo apresentar o estado da arte com relação ao processamento de palavras, destacando os estágios até agora identificados que precedem e que seguem o acesso lexical e apontando o papel crucial dos morfemas nesse processo.

ABSTRACT

Studies of word recognition show convergent evidence for an indistinguishable initial decomposition stage for transparent (SAILBOAT) and opaque (BOOTLEG) words, as predicted by full-decomposition models. Considering that both transparent and opaque words are decomposed, there should be a later stage of recomposition that would account for the difference between the meaning retrieval mechanisms of SAILBOAT and BOOTLEG. This study aims at presenting the state of the art concerning word processing, highlighting the stages that precede and follow lexical access, as well as the crucial role of morphemes in the process.

¹ Este artigo resume uma parte da tese de doutorado da autora (Programa de Pós-graduação em Linguística da UFRJ/apoio CAPES e Comissão *Fulbright*, sob orientação de Anieli França e Marcus Maia). O experimento reportado (GARCIA, 2013) foi realizado no Laboratório de Neurolinguística da Universidade de Nova York, em parceria com Teon Brooks, e teve a supervisão de Liina Pykkänen e Alec Marantz.

PALAVRAS-CHAVE

morfologia, composicionalidade semântica, acesso lexical, MEG

KEYWORDS

morphology, semantic compositionality, lexical access, MEG

O papel da morfologia no processamento de palavras

Entender o papel dos morfemas na arquitetura e no processamento da linguagem é um desafio relevante tanto para a linguística teórica quanto para a Psicolinguística. No escopo da teoria da Gramática Gerativa, ela se traduz primeiro em definir quais são os átomos memorizados que entram na numeração para então serem operados pela sintaxe. Para suas vertentes lexicalistas – que incluem a versão mais recente da Gramática Gerativa, o Minimalismo (CHOMSKY, 1995) –, as palavras estão armazenadas em sua forma plena no léxico e a computação se dá, portanto, no nível supraléxico. Há, por outro lado, propostas não-lexicalistas, que argumentam que a noção intuitiva de palavra não corresponde necessariamente às unidades mentais das operações linguísticas, uma vez que não parece se justificar um domínio especial de associação entre som e sentido no nível lexical (MARANTZ, 1997a; 1997b). A proposta não-lexicalista da Morfologia Distribuída (HALLE & MARANTZ, 1993; BORER, 2000; HARLEY & NOYER, 1998a; 1998b) prevê que a computação ocorra em múltiplas fases, distribuída em diferentes camadas funcionais. A sintaxe operaria, portanto, também no interior da palavra, fazendo a concatenação entre raiz e morfemas categorizadores.

Dentro do escopo da Psicolinguística, estudar essas questões envolve também compreender as etapas perceptuais envolvidas no reconhecimento lexical, que devem culminar na seleção da representação

relevante para a computação linguística. Essas etapas devem variar dependendo da modalidade – auditiva ou visual. Em estímulos auditivos, por exemplo, é natural que haja uma ativação incremental à medida em que os fones vão sendo percebidos no tempo. Sendo assim, deve existir um momento de composição fonética na etapa inicial do reconhecimento de uma palavra, de modo que ele envolva a ativação de várias outras palavras a partir da semelhança fonológica com cada segmento desvelado no curso da fala (cf. Modelo *Cohort* – MARSLEN-WILSON & WELSH, 1978). À medida que cada novo som da palavra é percebido, algumas palavras são desativadas e outras, com semelhança pelo meio, são ativadas. Estímulos visuais, por sua vez, devem prescindir dessa primeira etapa composicional, considerando-se que a leitura se dê por reconhecimento paralelo das letras (DAHEANE, 2009; RAYNER & POLLATSEK, 1987).

Após essa ativação inicial de nível mais baixo, a informação perceptual deve ser segmentada nas partes que servirão como as unidades mínimas da computação linguística. Determinar o tamanho e a natureza dessas unidades implica entender como o léxico mental é organizado. Estudos em processamento da linguagem têm se dedicado há pelo menos quatro décadas à investigação do papel da morfologia no reconhecimento de palavras. Inaugurados por TAFT & FORSTER (1975), esses estudos podem ser organizados em grupos distintos, de acordo com o que cada um deles postula com relação à decomponibilidade morfológica. Os modelos decomposicionais defendem que a palavra deve ser armazenada e acessada por meio de seus constituintes mínimos, os morfemas (TAFT & FORSTER, 1975; STOCKALL & MARANTZ, 2006). Por outro lado, modelos não-decomposicionais argumentam que a palavra esteja representada e seja acessada inteira (SEIDENBERG & GONNERMAN, 2000).

Modelos decomposicionais tardios postulam que a decomposição de palavras em seus constituintes pode ocorrer após o acesso lexical

à palavra inteira (GIRAUDO & GRAINER, 2000; MARINKOVIC, 2004). Modelos decomposicionais iniciais, por outro lado, defendem que os morfemas são acessados automaticamente logo no início do processamento da palavra complexa. Os modelos de decomposição inicial podem ainda ser divididos entre os que preveem um mecanismo de dupla rota (MARSLEN-WILSON et al., 1994; SHREUDER & BAAYEN, 1995; CARAMAZZA et al., 1998; PINKER, 2000; HAY & BAAYEN, 2005) e os que defendem a decomposição plena da palavra em morfemas durante seu reconhecimento (TAFIT, 2004; STOCKALL & MARANTZ, 2006).

Modelos não-decomposicionais consideram que palavras se relacionam no léxico mental por meio de conexões fonológicas e semânticas. Para esses modelos, portanto, não há um nível morfológico de representação, e efeitos de *priming* entre palavras relacionadas são explicados em termos de coincidência de traços fonológicos e de sentido. Essa abordagem tem sido desafiada por estudos que utilizam a técnica de *priming* encoberto (FORSTER, 1998)². Nesses estudos, palavras morfollogicamente relacionadas têm efeito diferenciado de palavras com semelhança fonológica ou semântica (LONGTIN ET AL., 2003; RASTLE et al., 2004; GARCIA, 2009). Os resultados têm mais recentemente se direcionado em favor dos modelos decomposicionais (GARCIA, 2009; RASTLE et al., 2004; LONGTIN et al., 2003; ZWITSERLOOD, 1994; STOCKALL & MARANTZ, 2006; FIORENTINO & POEPEL, 2007a).

Um fator importante que parece condicionar o modo como palavras complexas são acessadas é a transparência semântica, ou seja, o quão recuperável é a relação entre o sentido dos morfemas constituintes e o sentido da palavra inteira. Palavras semanticamente transparentes como SAILBOAT (barco a vela) e TEACHER (professor) têm sua ativação

² “Nessa técnica, o *prime* é apresentado muito rapidamente, de modo que não seja conscientemente percebido. Sendo assim, a visualização é suficiente para ativar etapas iniciais do reconhecimento lexical, e essa ativação afeta o reconhecimento da palavra alvo.” (GARCIA, 2013)

facilitada pelo sentido de seus constituintes (SAIL=vela, BOAT=barco, TEACH=ensinar). Essa facilitação parece, por outro lado, ser afetada pelo tipo de tarefa quando a relação entre os constituintes e o todo é opaca (BOOTLEG=falsificado; mas BOOT=bota e LEG=perna). Estudos com *priming* encoberto indicam haver um momento inicial do reconhecimento da palavra em que os constituintes são ativados, uma vez que há facilitação entre constituintes e a palavra inteira independente da transparência semântica (RASTLE et al., 2004; LONGTIN et al., 2003; GARCIA, 2009). Quando o *prime* é apresentado por um tempo mais longo, no entanto, essa facilitação ocorre apenas para palavras semanticamente transparentes. Isso indica a existência de uma fase no reconhecimento da palavra, anterior ao pareamento semântico, em que seriam acionadas apenas características morfo-ortográficas da palavra.

Palavras como *brother* (irmão), que podem ser “pseudo-segmentadas” em uma raiz e um sufixo, seriam decompostas tanto quanto palavras como *teacher*, em que a segmentação é relevante para o reconhecimento. Porém, ainda que essa decomposição ocorra independentemente das entradas lexicais, existiria sensibilidade visual para fatores de nível mais elevado como o reconhecimento de morfemas, de modo que palavras como *brothel* (bordel) não são segmentadas. Nesse caso, ainda que uma “pseudo-raiz” *broth-* (caldo) possa ser reconhecida, o material reminescente *-el* não é um sufixo existente na língua. Efeitos desse tipo se diferenciam do modelo de decomposição plena proposto originalmente por TAFT & FORSTER (1975), o *Affix Stripping*, para o qual a decomposição da palavra passa primeiramente por um momento de remoção obrigatória dos afixos, anterior ao acesso lexical. Essa remoção ocorreria independentemente de o material reminescente existir ou não. Para TAFT & FORSTER (1975), portanto, palavras como *winter* (vento) seriam decompostas. Palavras como *brothel*, por outro lado, não seriam (GARCIA, 2013).

Utilizando também a técnica de *priming* encoberto e trabalhando com palavras compostas, SHOOLMAN & ANDREWS (2003), FIORENTINO (2006) e FIORENTINO & POEPEL (2007b) encontram resultados semelhantes aos de RASTLE et al. (2004). ZWITSERLOOD (2004) reporta efeitos de decomposição independente de transparência semântica em um teste de *priming* parcial não encoberto. Não houve efeito significativo para relação meramente ortográfica nesses estudos.

O uso da técnica de rastreamento ocular apresenta um ganho importante por possibilitar que as medidas sejam feitas durante o processo de leitura, podendo ser útil para caracterizar diferentes estágios do processamento morfológico a partir do movimento dos olhos. Maia et al. (2007) apresentam resultados associando a técnica de rastreamento ocular com o efeito Stroop (STROOP, 1935). O estudo revela maior atividade ocular em condições com morfemas transparentes do que nas com palavras opacas e pseudo-derivadas, o que os autores interpretaram como indicativo em favor de modelos de dupla rota. Resultados nessa mesma direção foram encontrados por ANDREWS et al. (2004) e JUHASZ et al. (2003), que obtiveram efeitos de frequência para primeira fixação (*first fixation*) no primeiro constituinte de palavras compostas, e também efeitos na manutenção da fixação (*gaze duration*) para os segundos constituintes. Esses estudos também encontraram efeitos que indicam representação no nível da palavra inteira, tanto em manutenção da fixação quanto nos tempos totais de leitura. Todos esses resultados indicam que exista representação no nível sublexical, o que não pode ser explicado por modelos não-decomposicionais.

Se os resultados desses estudos mencionados acima apontam consistentemente para a existência de um estágio de decomposição inicial que parece não ser semanticamente condicionado, deve existir um momento posterior em que o sentido complexo da palavra é ativado. Essa ativação deve se dar por meio da combinação entre os sentidos dos

morfemas resultantes da decomposição inicial. Garcia (2013) teve como foco de análise exatamente esse estágio mais tardio do reconhecimento lexical. De modo a manipular esses efeitos posteriores ao acesso aos morfemas, a autora estabeleceu o Campo Medial Anterior (AMF) como medida dependente para essa atividade combinatória. O AMF é um componente de MEG que ocorre entre 350 e 450 milissegundos após o início da apresentação de um elemento crítico ligado a algum tipo de computação semântica (BEMIS & PYLKKÄNEN, 2011).

O processamento lexical no tempo e no espaço³

Investigações utilizando as técnicas de eletroencefalografia (EEG) e magnetoencefalografia (MEG) têm se mostrado convenientes como forma de mapeamento *online* das computações envolvidas no processamento lexical. Essas técnicas têm a vantagem de acompanhar essas computações com uma resolução temporal de milissegundos. No entanto, para que se possam formular hipóteses testáveis, é necessário que se tenha uma ideia clara de como componentes relacionados à atividade cerebral podem estar associados às computações específicas que queremos examinar. Estudos realizados recentemente com a técnica de magnetoencefalografia têm identificado alguns componentes na forma da onda obtida ao se medir a atividade eletromagnética associada ao reconhecimento visual de palavras (TARKIAINEN et al., 1999; EMBICK et al., 2001; HALGREN et al., 2002; PYLKKÄNEN et al., 2002; CORNELISSEN et al., 2003). Esses componentes estão relacionados a diferentes estágios desse processo, e são sensíveis a fatores específicos relacionados ao estímulo. De acordo com esses trabalhos, a rota para o acesso lexical aconteceria em pelo menos quatro estágios, durante 400 ms após o início da visualização da palavra. Abaixo apresentam-se os estágios relacionados à percepção *visual* em MEG,

³ Seção adaptada de Garcia (2013)

mas a sequência é qualitativamente semelhante para estímulos visuais e auditivos (SALMELIN, 2010). Esses diferentes componentes atuam de forma equivalente aos estágios do reconhecimento visual postulados por modelos de decomposição inicial. Sendo assim, eles têm se mostrado como importantes ferramentas para mapear as computações associadas a esses diferentes estágios, podendo ajudar a resolver, no tempo e no espaço, os dados conflitantes dos estudos baseados em tempos de resposta.

O primeiro componente, identificado originalmente por Tarkiainen et al. (1999), ocorre em torno de 100-150 ms após o início da apresentação da palavra e está associado a características visuais mais concretas, como tamanho e luminância (*Type I response*). Esse componente tem origem na região occipital e apresenta sensibilidade para propriedades ortográficas. Estudos mais recentes realizados por Dikker e Pylkkänen revelam ainda que pistas sintáticas e léxico-semânticas do estímulo já afetam esse componente (DIKKER & PYLKKÄNEN, 2011; DIKKER et al., 2010).

Após esse estágio, há um momento em que as palavras são decompostas em seus morfemas constituintes, caracterizada pelo componente M170. Essa decomposição, que é pré-lexical e tem origem no córtex temporal inferior, ocorre entre 150 e 200 ms e é sensível a sequências de letras (*Type II response* TARKIAINEN et al., 1999). Estudos com ressonância magnética funcional (fMRI) localizam uma região no córtex visual que estaria consistentemente implicada com a identificação de palavras, e que ficou conhecida como a Área da Forma Visual da Palavra (*Visual Word Form Area – VWFA*). Essa área seria resultado de uma reorganização funcional associando a capacidade da leitura com restrições estruturais do sistema visual (DEHAENE et al., 2002; McCANDLISS et al., 2003; FRANÇA et al., 2013). LEWIS et al. (2011) argumentam ainda que a área poderia se chamar Área da Forma Visual de Morfemas, já que, nesse estágio do reconhecimento, ocorre decomposição da palavra baseada na forma de raízes e afixos.

Essa afirmação se baseia em resultados de uma série de estudos com MEG que trouxeram uma caracterização temporal mais precisa para as propriedades identificadas nos estudos com neuroimagem sobre essa área (SALMELIN, 2007; ZWEIG & PYLKKÄNEN, 2009; SOLOMYAK & MARANTZ, 2009; 2010; LEWIS et al., 2011).

Zweig & Pylkkänen (2009) comparam diretamente palavras morfológicamente complexas e simples, dialogando com a literatura de *priming* encoberto (LONGTIN et al., 2003; RASTLE et al., 2004), e identificam o M170 como a primeira resposta visual sensível à complexidade morfológica. SOLOMYAK & MARANTZ (2009, 2010) e LEWIS et al. (2011) apresentam resultados que corroboram a existência de decomposição obrigatória baseada em relações de frequência entre raízes e afixos (HAY, 2001), a partir de modulações nas amplitudes do M170. LEWIS et al. (2011) encontram ainda evidências de que possa haver representação da palavra inteira já disponível nos estágios mais iniciais do reconhecimento lexical – apenas para palavras cuja decomposição não é determinante para o acesso lexical –, indicando que exista uma rota dupla de acesso atuando desde o início do reconhecimento da palavra.

Há ainda um terceiro componente, o M250, que ocorre entre 200 e 300 ms após o início da apresentação da palavra e parece ter origem na parte posterior do hemisfério esquerdo. Não há consenso quanto às possíveis computações associadas a esse componente, mas ele parece ter relação com certas propriedades fonológicas do estímulo (PYLKKÄNEN et al., 2002).

Todos esses estágios até aqui estão envolvidos com processamentos pré-lexicais, ou seja, eles antecedem o pareamento arbitrário entre forma e sentido que caracteriza o acesso lexical (HALLE & MARANTZ, 1993). Esse pareamento, por sua vez, parece ser captado por um componente mais tardio, em torno de 350 ms após a apresentação visual da palavra. Esse componente, sensível à frequência e à repetição (PYLKKÄNEN ET AL., 2002) – mas não à competição entre as representações ativadas pelo estímulo –, tem sido identificado como correlato do acesso lexical.

FIorentino & PoePpel (2007) utilizam esse componente para mapear estágios do curso temporal da decomposição em palavras compostas do inglês, manipulando as propriedades tidas como envolvidas no acesso lexical – e captadas pelo M350 – para comparar diretamente palavras compostas com palavras simples. Assumindo que o M350 é sensível a fatores relacionados ao acesso lexical, itens mais curtos e com maior frequência deveriam apresentar ativação mais rápida do M350 (*peak latency*). Sendo assim, os autores previram que, se houvesse ativação dos constituintes das palavras compostas, eles teriam latências menores do que as palavras simples – maiores e menos frequentes – com que foram comparadas. Como previsto, as latências do M350 associadas às palavras compostas foram mesmo menores, indicando ativação dos constituintes em estágios iniciais do reconhecimento de palavras complexas. Isso aponta para a existência de representações no nível sublexical. Aqui, é relevante observar que o componente neurofisiológico utilizado para caracterizar a decomposição não foi necessariamente o mesmo em palavras derivadas e palavras compostas. Nas palavras derivadas, essa constatação foi feita 1) comparando diretamente palavras derivadas com palavras simples (ZWEIG & PYLKKÄNEN, 2009), 2) comparando efeitos relacionados à forma visual da palavra com efeitos relacionados ao sentido (SOLOMYAK & MARANTZ, 2009), ou, ainda, 3) comparando frequências de sequências ortográficas em final de palavra com sequências correspondentes a afixos da língua (SOLOMYAK & MARANTZ, 2010).

Cada tipo de manipulação teve efeitos marcados em diferentes componentes na forma da onda. Manipulações ortográficas obtiveram resposta associada ao M100, ao passo que manipulações de morfemas e de sentido obtiveram resposta associada ao M170 e ao M350, respectivamente. Em palavras compostas, a decomposição foi constatada pela ativação lexical dos constituintes, indicada pelas

latências de pico do M350 mais rápidas que as latências associadas às palavras monomorfêmicas de mesmo tamanho e frequência. Não foram encontrados efeitos no componente M170, o que indica que houve acesso ao sentido dos constituintes, e não apenas um desmembramento pré-lexical baseado na forma visual das raízes. STOCKALL & MARANTZ (2006) também utilizam o M350 como índice de acesso lexical (ativação da raiz) para mostrar que o reconhecimento de palavras flexionadas envolve sua decomposição em morfemas lexicais e funcionais, independente da regularidade.

Resumindo, várias formas já foram usadas para verificar as etapas do reconhecimento visual de palavras. Independente da metodologia utilizada, existe um consenso sobre o fato de haver um momento do reconhecimento que envolve a ativação de constituintes. Basicamente, essa decomposição é constatada a partir de diferentes vias, utilizando diferentes metodologias. Recentemente, estudos eletrofisiológicos têm procurado localizar no tempo as etapas envolvidas no reconhecimento lexical.

Em palavras derivadas, essa decomposição parece mais bem caracterizada, envolvendo uma distinção mais clara entre o momento em que as palavras são decompostas com base na forma visual dos morfemas (captada por experimentos utilizando *priming* encoberto e detectável pelo componente M170) e o momento posterior em que se dá o acesso à raiz, com o pareamento entre forma e sentido (detectável pelo componente M350). Levando-se em consideração os resultados desses estudos, todas as palavras que podem ser segmentadas em morfemas existentes são segmentadas (teacher ✓; brother ✓; brothel ✕), nos momentos mais iniciais do reconhecimento.

Palavras compostas, por outro lado, não podem ser decompostas com base em uma relação entre afixos e raízes. Conforme reconhece Libben (1994), é mais difícil conceber uma heurística sistemática para

a segmentação de palavras compostas em seus morfemas constituintes. Por serem constituídos apenas por morfemas de classe aberta, a decomposição não pode ocorrer checando-se uma lista relativamente pequena, como no caso dos afixos. Além disso, palavras compostas possuem mais chance de serem segmentáveis em diferentes pontos, gerando múltiplos constituintes possíveis (*clamprod* pode ser segmentado em *clamp-rod* e em *clam-prod* – cf. LIBBEN et al., 1999). De qualquer forma, é praticamente consensual, entre os estudos realizados com palavras compostas, que existe acesso aos constituintes durante o curso temporal do reconhecimento. Essa decomposição, assim como ocorre para as palavras derivadas, acontece independente da transparência semântica (ZWITSERLOOD, 1994; FIORENTINO & POEPEL, 2007b). Ou seja, se as palavras compostas são segmentáveis, elas devem ser segmentadas (teacup ✓; hogwash ✓; penguin ✕).

Sendo assim, se todas as palavras com alguma estrutura interna possível são decompostas (teacup, hogwash, carpet / teacher, apartment, corner), uma questão que se configura mais recentemente diz respeito a quais seriam as computações subsequentes que dariam conta da interpretação correta das palavras complexas. Entender quais são as subrotinas do processamento lexical e quando elas ocorrem no curso temporal do reconhecimento – e apontar quais medidas dependentes captam essas diferentes subrotinas – é um trabalho ainda em progresso, e um desafio para pesquisas sobre o acesso lexical.

Pressupondo a ativação sublexical durante o processamento de palavras, GARCIA (2013) examinou a atividade neuronal associada a efeitos combinatórios pós-lexicais que devem ocorrer entre os morfemas ativados, identificando um possível componente que possa ser usado para testar futuras hipóteses sobre processos combinatórios no curso do reconhecimento de palavras compostas.

Como afirma GARCIA (2013),

A composicionalidade em palavras ou expressões linguísticas não é sempre indiscutível. Uma questão importante para o estudo da linguagem em uma perspectiva biolinguística deve ser, portanto, que tipo de composicionalidade faz parte do inventário de operações da linguagem natural. (GARCIA, 2013)

Estudos recentes têm se dedicado a isolar algumas operações combinatórias básicas. Manipulando a composicionalidade semântica em sintagmas adjetivais e em sentenças envolvendo coerção e violações semânticas, esses estudos têm encontrado atividade aumentada do AMF associada à composição de significados complexos (PYLKKÄNEN & McELREE, 2007; PYLKKÄNEN, MARTIN, McELREE & SMART, 2009; BRENNAN & PYLKKÄNEN, 2008; BRENNAN & PYLKKÄNEN, 2010; PYLKKÄNEN, OLIVERI & SMART, 2009; BEMIS & PYLKKÄNEN, 2011). GARCIA (2013) se soma a essa série de trabalhos, verificando que esses mesmos efeitos podem ser reproduzidos por processos de composição semântica entre os morfemas constituintes de palavras compostas.

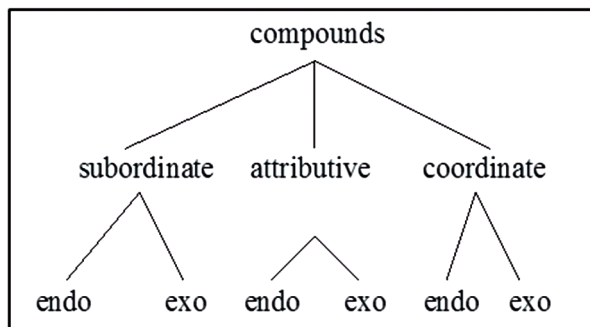
Tomando como base BEMIS e PYLKKÄNEN (2011), em que a correlação entre o AMF e processos combinatórios foi encontrada em sintagmas adjetivais envolvendo a combinação entre um nome e um adjetivo (*red boat* = barco vermelho), esse mesmo tipo de efeito foi investigado por GARCIA (2013) no âmbito de palavras compostas, como SAILBOAT.

Como essas palavras são constituídas pela concatenação de dois morfemas lexicais, analogamente ao que ocorre com a modificação em sintagmas como *red boat*, consideramos que seria o tipo de estímulo mais adequado a um primeiro teste que busca encontrar esses efeitos em palavras. (GARCIA, 2013)

Uma questão em aberto com relação ao processamento de palavras compostas é quanto a se existe um índice neurofisiológico independente para a composição entre seus morfemas constituintes. Existem estudos que apontam o papel da composição para dar conta, sobretudo, do processamento de palavras novas (SCHREUDER & BAAYEN, 1995; GAGNÉ, 2002; JI et al., 2011). Levando-se em consideração o curso da rota visual do reconhecimento de palavras (PYLKKÄNEN & MARANTZ, 2003), é possível pensar que exista algum índice, em torno do 400 ms, que possa estar envolvido com as operações combinatórias pós-lexicais.

Uma característica particular das palavras compostas é o fato de que seus dois constituintes são associados por uma relação gramatical não expressa (BISETTO & SCALISE, 2005). Por exemplo, a palavra *teacup* poderia ser associada à expressão *cup for tea*, caracterizando uma relação de subordinação entre seus constituintes. Os constituintes dos compostos podem ainda estar relacionados por uma relação de coordenação, como no caso do adjetivo *bittersweet*, não se esgotando nesses dois tipos as relações possíveis. BISETTO e SCALISE (2005) classificam as palavras compostas como sendo de diferentes tipos, baseando-se nos tipos de relações entre seus constituintes. A classificação como endocêntricas ou exocêntricas está relacionada com o núcleo da palavra composta, isto é, com o fato de o sentido da palavra inteira ser ou não um hiperônimo do seu núcleo. Essa classificação parece ser equivalente ao que LIBBEN (1998) chama de componencial e não-componencial, respectivamente. A classificação em subordinadas, atributivas ou coordenadas diz respeito à relação sintática entre seus constituintes.

FIGURA 1: Classificação das palavras compostas (BISETTO & SCALISE, 2005; p. 9)



Outro fator complicador para o entendimento de como as palavras compostas são processadas é a transparência semântica. Libben (1995) propõe um modelo em que a ativação das palavras compostas é separada em três níveis: (1) o nível do estímulo, para dar conta da habilidade que temos de interpretar palavras compostas novas; (2) o nível lexical, que daria conta do fato de sabermos que “palavras como *strawberry* (=morango) contêm o item lexical *straw* (=canudo), mas não o significado desse item” (LIBBEN, 2005: p. 58); e (3) o nível conceptual, para dar conta do fator componencialidade. Nesse modelo, as diferenças em termos de transparência semântica são caracterizadas a partir de uma ação combinada de ligações facilitatórias e inibitórias entre o nível lexical e o nível conceptual. Considerando que palavras complexas sejam representadas e acessadas por meio de seus constituintes, as palavras compostas podem ser abordadas como sendo o resultado de uma simples concatenação entre dois elementos lexicais. No entanto, a exata relação que pode existir entre esses dois elementos é bastante variável (JACKENDOFF, 2002; DRESSLER, 2006; JAREMA, 2006; BISETTO & SCALISE, 2005). De acordo com LIBBEN (2006), palavras compostas devem ser facilmente segmentáveis entre seus morfemas constituintes. Caso contrário, palavras compostas novas não seriam interpretáveis.

Por outro lado, existe um domínio idiossincrático que coloca essas palavras “em uma interseção entre itens lexicais e sentenças, refletindo as propriedades tanto das representações linguísticas na mente, quanto do processamento gramatical” (LIBBEN, 2006)⁴. Modelos eficientes de processamento lexical devem ser capazes de dar conta dessas diferenças.

Ainda que as relações entre os constituintes das palavras compostas possam ser de diferentes naturezas, pode-se afirmar que as palavras compostas, em geral, são o resultado do processo de composição entre o significado de seus constituintes. Isso deve ser necessariamente verdadeiro pelo menos para palavras novas, já que elas, em princípio, não devem ter seu sentido idiossincrático representado no léxico mental.

GARCIA (2013) realizou um teste de nomeação em que foram comparadas palavras compostas e palavras-controle simples da língua inglesa (e.g. SAILBOAT=barco a vela vs. SPINACH=espinafre). O tarefa foi realizada enquanto se monitorava a atividade eletromagnética cortical dos participantes. Considerando que o esforço computacional é maior em palavras compostas, uma vez que envolve a combinação dos sentidos dos constituintes, a amplitude do AMF deve ser aumentada durante o reconhecimento dessas palavras, em comparação com as palavras simples.

Conforme mostrado acima, outros estudos já utilizaram a técnica da magnetoencefalografia para entender os mecanismos envolvidos no reconhecimento de uma palavra. O componente M350, sensível a fatores como frequência e repetição, foi relacionado ao momento em que ocorre o acesso lexical, ou seja, em que ocorre o pareamento entre forma e sentido. Durante o reconhecimento da palavra SAILBOAT, então, ocorre a ativação da raiz SAIL e a ativação da raiz BOAT, indicadas pelo componente M350. Nesse estágio, o reconhecimento de palavras como SAILBOAT não é diferente do reconhecimento de

⁴ “compound words are structures at the cross-roads between words and sentences reflecting both the properties of linguistic representation in the mind and grammatical processing.” (LIBBEN, 2006; p. 3)

palavras como BOOTLEG. Sendo assim, ainda que em tese não exista relação semântica entre BOOT e BOOTLEG, a palavra é decomposta nos estágios mais iniciais do reconhecimento – como indicado por experimentos com *priming* encoberto e com MEG.

“O que diferencia a natureza dessas duas categorias de palavras (transparentes e opacas) deve ocorrer, portanto, em algum momento de composição semântica posterior ao acesso às raízes” (GARCIA, 2013). Olhado para esses momentos que seguem o acesso às raízes, o estudo aqui reportado buscou indícios da recombinação entre os morfemas ativados, que culminaria na ativação do sentido da palavra composta.

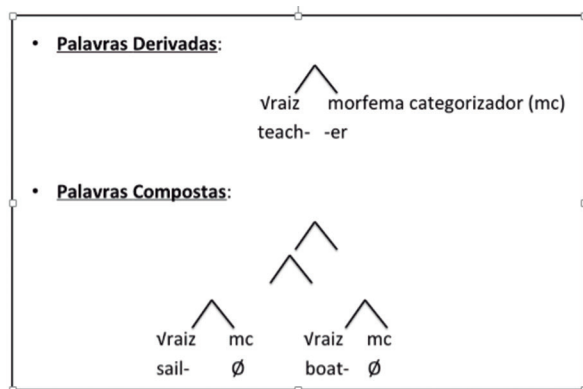
A opção por palavras compostas foi feita para fins de paralelismo com o trabalho de BEMIS e PYLKKÄNEN (2011), trazendo o foco para o contexto intralexical. O processo de formação da palavra, no entanto, pode ser relevante na segmentação e recomposição do estímulo e na ativação do sentido da palavra. Por exemplo, “palavras derivadas (e flexionadas), são formadas por uma raiz e por um morfema categorizador (e.g. [TEACH]ER) ou [PARK]ED)), ao passo que palavras compostas são formadas por pelo menos duas raízes” (GARCIA, 2013).

Considerando uma proposta não-lexicalista como a da Morfologia Distribuída (HALLE & MARANTZ, 1993), cada uma dessas raízes lexicais – que pode se relacionar com a palavra inteira de forma transparente, como em SAILBOAT, ou de forma opaca, como em BOOTLEG – deve ser concatenada a um morfema categorizador (e.g. [SAIL]Ø_N e [BOAT]Ø_N), e o significado dessas duas raízes deve ainda ser combinado para formar o significado da palavra composta SAILBOAT. Essa combinação pode ser por uma relação hierárquica de subordinação (SAILBOAT – [barco] [com vela]) ou por coordenação (BITTERSWEET – [amargo] e [doce]). A composição pode ainda ter palavras com a mesma categoria formal, como em SAILBOAT (nome-nome), ou com categorias diferentes, como em REDNECK (adjetivo-nome – red=vermelho; neck= pescoço; redneck=caipira).

FRANÇA et al. (2008) notam que há uma “invisibilidade das computações composicionais que se estabelecem a partir do ponto da arbitrariedade saussureana” (p.47), quando se utiliza a técnica de ERP. Esse trabalho compara palavras monomorfêmicas com palavras polimorfêmicas. Os resultados obtidos indicaram que, em um paradigma de *priming* auditivo com aferição da atividade elétrica relacionada à leitura dos alvos, a quantidade de camadas morfológicas não produz uma gradação na dificuldade de ativação da palavra. Sendo assim, a facilitação de **centralização** em relação a **centro** é equivalente à facilitação de **pureza** em relação a **puro**, havendo indicação de que o aumento das camadas morfológicas não acarretaria maior custo de ativação. Por outro lado, MAIA, LEMLE & FRANÇA (2007) reportam experimentos que investigam a decomposição morfológica na leitura de palavras isoladas, utilizando a técnica de rastreamento ocular. Esse trabalho revela maior atividade ocular (fixações e movimentos sacádicos) na leitura de palavras morfológicamente complexas, o que revelaria maior custo de processamento para palavras morfológicamente complexas. A sensibilidade do rastreador ocular para essas camadas composicionais pós-lexicais revela que há um custo de processamento para a soma das camadas funcionais após o acesso lexical. Essa composicionalidade, no entanto, deve ser de uma natureza diferente da encontrada em palavras compostas e captadas pelo AMF.

Como são constituídas por morfemas livres, as palavras compostas lidam com um inventário mais aberto de possibilidades. As palavras derivadas e flexionadas, por sua vez, podem contar com uma heurística mais sistemática na computação de cada nova camada funcional (*Affix Stripping*: TAFT & FORSTER, 1975).

FIGURA 2: Composicionalidade em palavras derivadas e compostas (In: GARCIA, 2013)



Estudar os diferentes níveis de complexidade lexical, assim como suas diferentes naturezas, pode contribuir com um maior entendimento sobre os mecanismos envolvidos no reconhecimento de palavras e, em maior escala, fornecer peças importantes para uma caracterização explicativa da dinâmica espaço-temporal do circuito cerebral de construção de sentido (cf. GARCIA, 2013).

Referências

- BEMIS DK., & PYLKKÄNEN L. **Simple Composition:** An MEG investigation into the comprehension of minimal linguistic phrases. *Journal of Neuroscience*, 31(8): 2801-2814. 2011.
- BISETTO A, SCALISE S. **Classification of compounds.** *Lingue e Linguaggio* 2: 319-332. 2005.
- BRENNAN J, PYLKKÄNEN L. **Processing Events:** Behavioral and neuromagnetic correlates of aspectual coercion. *Brain and Language*, 106, 132-143. 2008.

BRENNAN, J.; PYLKKÄNEN, L. **Processing Psych Verbs: Behavioral and MEG Measures of Two Different Types of Semantic Complexity.** *Language and Cognitive Processes*, 25 (6), 77-807. 2010.

CARAMAZZA, A.; SHELTON JR. **Domain-specific knowledge systems in the brain: The animate-inanimate distinction.** *Journal of Cognitive Neuroscience*, 10, 1-34. 1998.

CHOMSKY, N. **The Minimalist Program.** MIT Press. 1995.

COHEN, L.; LEHÉRICY S.; CHOCHON, F. et al. **Language-specific tuning of visual cortex? Functional properties of the visual word form area.** *Brain*. 125: 1054-1069. 2002.

DEHAENE S. **Reading in the brain.** Penguin Viking, 2009.

DEHAENE, S.; LE CLECH, G.; POLINE J. B.; LEBIHAN D.; COHEN, L. **The visual word form area.** A prelexical representation of visual words in the fusiform gyrus. *NeuroReport*, 13:321--325, 2002.

EMBICK, D., POEPPPEL, D. **Defining the relation between linguistics and neuroscience.** To appear in *Twenty-first century psycholinguistics: Four cornerstones*, Anne Cutler (ed.), Lawrence Erlbaum Associates. 2004.

FEDERMEIER, K.D.; KUTAS, M. **A rose by any other name: long-term memory structure and sentence processing.** *J. Mem. Lang.* 41, 469–495. (1999)

FIORENTINO, R.; POEPPPEL, D. **Compound words and structure in the lexicon.** *Language and Cognitive Processes*, 12, 953-1000. 2007.

FIORENTINO, R. POEPPPEL, D. **Processing of compound words: An MEG study.** *Brain and Language*, 103, 18-19. 2007b.

FIORENTINO, R. **Masked morphological priming of compound constituents:** Specifying the locus of morpho-orthographic segmentation effects using English compound words. Unpublished rawdata. 2006.

FRIEDERICI, A.; FRISCH, S. **Verb Argument Structure Processing:** The Role of Verb-Specific and Argument-Specific Information. *Journal of Memory and Language* 43. 2000.

GARCIA, D. C. **Elementos Estruturais no Acesso Lexical:** o Reconhecimento de Palavras Multimorfêmicas no Português Brasileiro. Dissertação de Mestrado. Rio de Janeiro: UFRJ/Faculdade de Letras. 2009.

GARCIA, D. C. **Efeitos composicionais no reconhecimento visual de palavras compostas em inglês:** um estudo com MEG. Tese de Doutorado. Rio de Janeiro: UFRJ/Faculdade de Letras. 2013.

GIRAUDO, H.; GRAINGER, J. **Effects of prime word frequency and cumulative root frequency in masked morphological priming.** *Language and Cognitive Processes*, 15, 421-444. 2000.

HAGOORT, P.; HALD, L.; BASTIAANSEN, M.; PETERSSON, K. M. **Integration of word meaning and world knowledge in language comprehension.** *Science*. Apr 16;304(5669):438-41. Epub 2004 Mar 18. 2004.

HALLE, M.; MARANTZ, A. **Distributed morphology and the pieces of inflection.** In K. Hale and S.J. Keyser, eds., *The View From Building 20: Linguistic Essays in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 111 – 176. 1993.

HANSEN, P.; KRINGELBACH, M.; SALMELIN R. **MEG:** An Introduction to Methods. Oxford University Press, Oxford, UK. 2010.

HARLEY, H.; NOYER, R. **Mixed nominalizations, short verb movement, and object shift in English.** In Proceedings of NELS 28, ed. by P.N. TAMANJI & K. KUSUMOTO, GLSA, Amherst, 143-157. 1998b.

HARLEY, H.; NOYER, R. **‘Licensing in the non-lexicalist lexicon: nominalizations, vocabulary items and the Encyclopaedia.’** In MITWPL 32: Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect, ed. Heidi Harley. MITWPL, Cambridge, 119-137. *Cognitive Sciences* 9(7). 342-348. 1998a.

KUTAS, M. & HILLYARD, S. A. **Reading senseless sentences:** Brain potentials reflect semantic incongruity. *Science*, 207, 203–205. 1980b.

LEWIS, G.; SOLOMYAK, O. & MARANTZ, A. **The neural basis of obligatory decomposition of suffixed words.** *Brain & Language*, 118, 118-127. 2011.

LIBBEN, G. **Semantic transparency in the processing of compounds:** consequences for representation, processing, and impairment. *Brain and Language*, 61, 30-44. 1998.

LIKERT, R. **A technique for the measurement of attitudes.** *Archives of Psychology*, Vol 22 140, 1932, 55.

LONGTIN, Catherine-Marie; SEGUI, Juan; HALLÉ, Pierre A. **Morphological priming without morphological relationship.** *Language and Cognitive Processes* 18.3:313-334. 2003.

MARANTZ, A. **‘No escape from syntax: Don’t try morphological analysis in the privacy of your own Lexicon.’** Proceedings of the 21st Annual Penn Linguistics Colloquium: Penn Working Papers in Linguistics 4: 2, ed. Alexis Dimitriadis et.al. 201-225. 1997a.

MARANTZ, A. **Stem suppletion, or the arbitrariness of the sign.** Talk given at the Universite' de Paris VIII. 1997b.

MARINKOVIC, K. **Spatiotemporal dynamics of word processing in the human cortex.** *Neuroscientist*, 10, 142–152. 2004.

MARIS, E.; OOSTENVELD, R. **Nonparametric statistical testing of EEG- and MEG-data NCI**, Biological Psychology, Radboud University Nijmegen, Nijmegen, The Netherlands F.C. Donders Center for Cognitive Neuroimaging, Radboud Universityf Nijmegen, Nijmegen, The Netherlands. 2007.

MARSLLEN-WILSON WD, Welsh, A. **Processing interactions and lexical access during word-recognition in continuous speech.** *Cognitive Psychology*, 16, 29-63. 1978.

MARSLLEN-WILSON, W., et al. **Morphology and meaning in the English mental lexicon.** *Psychological Review*, 101, 3-33. 1994.

PINKER S. Words and Rules: **The Ingredients of Language.** Phoenix: Harper Perennial. 2000.

MARR, D. **Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.** San Francisco: Freeman. 1982.

PYLKKÄNEN, L. & MCELREE, B. **The syntax-semantic interface: On-line composition of sentence meaning.** *Handbook of Psycholinguistics* (2nd Ed), 537-577. NY: Elsevier. 2006.

PYLKKÄNEN, L. & MCELREE, B. **An MEG Study of Silent Meaning.** *Journal of Cognitive Neuroscience*, 19, 1905-1921. 2007.

PYLKKÄNEN, L.; MARANTZ, A. **Tracking the time course of word recognition with MEG.** *Trends in Cognitive Sciences*, 7, 187-189. 2003.

PYLKKÄNEN, L.; BRENNAN, J.; BEMIS, D. K. **Grounding the Cognitive Neuroscience of Semantics in Linguistic Theory.** *Language and Cognitive Processes*. 2010.

PYLKKÄNEN, L.; MARTIN, A. E., MCELREE, B.; SMART, A. **The Anterior Midline Field: Coercion or Decision Making?** *Brain and Language*, 108, 184-90. 2009.

PYLKKÄNEN, L.; OLIVERI, B.; SMART, A. **Semantics vs. World Knowledge in Prefrontal Cortex.** *Language and Cognitive Processes*, 24, 1313-1334. 2009.

PYLKKÄNEN, L.; STRINGFELLOW, A.; MARANTZ, A. **Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density.** *Brain and Language*, 81, 666–678. 2002.

RASTLE, Kathleen; MATTHEW Davis; BORIS, New. **The Broth in my Brother's Brothel: Morpho-Orthographic Segmentation in Visual Word Recognition.** *Psychonomic Bulletin & Review* 11.6:1090-1098. 2004.

RAYNER, K.; POLLATSEK, A. **Eye movements in reading: A tutorial review.** In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp.327-362). Hillsdale, NJ: Erlbaum. 1987.

SCHREUDER, R.; BAAYEN, R. **Modeling morphological processing.** In Feldman, L. B. (ed), *Morphological Aspects of Language Processing*. Lawrence Erlbaum, Hillsdale, New Jersey, 131–154. 1995.

SEIDENBERG, M. S.; GONNERMAN, L. M. **Explaining derivational morphology as the convergence of codes.** Trends in Cognitive Sciences, 4, 353-361. 2000.

SHOOLMAN, N.; ANDREWS, S. **Racehorses, reindeers and sparrows:** Using masked priming to investigate morphological influences on word identification. *Masked priming: The state of the art.* Psychology Press. 2003.

SOLOMYAK, O.; MARANTZ, A. **Lexical access in early stages of visual word processing:** A single-trial correlational MEG study of heteronym recognition. *Brain and Language*, 108, 191–196. 2009.

STOCKALL, L.; MARANTZ, A. **A single route, full decomposition model of morphological complexity:** MEG evidence. *The Mental Lexicon*, 1(1), 85-123. 2006.

TAFT, M & KENNETH I. Forster. **Lexical storage and retrieval of prefixed words.** *Journal of Verbal Learning and Verbal Behavior* 14(6). 638-647. 1975.

Taft, M. **Morphological decomposition and the reverse base frequency effect.** *The Quarterly Journal of Experimental Psychology Section A*, 57, 745–765. 2004.

Zweig, E. & PYLKKÄNEN, L. **A visual M170 effect of morphological complexity.** *Language and Cognitive Processes*, 24, 412 - 439. 2009.

Zwitserlood, P. **The role of semantic transparency in the processing and representation of Dutch compounds.** *Language and Cognitive Processes*, 9, 341- 368. 1994.