

Universidade Federal do Rio de Janeiro

A TRANSPARÊNCIA SEMÂNTICA E O PROCESSAMENTO
MORFOLÓGICO EM PALAVRAS COMPOSTAS COM DOIS
IDEOGRAMAS EM JAPONÊS

Wendy Barile

2010

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



A TRANSPARÊNCIA SEMÂNTICA E O PROCESSAMENTO MORFOLÓGICO EM
PALAVRAS COMPOSTAS COM DOIS IDEOGRAMAS EM JAPONÊS

Wendy Barile

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Linguística da Universidade Federal do Rio de Janeiro como quesito para a obtenção do título de Mestre em Linguística.

Orientador: Professor Doutor Marcus Antonio Rezende
Maia

Rio de Janeiro
Janeiro de 2010

Barile, Wendy

A transparência semântica e o processamento morfológico em palavras compostas com dois ideogramas em japonês/ Wendy Barile - Rio de Janeiro: UFRJ/ Faculdade de Letras, 2010.

xi, 85f.: il.; 31 cm.

Orientador: Marcus Antonio Rezende Maia

Dissertação (mestrado) – UFRJ/ Faculdade de Letras/ Programa de Pós-graduação em Linguística, 2010.

Referências Bibliográficas: f. 78-82.

1. Morfologia. 2. Semântica. 3. Psicolinguística. 4. Língua Japonesa.
I. Maia, Marcus Antonio Rezende (Orient). II. Universidade Federal do Rio de Janeiro, Instituto de Pós-graduação em Linguística. III. A transparência semântica e o processamento morfológico em palavras compostas com dois ideogramas em japonês.

**A TRANSPARÊNCIA SEMÂNTICA E O PROCESSAMENTO MORFOLÓGICO EM
PALAVRAS COMPOSTAS COM DOIS IDEOGRAMAS EM JAPONÊS**

Wendy Barile

Orientador: Prof. Doutor Marcus Antonio Rezende Maia

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Linguística, da Faculdade de Letras da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários a obtenção do Título de Mestre em Linguística.

Examinada por:

Presidente, Professor Marcus Antonio Rezende Maia – Orientador

Professora Doutora Anieli Improta França – UFRJ

Professora Doutora Sumiko Nishitani Ikeda – Puc-SP

AGRADECIMENTOS

Gostaria de expressar os meus sinceros agradecimentos:

Ao professor Marcus Maia, que desde a Iniciação Científica não apenas me ensinou tanto sobre Linguística, como também e, principalmente, me mostrou o grande prazer presente na Pesquisa;

Às professoras Aniela França e Sumiko Ikeda, por todos os conselhos, os auxílios prestados e também por terem aceitado em participar da banca de defesa desta dissertação;

Ao atual coordenador da pós-graduação da Letras da UFRJ, Professor Mário Martellota, e aos professores: Celso Novaes, Miriam Lemle e Gastão Coelho, por toda a ajuda e ensinamentos;

Às professoras do Departamento de Japonês da UFRJ: Sonia Ninomiya, Yumiko Tominaga e, sobretudo, à professora Eli Yamada, por toda luta e incentivo ao longo dos anos;

À coordenadora Rika Hagino da Escola Modelo do Rio de Janeiro por todo o apoio e a parceria durante o período em que trabalhamos juntas;

Ao diretor, aos professores e às mães da Escola Japonesa situada em Santa Tereza que sempre se prontificaram em me auxiliar e em participar dos experimentos pilotos que vim realizando desde a época da graduação;

À professora Nícia Tanaka que tanto me ajudou na busca de participantes no Rio de Janeiro;

Aos amigos da UFRJ: Cris, Dani, Kátia, Marcela, Sara, Marije, Marília, Sabrina, Thiago;

Aos professores Robert Fiorentino, David Poeppel, Alec Marantz, Linnaea Stockall e ao aluno Pedro Alcocer por toda a ajuda prestada, seja de referências, conselhos, dicas ou apoio;

À JICA por ter tornado real o meu sonho de conhecer o Japão e ter feito possível a realização dos experimentos durante a minha estadia na Universidade de Tecnologia de Toquio;

Aos setores administrativos da JICA em São Paulo e em Yokohama;

À Associação de Descendentes no Exterior (*Kaigai Nikkeijin Kyoukai*) da JICA de Yokohama, sobretudo, ao Sr. Jun Yoshimura, à Sra. Tomoko Noguchi e à Sra. Junko Abe;

Ao professor Terry Joyce da Tama University que me orientou enquanto estava no Japão, e ao professor Hiroyuki Akama que me recebeu tão gentilmente em seu laboratório e me ajudou diversas vezes nas questões administrativas;

Aos quase cem alunos da Universidade de Tecnologia de Tóquio que se prontificaram em participar dos experimentos realizados para este estudo;

Ao CNPq, por ter financiado parte desta pesquisa;

Aos meus amigos Cecilia e Helio, por transformarem os meus dias de folga em momentos tão prazerosos;

Aos meus irmãos, Alan e Ivi, por sempre estarem ao meu lado;

Aos meus pais, Alessandro e Rutmi, pois, sem eles, nada disso seria possível.

RESUMO

A TRANSPARÊNCIA SEMÂNTICA E O PROCESSAMENTO MORFOLÓGICO EM PALAVRAS COMPOSTAS COM DOIS IDEOGRAMAS EM JAPONÊS

Wendy Barile

Orientador: Marcus Antonio Rezende Maia

Resumo da Dissertação de Mestrado submetida ao Programa de Pós-graduação em Lingüística, da Faculdade de Letras, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Lingüística.

O presente estudo procurou observar a natureza morfológica das palavras compostas em japonês e a interação deste fator com a transparência semântica. Foram elaborados dois experimentos de *priming* utilizando palavras compostas transparentes e opacas semanticamente em japonês apresentando cinco condições para o *prime*: o primeiro constituinte, o segundo constituinte, um constituinte não relacionado e uma condição em branco. No primeiro experimento, utilizando um SOA de 150ms, foram observadas diferenças entre os tipos de palavra. Embora tenha ocorrido efeito de facilitação nas palavras compostas transparentes precedidas pelo primeiro ou segundo constituinte, o mesmo não foi observado nas palavras compostas opacas. Por outro lado, o segundo experimento, com SOA de 60ms, demonstrou padrões de resposta similares entre as palavras transparentes e opacas: não houve interação, nem efeito significativo em relação ao tipo da palavra, mas houve um efeito extremamente significativo para as condições do *prime*. Os resultados parecem apontar para a participação das unidades morfológicas e um encontro tardio dos traços semânticos no processamento de palavras compostas em japonês. Este resultado está de acordo com a teoria da Morfologia Distribuída que defende a decomposição morfológica e a participação de traços semânticos tardiamente no curso temporal do reconhecimento de uma palavra. Adicionalmente, os experimentos também demonstraram diferenças significativas entre as duas condições controles (constituinte não relacionado e condição em branco). Uma vez que as palavras precedidas por constituintes não relacionados causaram inibição, indaga-se se o uso desta condição como controle é ideal para demonstrar resultados genuínos.

Palavras chaves: Morfologia, Semântica, Psicolinguística, Língua Japonesa, Compostos.

Rio de Janeiro
Janeiro de 2010

*ABSTRACT*SEMANTIC TRANSPARENCY AND MORPHOLOGICAL PROCESSING IN
JAPANESE COMPOUND WORDS WITH TWO CHARACTERS

Wendy Barile

Orientador: Marcus Antonio Rezende Maia

Abstract da Dissertação de Mestrado submetida ao Programa de Pós-graduação em Lingüística, da Faculdade de Letras, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Lingüística.

The present study investigates the morphology of Japanese compound words and its interaction with semantic transparency. Two priming experiments were conducted using semantically transparent and opaque Japanese compound words with primes counterbalanced with a Latin-square design across four conditions: first constituent; second constituent; unrelated constituent; and blank condition. Experiment 1, employing an SOA of 150ms, showed significant differences between the transparent and opaque words. Although priming effects were observed for the transparent words in both the first and second constituent conditions, this same pattern was not found for the opaque words. In contrast, Experiment 2, with an SOA of 60ms, showed similar patterns for both transparent and opaque words. Specifically, in the absence of interaction and word type effects, there was a significant effect for the prime condition. Accordingly, these results suggest an important role for morphological units and the subsequent involvement of semantic features in the processing of compound words in Japanese. This notion is generally consistent with the Distributed Morphology Theory, which not only argues for morphological decomposition, but also for the later participation of semantic features in the time course of word recognition. Additionally, the experiments also found significant differences between the two baseline conditions (the blank and unrelated conditions). Given that the unrelated condition exhibited significantly longer time reactions, these findings raise important questions about the advisability of using this relationship as a baseline condition in priming experiments that seek to identify genuine reaction times.

Key-words: Morphology, Semantic, Psycholinguistics, Japanese, Compounds

Rio de Janeiro
Janeiro de 2010

SUMÁRIO

1. Introdução.....	1
2. A Língua Japonesa	
2.1. Origens e estrutura.....	7
2.2. O sistema de escrita da língua japonesa.....	11
2.3. As palavras compostas na língua japonesa.....	15
3. O processamento das palavras	
3.1. Modelos Não-Decomposicionais e Decomposicionais.....	24
3.2. A técnica experimental <i>priming</i>	26
3.3. A interface entre a sintaxe e a semântica.....	28
3.4. Estudos anteriores: palavras derivadas.....	30
3.5. Estudos anteriores: palavras compostas.....	35
3.6. Estudos anteriores: o processamento morfológico do japonês.....	37
4. A Morfologia Distribuída	
4.1. A Morfologia Distribuída.....	41
4.2. As propriedades da Morfologia Distribuída.....	43
4.3. As listas distribuídas.....	44
4.4. A idiosincrasia semântica.....	49
4.5. O significado não-composicional especial.....	54
4.6. As palavras compostas.....	56
5. Os Experimentos	
5.1. Apresentação dos experimentos.....	60
5.2. Metodologia	
5.2.1. Material.....	61
5.2.2. Aparatos.....	64
5.2.3. Variáveis independentes e dependentes.....	65
5.2.4. Procedimentos.....	65
5.3. Experimento 1.....	66
5.3.1. Resultados dos tempos de respostas.....	67

5.3.2. Resultados dos índices de resposta.....	68
5.3.3. Discussão.....	69
5.4. Experimento 2.....	70
5.4.1. Resultados dos tempos de respostas.....	70
5.4.2. Resultados dos índices de resposta.....	72
5.4.3. Discussão.....	73
6. Conclusões.....	75
7. Referências.....	79

APÊNDICES

Palavras opacas.....	84
Palavras transparentes.....	85

LISTA DE FIGURAS

Figura	Descrição	Página
1	Evolução do ideograma de "cavalo".	11
2	Evoluções dos ideogramas de "em cima" e "embaixo".	11
3	Ideograma que representa “sol” em junção com o ideograma que representa “lua” forma o ideograma que representa “claro”.	12
4	Ideograma que representa “coração” participa como radical semântico em união com o ideograma de “tudo” que participa como radical fonético (leitura chinesa: <i>kai</i>) e formam o ideograma “arrependimento” (leitura chinesa: <i>kai</i>).	12
5	O uso misto dos diferentes sistemas de escrita em japonês.	13
6	A diferença no acesso lexical das palavras transparentes e opacas de acordo com Marslen-Wilson <i>et al.</i> (1994).	30
7	Condições experimentais do estudo de Maia, Lemle e França (2007).	33
8	A dicotomia entre o Léxico e a Sintaxe.	42
9	As listas da Morfologia Distribuída.	46
10	A primeira fase da derivação.	48
11	As listas distribuídas e os estágios de derivação.	49
12	O domínio dentro (<i>inner</i>) e for a (<i>outer</i>) da raiz.	51
13	Raiz concatenada ao categorizador nominal.	52
14	Raiz concatenada diretamente ao categorizador verbal.	52
15	Raiz concatenada ao categorizador verbal e, posteriormente, ao categorizador nominal.	53
16	A fronteira para o domínio de significado especial.	55
17	Palavra composta com significado composicional.	58
18	Palavra composta com significado não-composicional.	58
19	Exemplificação do experimento organizado em uma plataforma quadrado-latino.	63
20	Caixa de resposta do SuperLab.	64
21	Exemplificação dos procedimentos dos experimentos.	66
22	Exemplificação idealizada visualização dos experimentos.	73

LISTA DE GRÁFICOS

Gráfico	Descrição	Página
1	Médias dos tempos para as palavras transparentes e opacas no Experimento 1.	67
2	Médias dos tempos para as palavras transparentes e opacas no Experimento 2.	71

LISTA DE TABELAS

Tabela	Descrição	Página
1	Índices de resposta para as palavras transparentes e opacas no Experimento 1.	69
2	Índices de erros nas respostas entre as diferentes condições no Experimento 1.	69
3	Índices de resposta para as palavras transparentes e opacas no Experimento 2.	72
4	Índices de erros nas respostas entre as diferentes condições no Experimento 2.	72

1. Introdução

A partir da década de 1950, o campo dos estudos linguísticos conheceu uma série de mudanças importantes ocasionadas a partir da decadência da teoria behaviorista e do consequente reconhecimento da teoria gerativista. Esta última, liderada pelo linguista Noam Chomsky, diferenciou-se dos modelos anteriores primordialmente ao se preocupar com a diversidade sintática entre as línguas a partir de uma perspectiva mentalista computacional (Kato, 2002). Os pressupostos gerativistas esclareciam através dos modelos cognitivos que a linguagem é, de fato, um componente da mente humana representado fisicamente no cérebro. A partir deste momento, muitos estudos começaram a apresentar objetivos relacionados ao mapeamento da Faculdade da Linguagem e à compreensão sobre as capacidades específicas que um sujeito falante possui para adquirir e usar uma língua natural (Endo, 2007). Consequentemente, com a Teoria Gerativa, grande parte dos estudos sobre a linguagem passou a ter como foco principal o aspecto psicológico da linguagem, e não mais somente as propriedades absolutas das expressões linguísticas.

Entre as questões analisadas no campo da gramática gerativa no decorrer dos últimos 30 anos, um longo debate diz respeito à Hipótese Lexicalista, que propôs a separação completa entre a morfologia e a sintaxe, isto é, foi argumentado que o processo de formação de palavras e as operações sintáticas no nível da sentença não apresentam qualquer interação (Di Sciullo e Williams, 1987). Por conseguinte, os modelos de base Lexicalista (e.g. Chomsky, 1995), embora aceitem unidades menores do que as palavras, defendem que a derivação sintática só ocorre no nível sentencial e não no nível morfológico. Resumidamente, pode-se dizer que para os modelos desta natureza, o Léxico seria um repositório de palavras já prontas que “engatilhariam” o processo sintático ao formarem estruturas maiores.

Essa proposta foi considerada econômica, pois apresentava uma demanda reduzida do processamento sintático, já que este não ocorreria no nível das palavras; contudo, a proposta maximizava o trabalho da carga de memória. Os modelos não-lexicalistas, por outro lado, surgiram com propostas que defendiam o menor uso da carga de memória e a maior ocorrência de processamento linguístico. Estes modelos são em geral diferenciados entre os modelos híbridos (e.g. *Augmented Addressed Morphology*, Caramazza, Laudanna e

Romani, 1988; *Dual-Route Model*, Pinker, 1999) e aqueles que apostam na decomposição plena (e.g. *Distributed Morphology*, Halle e Marantz, 1993).

Os modelos híbridos embora ainda teorizem o armazenamento de algumas palavras, também propõem a decomposição morfológica para algumas palavras e, portanto, prevêm a importância da unidade morfológica no acesso lexical. Por outro lado, os modelos de decomposição plena, como a Morfologia Distribuída, apresentam uma postura mais radical, uma vez que predizem que a decomposição morfológica seja automática e geral para todos os vocábulos.

De acordo com a teoria da Morfologia Distribuída, parece mais econômico que o acesso lingüístico esteja relacionado aos mecanismos sintáticos do que simplesmente à carga de memória. Assim, a teoria propõe que o processamento sintático opera por fases desde os átomos lingüísticos (raízes e categorizadores) e lida inicialmente com unidades desprovidas de som que passam primordialmente por uma competição e uma seleção puramente morfológica, recebendo ao final de cada fase a inserção das peças de vocabulário, ou seja, do aparato fonológico. Posteriormente, ao final dos processos sintáticos e morfológicos, as palavras recebem seus significados na Enciclopédia, onde estão listados todos os significados arbitrários das estruturas armazenadas.

Desta forma, o termo morfema, para a Morfologia Distribuída e neste estudo, refere-se a um nó terminal sintático (ou morfológico) e seu conteúdo, mas não à expressão fonológica deste terminal. Portanto, morfemas são átomos de traços morfosintáticos-semânticos presentes no grupo oferecido pela Gramática Universal (Harley e Noyer, 1999). Esses traços semânticos, no entanto, são os traços mínimos, relevantes para a computação sintática. Nas palavras de Marantz (1996): “os nós terminais da sintaxe não contém, apropriadamente, traços fonológicos e traços semânticos irrelevantes sintaticamente”.

A teoria da Morfologia Distribuída, contudo, é bastante recente e vem debatendo intensamente com teorias diversas que buscam explicar o acesso e a representação das palavras de maneiras distintas. Por esta razão, nas últimas décadas os estudos sobre Morfologia vêm buscando respostas mais concretas através da realização de experimentos psicolingüísticos. Neste ponto, as línguas de natureza aglutinante oferecem perspectivas e desafios interessantes, pois apresentam processos produtivos tanto de adição de sufixos a

raízes verbais, como de composição. O japonês, mais particularmente, parece apresentar alguns fenômenos difíceis de serem explicados a menos que o processo de formação de palavras seja executado diretamente a partir do processamento sintático (Kageyama, 2002).

Nos últimos anos, vem se observando um grande crescimento dos estudos sobre o processamento da língua japonesa, que podem ser divididos entre os trabalhos que procuram analisar o processamento no nível de sentenças, e aqueles que avaliam o processamento das palavras. Os primeiros focam em questões tais como de que forma as ambiguidades sintáticas e semânticas são resolvidas, como as lacunas são preenchidas em uma sentença ou porque algumas sentenças são mais difíceis de serem processadas do que outras. Já os estudos no campo das palavras lidam com questões sobre o acesso e a representação da forma e do significado das palavras, e também sobre os diferentes sistemas de escritas e as rotas de processamento (Nakayama, 2002). Apesar de ser comum observar o emprego de formas derivadas nos estudos morfológicos, os estudos sobre o acesso e a representação das palavras em japonês utilizam em geral o uso de palavras compostas, uma vez que a língua japonesa apresenta um processo de composição bastante produtivo.

Em relação aos compostos, podemos observar que estes parecem estar em um nível intermediário entre as palavras e as sentenças: por um lado apresentam a união de mais de uma raiz; por outro lado demonstram uma fraca restrição sintática entre os elementos concatenados. Assim, o processo da composição oferece um meio efetivo e simples de se criar e transferir novos significados a partir de itens já existentes. Consequentemente, ao estudar os compostos, temos não só a oportunidade de examinar uma característica fundamental da morfologia, mas também de observar uma capacidade humana universal de criação de novos vocábulos (Libben e Jarema, 2005).

Ademais, os compostos também apresentam uma característica bastante pertinente entre os estudos morfológicos: o variado grau de transparência semântica. Em muitos casos, as palavras compostas apresentam significados semanticamente composicionais e, por isto, podem em princípio ser entendidas a partir da representação de seus constituintes. Todavia, é bastante frequente observar palavras compostas que não apresentam relação entre o significado da expressão como um todo e os significados de suas partes.

A falta da semântica dos constituintes no significado das palavras compostas tem sido matéria de grande debate nos estudos morfológicos ao longo das últimas décadas (Sandra, 1990; Zwitserlood, 1994; Fiorentino e Fund-Reznicek, 2009). Enquanto as teorias híbridas propõem a decomposição de palavras que apresentam significado transparente e o armazenamento como um todo das palavras que apresentam opacidade semântica, as teorias que seguem modelos de decomposição plena apostam que até mesmo as palavras com falta de composição semântica são decompostas morfológicamente.

Por outro lado, a Morfologia Distribuída considera que até mesmo as palavras simples são constituídas por átomos linguísticos e, portanto, precisam ser decompostas entre suas partes no processo morfológico. Além disso, também propõe que os significados idiossincráticos são designados tanto para expressões no nível de palavra, como no nível de sentenças (caso das expressões idiomáticas), ou seja, toda estrutura, até mesmo uma palavra simples contém um significado idiossincrático que precisa ser armazenado. Logo, apresentar ou não um significado composicional ou transparente é irrelevante para a decomposição morfológica: todas as estruturas da língua são decompostas em partes menores, uma vez que as menores unidades linguísticas são de natureza morfológica.

Uma metodologia já reconhecida como eficaz na busca da compreensão sobre o acesso às palavras é a metodologia de *priming*. Esta metodologia consiste na apresentação de uma palavra seguida de outra e permite analisar, em conjunto com a tarefa de decisão lexical, se a apresentação da palavra apresentada primeiro (*prime*) facilita ou inibe o acesso da palavra apresentada em seguida (*target* ou alvo). Assim, de acordo com o grupo de pares de palavras organizado, esta metodologia também possibilita a observação da participação de diferentes fatores como morfológico, semântico, ortográfico e fonológico. Adicionalmente, como veremos mais adiante, esta técnica também torna possível a tentativa de examinar a participação destes diversos fatores no curso temporal do processamento linguístico através do tempo de apresentação da palavra *prime*.

Desta forma, o presente estudo apresenta, através da realização e da análise de dois experimentos de *priming*, os seguintes objetivos:

- Investigar o acesso lexical das palavras compostas com dois ideogramas em Japonês, observando se a unidade morfológica é relevante ou não neste processo;
- Examinar a interface sintaxe-semântica ao comparar palavras compostas transparentes e opacas no japonês, questionando se o processamento morfológico apresenta restrições semânticas ou não;
- Analisar o curso temporal do acesso lexical utilizando dois tipos de tempos de apresentação da palavra *prime*;
- Avaliar o grupo controle empregado em uma metodologia de *priming* ao utilizar dois grupos controles diferentes.

Guiados pela teoria da Morfologia Distribuída, argumentamos a favor da unidade morfológica e do processo de decomposição inicial sem a interação de fatores semânticos. Por conseguinte, nossas hipóteses predizem efeitos de facilitação similares para as palavras transparentes e opacas quando a palavra *prime* for apresentada por um tempo curto. Contudo, esperamos padrões de respostas diferentes entre as palavras transparentes e opacas quando a palavra *prime* permanecer por um tempo mais longo. Estes resultados indicariam que no primeiro estágio do acesso lexical, apenas os traços sintáticos (morfológicos) seriam acessados e, portanto, não se observaria diferenças entre os dois tipos de palavra. Em contrapartida, em estágios posteriores os traços semânticos são acessados, resultando, em consequência, diferenças nos resultados do *priming*.

Em adição, de acordo com Forster, Mohan e Hector (2003), algumas vezes é argumentado que um grupo controle mais apropriado para um experimento de *priming* seria a apresentação de uma tela em branco no lugar da palavra *prime*, pois mesmo o uso de uma palavra não relacionada como *prime* poderia inibir a ativação do alvo, resultando, em consequência, uma superestimação em relação aos dados. Por esta razão, procuramos analisar melhor esta questão apresentando os dois tipos de grupo controle.

Propomos esta dissertação da seguinte maneira: o capítulo dois apresenta algumas informações estruturais e históricas básicas sobre a língua japonesa, assim como um pequeno detalhamento sobre a natureza dos compostos japoneses; o capítulo três expõe o debate sobre a dicotomia entre a Morfologia e a Sintaxe, juntamente com a menção de parte

da literatura na área do processamento das palavras e a interface sintaxe/semântica; o capítulo quatro apresenta as características principais da teoria da Morfologia Distribuída que temos como base no nosso trabalho; o capítulo cinco expõe os experimentos realizados neste estudo esclarecendo a metodologia utilizada, os procedimentos, os resultados e as discussões; por fim, propomos uma breve conclusão resumindo os resultados encontrados neste estudo no capítulo seis.

2. A Língua Japonesa

2.1. Origens e estrutura

Em contraste às outras línguas dos continentes Europeu, Asiático e das Américas, a língua Japonesa é falada em uma ilha e, deste modo, está fisicamente isolada das demais. Por conseguinte, não se observa uma “continuidade linguística” como na Europa ocorre entre as línguas de origem anglo-saxônica ou entre as línguas neolatinas, por exemplo. Além disso, ao contrário de grande parte das línguas européias, o japonês é basicamente falado apenas dentro dos confinamentos de seu território e não é primeira, nem segunda língua oficial de nenhum outro país exceto do próprio Japão (Shibatani, 1990).

A natureza exclusivista do japonês é acentuada diante do mistério em relação às suas origens: até hoje não se sabe ao certo à qual família linguística genética o japonês pertence. Desde os meados do século XIX, pesquisadores japoneses e estrangeiros vêm propondo uma larga variedade de teorias sobre este tópico. Entre estas teorias, um dos debates mais persuasivos diz respeito à conexão entre o japonês e o coreano, juntamente com a atribuição deste subgrupo à família das línguas altaicas (originárias do Norte Asiático). Contudo, nenhuma das teorias propostas até hoje apresentou justificativas suficientes para definir a afiliação do japonês com outras línguas. Shibatani (1990) argumenta que este mistério só será provavelmente solucionado a partir de estudos mais detalhados nos campos linguísticos das línguas altaicas, austronésias (originárias do Sudeste Asiático) e também do japonês antigo.

Em relação à tipologia, a língua japonesa apresenta a ordem dos constituintes SOV, ou seja, sujeito – objeto – verbo. Esta ordem vocabular é relativamente livre, excetuando-se a permanência do verbo ao final. Porém, de acordo com Yamashita (1997), a ordem canônica é ainda predominante e parece ocorrer 17 vezes mais do que as outras sentenças com movimentos de constituintes. Já a ordem dos outros elementos gramaticais também segue consistentemente o padrão modificador-núcleo.

No que diz respeito às características fonológicas do japonês, o dialeto de Tóquio é atualmente considerado a língua oficial. Sua forma acústica contém cinco vogais, um

simples grupo de consoantes e apresenta estrutura silábica básica CV. Também pode apresentar dois tipos de consoantes que fecham uma sílaba: 1) uma prolongação de uma articulação consonantal com as características determinadas pela consoante seguinte (ex: *hakkiri*); 2) uma consoante nasal que pode vir a fechar também uma palavra (ex: *shinbun*). Essas duas unidades são contadas como mora, assim como o prolongamento das vogais (ex: *ōkii*)¹. Na verdade, a mora é uma unidade tão importante quanto a sílaba no japonês: funciona como uma unidade rítmica na composição de poemas e, em muitos dialetos, pode apresentar também mudança de *pitch*. Desta forma, na estrutura fonológica japonesa uma palavra como *shinbun* contém quatro unidades (*shi-n-bu-n*).

Em muitos livros, é comum observar a definição do japonês como uma língua tonal. Porém, o sistema acentual japonês é particularmente distinto das línguas tonais prototípicas como o chinês. No caso das últimas, cada sílaba precisa ter o tom especificado, pois não há como prever o tom de cada sílaba de uma palavra ou de uma frase a partir de outro artifício. No entanto, este não é o caso do japonês, que permite prever através de um grupo de regras internas o *pitch* de cada sílaba e frase. Esta diferença entre o japonês e o chinês requisitou certa adaptação na fonologia dos empréstimos lexicais. Justamente a falta de contrapartes tonais no japonês para os tons chineses resultou na perda das marcas tonais distintivas nas palavras chinesas importadas, originando, em consequência, uma ampla homofonia entre as palavras sino-japonesas. A palavra “*kikan*”, por exemplo, com o mesmo *pitch* acentual pode significar “período”, “orgão” ou “instituição” de acordo com o contexto em que se encontra.

A partir da história do Japão, sabemos que o contato com a cultura chinesa trouxe não só um grande enriquecimento nas artes e na religião, como também na linguagem. O empréstimo sistemático das palavras chinesas parece ter ocorrido em três etapas. A primeira etapa iniciou-se no período pré-Nara (antes de 710 d.C.) quando o Budismo foi introduzido no Japão. A segunda ocorreu já durante o período Nara (início em 710 d.C.) momento em que muitos estudantes e oficiais da Corte foram estudar em duas grandes

¹ O sistema Hepburn (tradicional e revisado) de transcrição de caracteres japoneses para caracteres romanos estipula, nas palavras de origem chinesa e japonesa, o uso do sinal diacrítico para o alongamento das vogais *o* e *u*, enquanto que a vogal longa *e* é escrita *ei*, e a vogal longa *i* é escrita *ii*. No caso das vogais longas presentes em palavras estrangeiras, estas são todas escritas com o uso do sinal diacrítico.

metrópoles da dinastia Tang. Na transição entre os períodos Nara e Heian (início em 794 d.C.) ocorreu também um dos grandes acontecimentos da língua japonesa que intensificou o empréstimo lexical: a aquisição do sistema de escrita chinês. Posteriormente, durante o século XIV uma terceira etapa aconteceu quando novas palavras foram introduzidas pelos seguidores do Zen Budismo.

Estes empréstimos foram realizados não só em diferentes momentos, mas também a partir de localidades variadas da China e, portanto, a partir de diferentes dialetos. É por esta razão que, atualmente, a maior parte dos ideogramas japoneses apresenta mais de uma leitura, refletindo os diversos empréstimos lexicais, bem como o próprio vocabulário nativo. Entretanto, esses fatos históricos são muitas vezes desconhecidos pelos falantes nativos de japonês, que aprendem os ideogramas e as diversas leituras sem ter conhecimento dos fatores etimológicos.

Ao observar o empréstimo das palavras chinesas no japonês, é natural que muitos estudiosos relacionem este ao empréstimo do vocabulário latino à língua inglesa (Shibatani, 1990; Kess e Miyamoto, 1999). Na verdade, tanto o nível de formalidade quanto a proporção parecem ser bastante similares entre as palavras sino-japonesas e as palavras latinas no inglês. Em relação ao prestígio, semelhantes às palavras latinas no inglês, as palavras sino-japonesas tendem a expressar conceitos abstratos e ocorrem em maior número no contexto acadêmico. Ueno (1980) aponta que a proporção de palavras latinas em inglês é de 55%, enquanto que as palavras de origem anglo-saxônica apresentam apenas 35%, destinando os 10% restantes aos outros empréstimos estrangeiros. Já o estudo feito pelo Instituto Nacional de Linguagem (*Kokuritsu Kokugo Kenkyūjo*) de 1964 observou em revistas científicas um total de 59,3% de palavras sino-japonesas, 36,7% de palavras nativas e 3,9% de outros estrangeirismos e hibridismo. No entanto, o comportamento observado atualmente, tanto em relação às palavras latinas quanto às palavras sino-japonesas, é o de que tais empréstimos lexicais já se tornaram intrínsecos em cada língua e a etimologia original das palavras tornou-se apenas uma memória distante (Kess e Miyamoto, 1999).

Além dos empréstimos de origem chinesa, a língua japonesa também apresenta vocábulos de outras origens que foram introduzidos no contato com o mundo ocidental.

Inicialmente, em 1543 ocorreu o contato com a cultura européia quando um navio mercante português alcançou a ilha de Kyushu (sul do Japão). Neste momento, foram introduzidas não só as armas e o Cristianismo, como também um grande número de palavras. Em seguida, espanhóis e holandeses também chegaram às ilhas nipônicas e fizeram empréstimos linguísticos antes do fechamento do país no período Edo (início em 1603). Com a abertura do país no período Meiji (início em 1868), o inglês, o alemão e o francês ganharam mais espaço no ensino de língua estrangeira, ocasionando também acréscimos linguísticos. Porém, foi após a Segunda Guerra Mundial (1939 - 1945) que o empréstimo de palavras estrangeiras passou a vir em grande escala e quase totalmente dos Estados Unidos.

Esta sistemática introdução de vocábulos estrangeiros também resultou em novos procedimentos de formação de palavras e em uma grande diversidade em termos de estrutura morfológica das palavras compostas. Adicionalmente, do ponto de vista morfológico, o japonês é classificado como uma língua aglutinante, pois envolve um grande número de afixações na sua morfologia verbal e um mecanismo de formação de compostos naturalmente bastante produtivo. Toda esta diversidade faz com que o japonês ofereça uma perspectiva interessante nas questões sobre a representação e o acesso de palavras complexas (Joyce, 2002).

O sistema ortográfico do japonês também traz um acréscimo produtivo nos estudos linguísticos, uma vez que a maior parte dos estudos no campo do processamento da linguagem lida com a modalidade escrita. A observação do processo morfológico a partir da leitura pode apresentar discussões sobre o reconhecimento visual das palavras e a universalidade ou não deste processo entre os diversos sistemas de escrita. Assim, o fato da língua japonesa apresentar um sistema ortográfico misto (originário e derivado da escrita Chinesa), com características tão intrigantes, faz com que a observação do processo morfológico na modalidade escrita em japonês apresente ainda mais relevância. Apesar deste trabalho não se aprofundar nas questões acerca dos diferentes sistemas de escrita, acreditamos que seja interessante fazer uma breve explicação sobre os sistemas de escrita do japonês, já que trabalhamos com o reconhecimento visual das palavras nos experimentos que serão apresentados posteriormente.

2.2. O sistema de escrita da língua japonesa

A escrita é uma das mais significantes realizações culturais do ser humano, pois torna possível gravar e transmitir informações à distância, não só no espaço, mas também no tempo. Todavia, não se pode confundir escrita com linguagem. Afinal, por muito tempo na história da humanidade a maioria das pessoas era iletrada. Ainda hoje é possível observar pessoas iletradas ao redor do mundo que levam vidas produtivas e satisfatórias (Rogers, 2005). Deste modo, de um lado temos a linguagem que é adquirida automaticamente e faz parte de um sistema inato e complexo do cérebro humano, possibilitando a produção e a compreensão de declarações. Do outro lado temos a escrita que consiste no uso de marcas gráficas para representar declarações e que precisa ser conscientemente aprendida.

No Japão, o sistema de escrita utilizado foi parcialmente importado da China e os ideogramas receberam o nome de *kanji* (“kan” é originário da pronúncia da palavra Han, como era chamado o povo chinês devido à Dinastia Han, já o morfema “ji” significa “letra”, ou seja, “letra do povo Han”). Os *kanji* são tradicionalmente classificados em quatro grupos de acordo com a sua formação:

- *Kanji* ideográfico: versão estilizada de um objeto ou um conceito;



Figura 1: Evolução do ideograma de “cavalo”.

- *Kanji* diagramático: apresenta uma relação lógica, geométrica ou conceitual;

☺	=	⊥	上
☹	=	⊥	下

Figura 2: Evoluções dos ideogramas de “em cima” e “embaixo”.

- *Kanji* de composição semântica: apresenta dois ou mais caracteres unidos para representar o valor adicional de seus significados;

日 + 月 = 明

Figura 3: Ideograma que representa “sol” em junção com o ideograma que representa “lua” forma o ideograma que representa “claro”.

- *Kanji* fonético-semântico: contém um radical semântico com a aproximação do significado e um radical fonético que sugere a leitura chinesa (em geral, o radical semântico se encontra do lado esquerdo e o fonético do lado direito).

心 + 悔 = 悔

Figura 4: Ideograma que representa “coração” participa como radical semântico em união com o ideograma de “tudo” que participa como radical fonético (leitura chinesa: *kai*) e formam o ideograma “arrependimento” (leitura chinesa: *kai*).

Entretanto, o sistema de escrita japonesa é considerado misto, pois apresenta dois outros padrões em que a unidade básica é a mora. Estes dois sistemas, chamados de *kana*, foram criados a partir dos caracteres chineses e cada um apresenta 46 símbolos básicos e 25 marcas diacríticas adicionais. Apesar de ambos terem sido criados a partir dos caracteres chineses, possuem nomes distintos (*katakana* e *hiragana*) e apresentam diferenças na forma: o *katakana* é escrito com traços retos; já o *hiragana* é escrito de forma arredondada. Cada símbolo de *katakana* ou *hiragana* apresenta apenas uma leitura, ou seja, os *kana* apresentam relação de um para um com os sons designados.

Por outro lado, os diversos *kanji* possuem mais de uma leitura refletindo os diferentes empréstimos, além das palavras nativas. O ideograma 行, por exemplo, recebe a leitura japonesa “*i*” quando vem acompanhado de morfemas gramaticais escritos em *hiragana* (ex: 行く, *iku*), mas pode apresentar outras três leituras chinesas distintas, como pode ser visto nas palavras: 行儀 (*gyōgi*), 行進 (*kōshin*) e 行脚 (*angya*). Infelizmente, não

existe uma regra sistemática que prediz o uso das leituras. Em geral, as palavras escritas com mais de um *kanji* são lidas com as leituras chinesas e aquelas que vêm acompanhadas de morfemas em *hiragana* recebem a leitura japonesa, embora esta regra não seja sempre válida.

Ultimamente, os *kanji* são utilizados para as palavras com conteúdo; já o *hiragana* apesar de também ser utilizado para escrever algumas palavras nativas, é geralmente utilizado para os morfemas gramaticais como partículas e flexões verbais. Já o *katakana* serve para escrever as palavras estrangeiras, certas onomatopéias e, às vezes, também é utilizado para demonstrar ênfase. Em adição a esses sistemas de escrita, o alfabeto romano, chamado *romaji*, também é utilizado, mas em menor quantidade. Portanto, o japonês apresenta basicamente quatro formas de escrever sua língua. Assim, uma palavra como montanha, por exemplo, pode ser escrita de quatro maneiras:

Kanji: 山

Hiragana: やま

Katakana: ヤマ

Romaji: yama

Ainda que o domínio no uso de cada um desses sistemas de escrita seja bastante distinto dentro da língua japonesa, os quatro tipos podem ser encontrados em apenas uma sentença, como pode ser visto abaixo (fig.5).²

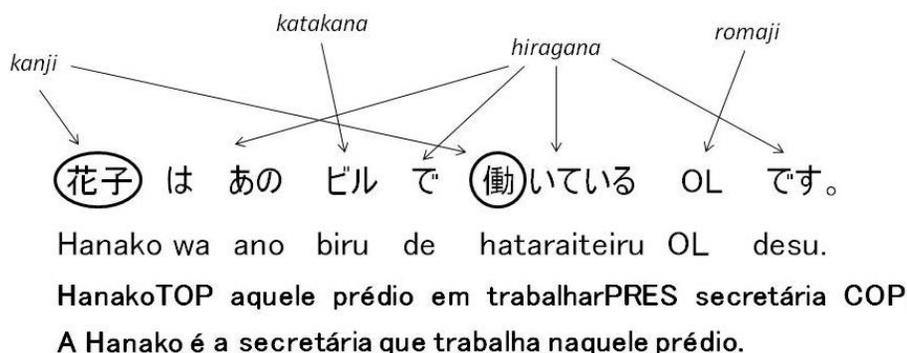


Figura 5: O uso misto dos diferentes sistemas de escrita em japonês.

² TOP = marca de tópico; PRES = marca de presente; COP = cópula.

Algumas outras características que tornam a escrita japonesa tão diferente da ocidental são a escrita na vertical, da direita para a esquerda e o manuseio dos livros de trás para a frente a partir da visão Ocidental. São encontradas atualmente, tanto a escrita na vertical como na horizontal, contudo a maneira mais formal ainda é a na vertical como se vê em artigos de jornais e em cartas formais (Shibatani, 1990).

Devido à complexidade da escrita causada pelo uso dos quatro sistemas e pelas diferentes leituras dos ideogramas, muitos movimentos tentaram abolir a escrita chinesa no Japão. Porém, estes movimentos não tiveram sucesso e, de acordo com a opinião de alguns autores (Shibatani, 1990; Kess e Miyamoto, 1999), a utilização dos ideogramas chineses provavelmente permanecerá. O governo japonês, entretanto, simplificou a forma de grande parte dos ideogramas e limitou o número dos *kanji* utilizados mais frequentemente. A lista de ideogramas básicos proposta pelo governo japonês, revisada em 1981, contém 1945 ideogramas e é utilizada como limite para a grande maioria dos jornais, além de servir como base para o ensino da educação primária e secundária no país.

Como foi mencionado anteriormente, diante da incrível diversidade da escrita japonesa, muitos estudos sobre processamento lexical apresentam questões também sobre o sistema ortográfico, pois, uma vez que muitos experimentos psicolinguísticos envolvem a leitura, a discussão sobre o reconhecimento visual da palavra é bastante pertinente. Em especial, os estudos sobre o processamento das palavras na modalidade escrita em japonês geralmente apresentam pelo menos uma das seguintes questões:

- 1) Em relação à correspondência entre a forma e o som nos ideogramas: há competição entre as diferentes leituras que um *kanji* apresenta? Existe preferência sobre alguma leitura?
- 2) Em relação à informação semântica do *kanji*: a forma visual do *kanji* pode levar ao acesso do significado da palavra diretamente sem passar pelos traços fonológicos?
- 3) Na língua japonesa, os compostos fazem parte de mais de 50% das palavras nos dicionários. Deste modo, pergunta-se: as palavras compostas são reconhecidas e processadas como unidades integradas ou ocorre o reconhecimento dos constituintes individualmente?

A resposta para cada uma dessas questões poderia não só explicar melhor o processamento da língua japonesa, como também ajudar a entender o mecanismo humano de reconhecimento de palavra. O presente estudo pretendeu analisar somente a terceira questão apresentada antes, ao investigar o papel dos constituintes morfológicos e a integração destes com os fatores semânticos no processamento das palavras compostas. Em seguida, apresentaremos uma breve explicação sobre a formação de compostos na língua Japonesa.

2.3. As palavras compostas na língua Japonesa

De acordo com alguns autores (Shibatani,1990; Libben e Jarema, 2005), a composição é o processo mais produtivo de criação de novas palavras. Particularmente, em japonês, a composição é um processo bastante rico, pois combina quase todas as categorias gramaticais. Deste modo, muitas palavras comuns do japonês moderno são representadas por um composto polissilábico de dois a quatro *kanji* com leituras chinesas. Ademais, como vimos anteriormente, o japonês apresenta uma vasta gama de palavras importadas que combinam-se com palavras nativas e de origem sino-japonesa de forma bastante livre. Vejamos a seguir alguns exemplos de formação de compostos de acordo com suas origens, retirados de Shibatani (1990):

- Compostos derivados de palavras nativas;

- a) N + N³ 秋空
 akizora
 outono-céu (céu de outono)
- b) A⁴ + N 近道
 chikamichi
 perto-rua (atalho)

³ N = Nome; A = raiz para Adjetivo ou Advérbio, V = Verbo.

⁴ Shibatani (1990) não identifica a qual categoria a sigla A se refere: se a de adjetivo ou se a de advérbio. Em Japonês, uma mesma raiz, como *chika* (perto), torna-se adjetivo com o acréscimo do morfema *-i*, ou torna-se advérbio com o acréscimo do morfema *-ku*. Pode até mesmo tornar-se um nome com o acréscimo do morfema *-sa*. No entanto, as palavras concatenadas na formação dos compostos, em geral, não apresentam os morfemas categorizadores realizados fonologicamente.

c) V + N 飲水
nomimizu
 beber-água (água de beber)

- Compostos derivados de palavras sino-japonesas;

a) N + N 父母
fubo
 pai-mãe (pais)

b) A + N 高利
kōri
 alto-interesse (interesse alto)

c) V + N 止血
shiketsu
 parar-sangue (parar de sangrar)

- Compostos híbridos, ou seja, junções de palavras de diferentes origens;

a) Sino-japonesa + Nativa 台所
daidokoro
 lugar-lugar (cozinha)

b) Nativa + Sino-japonesa 時計
tokei
 tempo-medida (relógio)

c) Sino-japonesa + Estrangeira 石油ストーブ
sekiyusutobu
 óleo-fogão (fogão à óleo)

- c) V + V 走り疲れ
hashirizukare
 correr-cansar (cansaço causado por correr)
- d) N + A 腹痛
haraita
 barriga-dor (dor de barriga)
- e) A + A 高低
takahiku
 alto-baixo (altos e baixos)
- f) V + A 切れ長
kirenaga
 cortar-longo (fenda comprida)

A grande maioria das palavras compostas resultantes a partir destas formações está nas categorias de nomes ou nomes verbais. A categoria de “nome verbal” é uma classificação especial para um nome que tem aspecto verbal, mas para se tornar verbo requer o acréscimo de um morfema verbal *-suru*. Em geral, estas palavras são de origem sino-japonesa (ex.1), mas é possível também encontrar palavras nativas (ex.2) e estrangeiras (ex.3).

- 1) *kenkyū* (pesquisa) – *kenkyū suru* (pesquisar)
- 2) *ajitsuke* (tempero) – *ajitsuke suru* (temperar)
- 3) *kopii* (cópia) – *kopii suru* (copiar)

Além disso, a ordem dos elementos na grande maioria dos compostos japoneses reflete a ordem normal dos elementos na sentença, ou seja, modificador-núcleo (exs. 4, 5 e 6).

Sintagma Simples	Palavra Composta
4) 秋の空	秋空
<i>aki no sora</i>	<i>akizora</i>
outono de céu	outono-céu
(céu de outono)	(céu de outono)
5) 渡る鳥	渡り鳥
<i>wataru tori</i>	<i>wataridori</i>
migrar pássaro	migrar-pássaro
(pássaro que migra)	(pássaro migratório)
6) ハワイで焼ける	ハワイ焼け
<i>Hawai de yakeru</i>	<i>Hawaiyake</i>
Havaí em bronzear	Havaí-bronzear
(bronzear no Havaí)	(bronzamento Havaiano)

Neste ponto, é possível identificar uma diferença entre os compostos de origem nativa e aqueles de origem sino-japonesa: como o chinês apresenta uma ordem vocabular VO, os compostos chineses apresentam ordem V-N (ex.7), enquanto que os compostos japoneses apresentam ordem N-V (ex.8).

- 7) 着色
- chakushoku*
- aplicar-cor (aplicação de cor)
-
- 8) 色付け
- irozuke*
- cor-aplicar (aplicação de cor)

Contudo, de acordo com Shibatani (1990), o padrão sino-japonês V-N já está fortemente enraizado na língua Japonesa, pois até a formação de palavras compostas recentes é muitas vezes realizada desta maneira. Os compostos transitivos longos, no entanto, ainda são escritos em geral com a ordem básica dos constituintes no japonês (exs. 9 e 10).

9) 輸出制限

Yushutsuseigen

Exportação-limite (limite de exportação)

10) 物価統制

Bukkatōsei

preço-controle (controle de preço)

Há também dois tipos adicionais de compostos em japonês que não são frequentemente observados nas línguas européias, são eles: 1) os compostos *dvandva* que expressam relação coordenada entre os elementos (ex.11); 2) os compostos verbais, que podem ser combinações de nome + verbo (ex. 12), advérbio + verbo (ex.13) ou verbo + verbo (ex.14).

11) 男女

danjo

homem-mulher (homem e mulher)

12) 勇気付ける

yūkizukeru

coragem-afixar (encorajar)

13) 近寄る

chikayoru

perto-chegar (aproximar)

14) 噛み切る

kamikiru

morder-cortar (cortar mordendo)

Além das questões estruturais, ainda há muitas questões sobre a semântica dos compostos que precisam ser mais bem estudadas, como por exemplo, a relação semântica entre os elementos do composto e o significado total da palavra final. Visto que o significado de cada ideograma é bastante estável, em geral, as palavras ou os morfemas mantêm seus significados ao se concatenarem com outros elementos e, em consequência, originam palavras compostas com significados transparentes e composicionais. Por esta razão, é possível, em muitos casos, adivinhar o significado de uma palavra composta mesmo que seja a primeira vez que se a encontre (ex.15).

15) 閉店

heiten

fechar-loja (fechamento da loja)

Porém, vale ressaltar que muitas vezes as palavras constituintes não demonstram tal transparência semântica. Os morfemas finais *-nin* e *-sha* em compostos japoneses, por exemplo, têm o significado de “pessoa” e designam, normalmente, o agente de uma ação que é denotada pelo primeiro elemento do composto (ex.16). No entanto, em algumas situações, esta composição entre “ação” e “agente” não parece ser tão clara (ex.17).

16) 案内人

annainin

guiar-pessoa (guia)

17) 使用人

shiyōnin

usar-pessoa (empregado)

O grau de transparência semântica entre as palavras constituintes e os compostos resultantes não é um assunto novo no meio linguístico, e nem exclusividade da língua japonesa. Pelo contrário, como foi discutido na introdução e posteriormente melhor explicitado neste trabalho, muitos estudos vêm tentando observar a relação semântica entre as partes e o todo das palavras. A língua japonesa, por apresentar um processo de composição extremamente rico, torna possível observar diferentes graus de transparência semântica em suas palavras compostas. Os exemplos a seguir procuram demonstrar tal afirmação (veja também os Apêndices um e dois).

Palavras que apresentam significado transparente:

18) 男子

danshi

homem-criança (menino)

19) 食品

shokuhin

comer-produto (alimento)

20) 外国

gaikoku

fora-país (estrangeiro)

Palavras que apresentam significado opaco:

21) 小説

shōsetsu

pequeno-teoria (romance)

22) 文化

bunka

frase-transformação (cultura)

23) 目次

mokuji

olho-próximo (índice)

A partir dos exemplos anteriormente citados, observa-se que a língua japonesa possui uma larga idiossincrasia na composição dos significados de seus compostos. A idiossincrasia semântica foi por muito tempo uma justificativa da corrente Lexicalista para definir que as palavras eram armazenadas inteiramente e que não eram abertas aos mecanismos sintáticos. Todavia, a concepção tradicional do Léxico como um depósito de palavras tem sido reanalisada nas últimas décadas voltando maior atenção às palavras e suas estruturas internas. Vejamos, em seguida, os modelos teóricos propostos e alguns estudos realizados anteriormente que buscaram analisar estas questões.

3. O processamento das palavras

3.1. Modelos Não-Decomposicionais e Decomposicionais

Por muito tempo, os linguistas trabalharam observando fenômenos em estruturas do tamanho de sentenças, relegando uma importância menor às palavras. Entretanto, parece bastante intuitivo que conhecer uma língua envolve conhecer as palavras desta língua⁵. Consequentemente, com o passar dos anos e o avanço das teorias linguísticas, algumas propostas em relação ao armazenamento e ao acesso às palavras começaram a ganhar relevância. Entre estas, dois tipos de modelos são em geral observados: Lexical ou Não-Decomposicional, e Decomposicional ou Não-Lexical. Os primeiros postulam que as formas lexicais são armazenadas inteiramente, por conseguinte, argumentam em favor de maior trabalho da carga de memória e menor participação de processos computacionais. Por outro lado, os modelos de base Decomposicional estão no extremo oposto, pois propõem que todas as formas complexas passam por um processo de decomposição, ou seja, defendem uma maior demanda de computação e a redução de armazenamento.

Entre estes extremos teóricos, há também os modelos híbridos que assumem tanto a representação lexical das palavras inteiramente, como o armazenamento separado dos constituintes. Alguns modelos desta natureza defendem um mecanismo de duas rotas paralelas (e.g. Schreuder e Baayen, 1995; Pinker, 1999); isto é, de acordo com a palavra, esta pode ser decomposta entre suas partes ou acessada inteiramente. Contudo, outros modelos apostam que o mecanismo *default* é o acesso lexical inteiro, enquanto que a decomposição é somente realizada no encontro de palavras novas (e.g. Caramazza, Laudanna e Romani, 1988; Niemi, Laine e Tuominen, 1994).

Na verdade, todos esses modelos propõem a existência da informação morfológica; todavia, se diferenciam no grau de importância relegado a este tipo de informação. Enquanto os modelos lexicalistas argumentam que a informação morfológica está presente dentro da palavra e é armazenada juntamente como um bloco, os modelos de base

⁵ "It's somehow intuitive to think that knowing a language involves knowing the words of the language" (Marantz, 2001, p.1).

Decomposicional sugerem um estágio individual para o processamento da morfologia (Chialant e Caramazza, 1995). Por exemplo, de acordo com a teoria Lexicalista, a marca de plural da palavra “livros” no português é armazenada juntamente com toda a informação (fonológica, morfológica e semântica) relevante da palavra. No entanto, do ponto de vista Decomposicional, a palavra não é a menor unidade linguística; pelo contrário, a informação morfológica também é uma unidade relevante para o processamento e, por isto, a marca de plural e as outras informações morfológicas serão processadas separadamente dos outros traços presentes em uma palavra.

Embora o extensivo uso da Teoria Lexicalista como modelo base durante muitos anos, as análises mais recentes sugerem que a visão não-decomposicional parece ser um equívoco. Inicialmente, sabemos que a produtividade é uma das propriedades mais memoráveis da linguagem: um pequeno número de elementos é re combinado de diferentes maneiras e forma um número ilimitado de palavras e sentenças. Adicionalmente, as línguas demonstram bastante sistematicidade entre as regras de derivação e flexão. Assim, além de bastante redundante, parece pouco provável que seja necessário o armazenamento de cada combinação individualmente. Shoolman e Andrews (2003) ainda afirmam que um estudante de Ensino Médio adquire em torno de 80.000 palavras durante seus anos de estudo, esta quantidade, porém, parece ser impossível de ser adquirida caso cada vocábulo solicitasse um armazenamento individual.

Por esta razão, nos últimos 30 anos, um extenso número de investigações empíricas vem sendo realizado na tentativa de analisar as variadas hipóteses propostas. Uma técnica bastante reconhecida e utilizada nos estudos morfológicos atualmente é a técnica de *priming* que se caracteriza pela apresentação simultânea de dois estímulos e a observação sobre o efeito que o primeiro estímulo pode causar no segundo. A introdução desta técnica costuma ser atribuída a Meyer e Schvaneveldt (1971), mas foi a partir dos experimentos realizados por Taft e Foster (1975) que a técnica recebeu maior atenção, quando os autores argumentaram em favor da decomposição on-line de formas complexas a partir dos efeitos de constituição morfológica.

Posteriormente, entretanto, o estudo de Butterworth (1983) propôs que um processo de decomposição geral não seria possível, já que as palavras admitem grande número de

idiosincrasias e, portanto, deveriam ser armazenadas inteiramente. Consequentemente, ambos os estudos geraram um grande debate e a profusão de muitas pesquisas com base experimental procurando investigar as diversas teorias. Uma vez que também utilizamos a técnica de *priming* neste trabalho, gostaríamos de explicitá-la melhor antes de apresentar parte da literatura dos estudos sobre palavras na interface sintaxe-semântica.

3.2. A técnica experimental *priming*

É natural imaginar que o acesso a uma palavra pode facilitar o acesso de uma segunda palavra que contenha semelhança no significado ou na forma. Entretanto a pergunta que ainda se faz nos estudos sobre palavras é: a partir de que fator ou a partir de quais fatores esta facilitação foi gerada? Os estudos que utilizam a técnica de *priming* vêm buscando responder tais perguntas ao investigar os processos de percepção e do reconhecimento das palavras.

Para tal finalidade, na organização de um experimento de *priming*, o experimentador necessita construir uma lista de estímulos com palavras em pares: a primeira palavra é chamada *prime*, enquanto a segunda é chamada alvo (*target*). Como explicita França *et al.* (2008), estas palavras podem apresentar relação semântica (médico – enfermeira), fonológica (cara – carinho), morfológica (comércio – comerciante), ortográfica (pente – lente) ou sintática (fazia – tinha). É preciso também que se crie um grupo controle para servir como base para comparação. O grupo controle ideal em um experimento de *priming* ainda tem sido matéria de discussão entre as pesquisas; todavia, os mais utilizados atualmente são: pares de palavras que não apresentam qualquer relação entre si (mesa – frio) ou palavras alvos apresentadas sem *prime*, ou seja, antecedidas apenas por uma tela em branco.

Para medir a reação dos voluntários nos experimentos psicolinguísticos, os participantes realizam uma determinada tarefa em relação ao estímulo construído. Esta tarefa, em geral, não possui relação com o objetivo real da pesquisa, pois é ideal que o participante esteja desatento às reais intenções da pesquisa e que não crie, deste modo, expectativas sobre o estímulo construído. Assim, no caso dos estudos com palavras, uma

tarefa bastante realizada é a tarefa de decisão lexical que solicita que o sujeito julgue se a palavra alvo é realmente uma palavra ou não. Por esta razão, é preciso que a lista de palavras controladas seja mesclada também com um grupo de pares com não-palavras (feliz – mreco). Além disso, todas as palavras são apresentadas em ordem aleatória.

Os tempos de resposta, assim como a acuidade, podem demonstrar a influência que o *prime* exerceu em relação à palavra alvo. Por exemplo, em um par sem nenhuma relação como “mesa – frio”, é pouco provável que a palavra *prime* “mesa” ajude no acesso da palavra alvo “frio”. No entanto, no caso de palavras que apresentam relação semântica, é possível que a palavra “médico” facilite o acesso à palavra alvo “enfermeira”. O estudo de Gomes (2009), porém, propõe que mesmo duas palavras associadas semanticamente são mediadas por uma estrutura sintática que surge espontaneamente quando as palavras são apresentadas seguidamente. Esta última análise condiz com os estudos mais recentes que vêm procurando observar efeitos de facilitação a partir de fatores sintáticos/morfológicos, como veremos adiante.

A técnica de *priming* oferece também variações de acordo com a modalidade em que as palavras são apresentadas ou em relação ao tempo em que a palavra *prime* permanece na tela. Em geral, muitos experimentos apresentam as palavras (tanto *prime* como alvo) na modalidade escrita. Contudo, a análise a partir destes experimentos esbarra em questões ortográficas, pois se pergunta até que ponto os efeitos de facilitação são de natureza puramente morfológica e não ortográfica. Ademais, é muito relevante também a análise sobre o processamento morfológico à parte do sistema de escrita. Por esta razão, passou-se a utilizar igualmente o *cross-modal priming* ou *priming* inter-modal, em que uma palavra é ouvida e a outra é lida. Adicionalmente, outra maneira utilizada para tentar eliminar o efeito ortográfico é a apresentação das palavras com a fonte ou o tamanho das letras diferenciado.

Uma segunda característica importante nas variações dos estudos de *priming* diz respeito ao tempo de permanência da palavra *prime*, o qual é chamado de SOA (*stimulus onset asynchrony*). Estudos anteriores como Rastle *et al.* (2000) e Pylkkänen e Okano (a aparecer) predizem que se a palavra *prime* for apresentada por tempos mais curtos que 40ms (milésimos de segundo), ou seja, não vista conscientemente, esta palavra será

processada apenas estruturalmente, pois não há tempo suficiente para processar os traços semânticos. Por outro lado, se esta palavra permanecer na tela por um tempo maior, como 100ms, traços semânticos provavelmente poderão ser acessados, influenciando a decisão sobre uma palavra apresentada posteriormente. Além disso, se o tempo de permanência do *prime* for bastante extenso como 2000ms, o processo será mais reflexivo (Joyce, 2002), o que não é ideal em um experimento de *priming*, pois o teste procura visualizar a influência dos fatores linguístico no acesso automático e reflexo (e não reflexivo).

Através do SOA da palavra *prime*, portanto, pode-se tentar analisar também o curso temporal do processamento da linguagem. Se a apresentação durante apenas 40ms de uma palavra pode facilitar uma palavra relacionada morfológicamente, isto pode indicar que o acesso morfológico ocorre nos primeiros milissegundos do processamento lexical. Por outro lado, se um experimento com um SOA mais longo (100ms, 150ms) demonstrar facilitação apenas entre os pares relacionados semanticamente, o resultado apontaria para a ocorrência do processamento semântico em um estágio um pouco posterior.

Por fim, pode-se ressaltar que alguns experimentos neurolinguísticos também vêm utilizando a técnica de *priming* em conjunto com aparelhos neurológicos como MEG (*magnetoencephalogram*) e EEG (*electroencephalogram*). No caso dos estudos Psicolinguísticos com objetivos morfológicos que utilizam a técnica de *priming*, dois fatores têm sido considerados de muita importância na organização dos estímulos; são eles: os fatores semânticos e os fatores morfológicos. Uma vez que a interface sintaxe e semântica não foi completamente esclarecida, muito ainda se pergunta até que ponto um fator influencia o outro ou quais são os momentos de participação de cada fator no processamento linguístico. O fator semântico, em especial, apresenta grandes questões para o processamento morfológico, pois, muitas vezes, os domínios morfológicos não se estendem igualmente aos domínios semânticos e vice-versa.

3.3. A interface entre a sintaxe e a semântica

A questão da idiosincrasia semântica vem sendo rica matéria de debate nos estudos sobre processamento lexical, pois a idéia de decomposição morfológica parece se esvaír

quando os morfemas não apresentam os significados comuns. Por exemplo, porque uma palavra como “restaurante” seria decomposta entre a raiz verbal de “restaurar” e o afixo “-ante”, se a palavra como um todo não se refere mais à sua raiz? Logo, analisar como o sistema lida com a opacidade semântica poderia trazer novas perspectivas sobre o sistema cognitivo, observando não só como ocorre a representação das palavras, mas também de que forma fatores semânticos e morfológicos exercem impacto no processo computacional de produção e compreensão on-line da linguagem (Libben e Jarema, 2005).

Algumas teorias híbridas argumentam que as palavras transparentes semanticamente (ex: comerciante) são decompostas entre seus constituintes, pois ainda estão fortemente relacionadas às suas raízes. Todavia, palavras opacas semanticamente (ex: restaurante) são armazenadas e acessadas por inteiro, pois não há mais relação entre o significado da raiz e o significado da palavra como um todo, mesmo que etimologicamente esta relação tenha existido. Para modelos desta linha, portanto, o fator semântico é crucial e norteia a decomposição morfológica.

Por outro lado, teorias que seguem a linha *full-parsing*, ou seja, “parseamento total” (como a Morfologia Distribuída) prevêm que a decomposição inicial é puramente morfológica, sem participação de fatores semânticos e, por isto, é total e geral. Fatores semânticos só participam posteriormente, isto é, a decomposição morfológica inicial é indiferente a transparência ou a opacidade semântica da palavra. Esta decomposição seria apenas um dos estágios presente no mecanismo do processamento linguístico. Um estágio subsequente seria a combinação morfológica que envolveria talvez, não só a ação de ligar as partes novamente, como também processos interpretativos e combinatórios.

Muitos estudos defendendo as diversas teorias procuram esclarecer a representação e o processamento lexical a partir da observação dos fatores morfológicos e semânticos. Cada um dos estudos que apresentaremos a seguir pressupõe teorias diferentes, mas apresentam um ponto em comum: prevêm que a menor unidade no processo linguístico é o morfema, logo, há um estágio de decomposição no processamento das palavras.

3.4. Estudos anteriores: palavras derivadas

O estudo de Marslen-Wilson *et al.* (1994) recebeu bastante relevo no meio da pesquisa morfológica ao propor que a transparência semântica seria um fator importante para determinar como uma palavra morfológicamente complexa é representada no nível lexical. O estudo organizou experimentos de *priming* com decisão lexical na tentativa de observar se a apresentação de uma palavra derivada facilitaria no acesso da palavra base apresentada em seguida. Foram utilizados pares que apresentavam transparência semântica como *government* (governo) e *govern* (governar); e opacidade semântica como *apartment* (apartamento) e *apart* (à parte).

Os resultados demonstraram que houve facilitação no encontro de pares como *government* – *govern*, mas não em pares como *apartment* – *apart*. Os autores argumentaram que no caso das palavras transparentes (como *government*), o significado é sincronicamente composicional, deste modo, as palavras derivadas ainda estão fortemente ligadas às palavras bases e demonstram facilitação uma vez que a raiz já foi acessada no encontro da primeira palavra. Por outro lado, as palavras opacas (como *apartment*), mesmo que historicamente tenham apresentado alguma relação semântica com a palavra base, sincronicamente não apresentam significado composicional e, por isso, são representadas individualmente e à parte das palavras bases. Isto é, os autores postulam que as palavras opacas alvo necessitam um acesso diferente daquele ocorrido na palavra *prime*, e, por conseguinte, não demonstram qualquer tipo de facilitação.

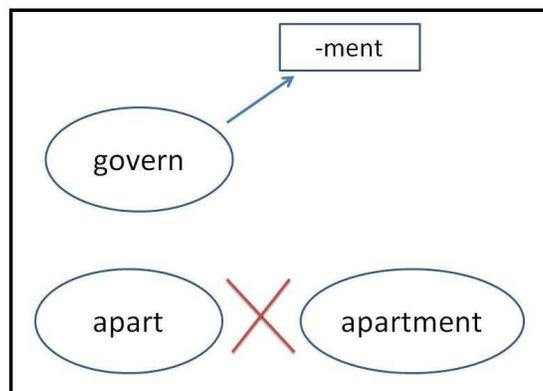


Figura 6: A diferença no acesso lexical das palavras transparentes e opacas de acordo com Marslen-Wilson *et al* (1994).

Em conseqüência, com esta análise foi argumentado que a transparência semântica é um fator importante na organização do léxico: uma palavra opaca não é mais relacionada à sua família morfológica e, por isto, é armazenada separadamente; já as palavras transparentes semanticamente são morfológicamente complexas, logo, sofrem decomposição. Entretanto, este estudo gerou grandes controvérsias no campo da morfologia e pesquisas posteriores tentaram demonstrar, em contraposição, que o processamento inicial da palavra não sofre influências de fatores semânticos, pois é exclusivamente morfológico.

Logtin, Segui e Hallé (2003) organizaram um estudo com palavras em francês seguindo parcialmente o estudo de Marslen-Wilson *et al.*(1994). Foi realizado também um experimento de *priming* apresentando palavras derivadas seguidas por suas bases. Assim como o estudo anterior, foram contrastados pares transparentes semanticamente como *fillete* (garota pequena) e *fille* (garota); e pares com opacidade semântica como *fauvette* (gorjeio) e *fauve* (gato selvagem). Além desses pares, Longtin, Segui e Hallé (2003) introduziram um novo grupo de palavras para o estímulo do experimento que o diferenciou do estudo anterior: um grupo de palavras “pseudo-derivadas”. Estas palavras têm como característica a decomposição em estruturas que parecem morfemas, mas na realidade não são nem sincronicamente, nem diacronicamente, formadas por tais morfemas. Por exemplo, a palavra *baguette* (pão francês) poderia ser decomposta em duas partes (*bague-tte*), contudo *baguette* não possui qualquer relação com o morfema *bague* (sino) e nem apresenta a semântica do afixo *-tte* (diminutivo em francês).

O experimento também apresentou um quarto grupo de palavras com semelhança ortográfica: aquelas que possuíam aparentemente um radical em comum, mas sem um afixo significativo na língua. Por exemplo, *-tte* é um sufixo diminutivo no francês (*fauvette*), mas *-cot* não é um sufixo. Assim, *abricot* não parece ser decomposto em *abri-cot*. Deste modo, o grupo de palavras com semelhança ortográfica apresentou pares como *abricot* (damasco) – *abri* (abrigo).

Os resultados do experimento demonstraram um efeito de facilitação significativo nos pares de palavras que apresentavam relação morfológica, mesmo que aparente: os pares com transparência semântica, opacidade e pseudo-derivação. Em contraste, os pares com relação ortográfica apresentaram maior demora na decisão lexical. Ou seja, os efeitos de

facilitação encontrados anteriormente não são decorrentes da semelhança ortográfica, nem de relação semântica, mas provavelmente derivados de aspectos puramente morfológicos. Os autores sugeriram que, inicialmente, um processo automático de decomposição é “engatilhado” pelas palavras que são morfológicamente complexas em sua estrutura superficial. Porém, somente se a palavra é verdadeiramente e sincronicamente complexa morfológicamente, ou seja, se sua composição morfológica é semanticamente transparente, a decomposição é útil e as unidades morfológicas podem permanecer ativadas.

Muito parecido com o estudo de Longtin, Segui e Hallé (2003) é o estudo de Rastle, Davis e New (2004), que também elaborou um experimento de *priming* com três diferentes grupos de palavras na língua inglesa: um grupo de pares transparentes semanticamente, um grupo de pares que apresentam relação morfológica mas não semântica (pseudo-derivadas e opacas, neste estudo, ficam no mesmo grupo) e um grupo de pares que possuem um radical idêntico mas a palavra derivada não apresenta um afixo significativo na língua (como o grupo ortográfico visto em Longtin, Segui e Hallé, 2003).

Os resultados encontrados por Rastle, Davis e New (2004) não demonstraram diferenças significativas entre a facilitação observada nos dois primeiros grupos de palavras, ou seja, nos grupos de pares morfológicamente relacionados. Apenas o grupo com condição ortográfica é que não mostrou efeitos de facilitação. Os autores argumentaram que o resultado aponta para um processo morfológico de segmentação que opera em qualquer palavra com uma raiz e afixos, independente da relação semântica que esta apresenta com suas partes. Ou seja, o reconhecimento inicial da palavra contém um estágio inicial de segmentação morfêmica, mas que, provavelmente, é seguido por estágios com participação de fatores semânticos.

O estudo de Maia, Lemle e França (2007) também demonstrou, através de dois paradigmas experimentais, efeitos de decomposição interna nas palavras do português brasileiro. No primeiro experimento, solicitava-se a identificação da cor de uma letra em um morfema presente de uma palavra. A palavra era apresentada dividida por duas cores, mas o corte entre as cores poderia ser morfêmico ou não (e.g. jornal-eiro X jornale-iro). Havia também um grupo de palavras com pseudo-morfemas, ou seja, palavras que apresentam uma parte com coincidência fonológica a um morfema (e.g. entrevista).

Um fator adicional também incluído no *design* desse experimento foi a verificação de eventuais diferenças em relação à renegociação de significado acarretada pela adição do sufixo à raiz, contrastando-se formas como “jornalista” com formas como “frentista”. Note-se que, no primeiro exemplo, o sufixo “-ista” tem sua computação feita tomando por base a palavra “jornal”, enquanto que em “frentista” o significado da palavra “frente” não é o ponto de partida da computação semântica causada pela introdução do sufixo “-ista”, embora as duas palavras compartilhem a raiz “frent-”. O objetivo do experimento era verificar em que medida no processo da leitura a identificação implícita do morfema no interior da palavra fonológica exerceria efeito de facilitação na realização da tarefa de identificação cromática (por exemplo, a cor da letra “i” da forma “-ista”).

Corte/Tipo	Morfema+Palavra MP	Pseudo-morfema PM	Morfema+Raiz MR
morfêmico C	MAL <u>I</u> NHA BALE <u>I</u> RA	ESP <u>I</u> NHA MADE <u>I</u> RA	CAN <u>I</u> NHA COPE <u>I</u> RA
Não morfêmico N	MAL <u>I</u> NHA BALE <u>I</u> RA	ESP <u>I</u> NHA MADE <u>I</u> RA	CAN <u>I</u> NHA COPE <u>I</u> RA

Figura 7: Condições experimentais do estudo de Maia, Lemle e França (2007)

Os resultados obtidos indicaram que os sujeitos reconhecem mais acertadamente e com maior rapidez a cor da letra alvo nas condições com recorte morfêmico, esteja o morfema em concatenação com uma palavra ou com uma raiz. Por outro lado, não se observou efeito de recorte cromático significativo, nem nos índices, nem nos tempos de decisão acertada, nas condições com pseudo-morfema. Esses resultados sugerem que os leitores utilizariam um procedimento de *parsing* morfológico pleno, isolando os morfemas que compõem uma palavra, quer esses morfemas estejam em relação de transparência, quer estejam em relação de opacidade com a base.

Já o segundo experimento investigou o rastreamento ocular das mesmas palavras, pretendendo identificar, preliminarmente, os pontos de fixação e as sacadas na primeira passagem do olhar, bem como nos movimentos regressivos. Os resultados sugeriram uma

correlação entre a computação morfológica no interior da palavra, os tempos de fixação médios e os índices de movimento sacádico. As palavras com sufixos concatenados apresentaram tempos de fixação médios mais elevados e, também, maiores índices de movimentos sacádicos do que palavras com pseudo-morfemas, confirmando parcialmente a hipótese de que a concatenação morfêmica requer maiores latências. Os resultados obtidos nos dois experimentos permitem reunir evidências para argumentar que no processo de leitura, palavras complexas são processadas morfológicamente, concatenando-se raízes a afixos, em contraposição aos modelos que postulam o acesso às palavras por inteiro.

O estudo de Garcia (2009) também propôs uma análise experimental sobre a constituição morfológica de palavras no português brasileiro. Foi utilizada a técnica de *priming* com decisão lexical e quatro diferentes tipos de relação entre as palavras *prime* e alvo: a) palavras morfológicamente relacionadas: fila/fileira; b) palavras apenas semanticamente relacionadas: ordem/fileira; c) palavras apenas fonologicamente relacionadas: filé/fileira; d) palavras não relacionadas. Os resultados demonstraram que os pares morfológicamente relacionados foram decididos significativamente mais rapidamente do que os outros pares, conseqüentemente, apontando para os modelos de base Decomposicional. A autora sugeriu a partir dos resultados que as palavras morfológicamente relacionadas repartem a mesma raiz e, portanto, apresentam um momento da derivação igual. A palavra “fileira”, por exemplo, passou pelo estágio “fila” em um momento da derivação. Por esta razão, quando a palavra “fileira” é apresentada após a palavra “fila”, há uma facilitação ocasionando menores tempos de decisão.

Além dos estudos relacionados acima, nos últimos 30 anos foram propostas muitas outras pesquisas sobre morfologia tentando observar o processamento lexical através de palavras flexionadas e derivadas. Por outro lado, em menor quantidade estão os estudos sobre as palavras compostas observando também a decomposição morfológica. Vejamos a seguir alguns desses estudos.

3.5. Estudos anteriores: palavras compostas

O estudo de Fiorentino e Poeppel (2007) apresenta, assim como os estudos anteriores, as questões sobre a realidade da decomposição morfológica. O estudo apresentou um experimento com decisão lexical e o uso simultâneo do aparelho MEG, utilizando como estímulo palavras divididas em diferentes grupos:

- Palavras compostas
Ex: *flagship* (*bandeira* + *navio* = *nau*)
- Palavras simples (com os tamanhos controlados e semelhantes ao dos compostos)
Ex: *crescent* (*meia-lua*)
- Não-Palavras com estrutura de palavras compostas, formadas por “palavra e não-palavra”
Ex: *crowstep*

Os resultados obtidos demonstraram diferenças significativas entre os grupos: as palavras compostas apresentaram tempos mais curtos de decisão do que as palavras simples que, por sua vez, apresentaram tempos mais curtos do que as “não-palavras”. Este resultado sugere que o acesso aos constituintes morfológicos facilitou os tempos de resposta, já que as palavras compostas foram decididas mais rapidamente. Ademais, em relação às medidas do MEG, no estudo de Fiorentino e Poeppel (2007), o componente M350 (supostamente envolvido com o acesso lexical) foi o único sensível à manipulação dos estímulos e demonstrou um pico anterior para as palavras compostas do que àquele encontrado nas palavras simples.

Em suma, o resultado dos tempos de resposta demonstrou que há decomposição morfológica, uma vez que os tempos de resposta entre palavras compostas e simples foram diferentes. Se não houvesse decomposição e as palavras fossem acessadas como um todo, o mais esperado seria não encontrar diferenças entre os dois grupos. Adicionalmente, a análise das observações do MEG também apontou para a decomposição no processamento, já que o componente em 350ms parece refletir aspectos do acesso lexical e demonstrou um pico acentuado no encontro às palavras compostas.

O estudo de Fiorentino e Fund-Reznicek (2009), por outro lado, observou não só o processamento morfológico, mas também a propriedade de transparência/opacidade semântica das palavras compostas na língua inglesa. Foram realizados dois experimentos de *masked-priming* (SOA extremamente curto) utilizando como estímulo os seguintes grupos de palavras:

- palavras compostas com relação semântica transparente entre seus constituintes e a palavra como um todo
Ex: *teacup* (chá + copo = xícara)
- palavras compostas com relação semântica opaca entre seus constituintes e a palavra como um todo
Ex: *honeymoon* (mel + lua = lua-de-mel)
- palavras simples que apresentam um pseudo-morfema
Ex: *penguin* (*pen-* em *penguin* não é um morfema, é apenas parte da palavra)

Os dois experimentos realizados apresentaram a tarefa de decisão lexical e, enquanto no primeiro experimento as palavras compostas eram seguidas pelos constituintes da esquerda (ex: *teacup* – *tea*), no segundo experimento as palavras compostas eram seguidas pelos constituintes da direita (*teacup* – *cup*).

Os resultados de ambos os experimentos demonstraram efeitos de facilitação nos pares morfológicos (primeiro e segundo grupos), mas não nos pares com pseudo-morfema. Os autores argumentam que estes resultados sugerem um processo de segmentação inicial no nível morfológico, pois os efeitos de facilitação não estão relacionados nem a semântica, nem a ortografia. Além disso, a partir deste estudo, foi possível argumentar que o processo decomposicional puramente morfológico parece ser geral e automático e, por isso, ocorre também no encontro aos compostos.

As palavras compostas também têm sido observadas nas línguas asiáticas como chinês (Zhou *et al.*, 1999; Huang, Chen e Lin, 2009; Zhang e Heuven, 2009 e outros) e japonês (Tamaoka e Hatsuzuka, 1998; Joyce, 1999, 2002; Joyce e Ohta, 1999 e outros). No presente trabalho, destacamos alguns estudos que observaram os compostos japoneses na interface sintaxe-semântica.

3.6. Estudos anteriores: o processamento morfológico do japonês

Como exposto no capítulo dois, o extensivo processo de empréstimo lexical do chinês e a grande produtividade de formação de palavras proporcionaram ao japonês uma rica variedade na estrutura morfológica das palavras compostas. Todavia, inicialmente, muitos estudos que buscaram observar o reconhecimento do *kanji* em compostos sugeriram que este reconhecimento ocorreria no nível da palavra e não do constituinte.

Os estudos de Morton e Sasanuma (1984); Sakuma, Itoh e Sasanuma (1989) e Morton *et al.* (1992) realizaram experimentos com *priming* e tentaram analisar como se relacionavam a forma do *kanji* sozinho com a forma deste dentro do composto. Também foi observado se o composto com *kanji* era reconhecido como uma unidade integrada ou se a partir do reconhecimento individual de seus constituintes. Os experimentos nestes estudos não demonstraram facilitação entre os pares de compostos e seus primeiros constituintes, direcionando os autores à conclusão de que o acesso lexical das palavras japonesas escritas em *kanji* utilizaria unidades do tamanho de palavras. Ou seja, foi proposto que os compostos japoneses seriam armazenados e acessados como uma unidade inteira. Assim, mesmo as palavras que possuíssem um constituinte semelhante deveriam ser armazenadas em entradas diferentes no léxico mental. Embora muitas palavras compostas com dois *kanji* (e.g. 東方 – *tōhō*, leste + lado = lado leste) apresentem um *kanji* que pode ser por si só uma palavra (東 - *higashi*, leste), ainda assim as palavras expressas tanto pelo *kanji* sozinho, como pelo composto não compartilhariam nenhuma etapa no acesso lexical.

No entanto, posteriormente, outros estudos foram realizados procurando demonstrar em contrapartida efeitos de decomposição morfológica. O estudo de Tamaoka e Hatsuzuka (1998), por exemplo, procurou observar como as variações semânticas no nível de morfema interagem com a semântica no nível da palavra através de dois experimentos: um com decisão lexical e outro com nomeação lexical.

As palavras foram separadas em três conjuntos:

- Palavras que apresentam conceitos opostos
長短 (*chōtan*)
comprido - curto (comprimento)

- Palavras que apresentam conceitos similares
柔軟 (*jūnan*)
flexível - leve (leveza)
- Palavras com conceitos fortemente relacionados
移住 (*ijū*)
mudar - moradia (migração)

De acordo com os resultados, as palavras constituídas por ideogramas com idéias similares parecem ter sido processadas mais rapidamente do que aquelas com conceitos opostos. Além disso, as palavras com conceitos fortemente relacionados foram respondidas mais rapidamente do que as palavras dos outros dois grupos. Deste modo, parece que as estruturas semânticas das palavras constituintes participam nas tarefas de decisão e nomeação dos compostos. Os autores sugerem que estes achados também indicam que a unidade do *kanji* é ativada como uma unidade morfológica, assim como a unidade palavra no nível semântico.

Ademais, os grupos de palavras com idéias similares e opostas apresentaram efeitos inibitórios que, provavelmente, ocorreram porque as representações semânticas dos dois morfemas competem entre si no nível semântico. Essa interferência é particularmente maior quando as representações morfológicas de significados opostos são ativadas. O grau de competição parece diminuir quando há similaridade semântica entre os ideogramas. Assim, os autores sugeriram através desses resultados que as palavras não são representadas apenas como um todo, pois ocorre também a representação no nível morfológico.

Procurando também demonstrar a decomposição no interior das palavras japonesas, o estudo de Joyce (2002) realizou dois experimentos de *priming* com diferentes tempos de permanência da palavra *prime*, ou seja, diferentes SOA (*stimulus onset asynchrony*): um com 3000ms e outro com 250ms. O estudo teve como objetivo observar se as palavras demonstrariam efeitos de facilitação entre palavras compostas e seus morfemas constituintes, inicialmente apresentando o morfema da esquerda como *prime* (山 *yama* - 山桜 *yamazakura*) e, em uma segunda realização dos experimentos, com o constituinte da direita como *prime* (桜 *sakura* - 山桜 *yamazakura*). As palavras foram divididas em cinco

grupos, de acordo com sua estrutura interna: os três primeiros grupos divididos por categoria sintática e os dois seguintes de acordo com a natureza semântica.

- Modificador – modificado
山桜 (*yamazakura*)
montanha – cerejeira (flores de cerejeira nascida nas montanhas)
- Verbo – complemento
登山 (*tozan*)
subir – montanha (subir a montanha)
- Complemento – verbo
外食 (*gaishoku*)
fora – comer (comer fora)
- Pares associativos
男女 (*danjo*)
Homem – mulher (homem e mulher, ambos os sexos)
- Pares sinônimos
山岳 (*sangaku*)
Montanha – montanha (montanhas)

No primeiro experimento, com SOA estipulado em 3000ms, não foram observadas diferenças entre os grupos descritos acima. Contudo, os resultados demonstraram tempos mais curtos para os pares acima descritos do que para o grupo controle (condição em branco). Em relação à diferença entre o constituinte da esquerda e o da direita, o autor só observou diferença significativa na condição “modificador – modificado”. O segundo experimento, com SOA estipulado para apenas 250ms, demonstrou resultados bastante parecidos, até no que diz respeito à análise entre os constituintes: diferenças significativas apenas nos pares com estrutura “modificador – modificado”. Os pares com *prime* apresentaram tempos mais curtos do que o grupo controle e não houve diferenças significativas entre os pares de palavras controladas.

Em resumo, os dois experimentos demonstraram efeitos de facilitação nos pares com constituintes como *prime*, indicando que os fatores morfológicos são importantes no

reconhecimento de uma palavra composta em japonês. Ademais, também não foram encontradas muitas diferenças significativas entre os constituintes da esquerda e da direita, concluindo-se, assim, que tanto o primeiro quanto o segundo constituinte facilitam o acesso da palavra em níveis iguais na maioria dos casos. Consequentemente, o autor sugeriu que a representação da informação morfológica é uma questão fundamental para os modelos sobre o processamento das palavras.

Um modelo de gramática que propõe o morfema como menor unidade do processamento linguístico é o modelo da Morfologia Distribuída. No capítulo seguinte, procuraremos observar mais de perto este modelo teórico que utilizamos como base para o nosso estudo.

4. A Morfologia Distribuída

4.1. A Morfologia Distribuída

O capítulo anterior expôs uma pequena amostra da variedade de pesquisas sobre a complexidade morfológica que apresenta resultados defendendo as diversas teorias existentes. Ultimamente, no entanto, estas questões tornaram-se ainda mais cruciais tanto para a Psicolinguística como para as perspectivas cognitivas mais amplas, pois as diferentes propostas sugerem argumentos bastante díspares sobre a natureza da representação linguística na arquitetura cognitiva da linguagem (Fiorentino e Poeppel, 2007).

As propostas teóricas iniciais assumiam que as palavras seriam armazenadas inteiramente fazendo parte de uma lista de vocábulos à parte da computação das sentenças. Ou seja, as palavras seriam signos *saussureanos*, correspondentes de som e significado, que precisariam ser memorizados; enquanto que as sentenças seriam objetos de processos internos da língua. Contudo, como vimos, as propostas desta natureza foram bastante questionadas, pois, apesar da economia de recursos computacionais, há uma notável demanda de armazenagem mnemônica. Boa parte das palavras derivadas apresenta visivelmente uma forte relação com suas raízes e, por isso, não parece ser necessário o armazenamento de cada item separadamente; é o caso das palavras do português “globo”, “globalizar” e “globalização”.

O visível custo destinado à memorização fez com que algumas teorias expandissem um pouco a concepção do Léxico ao aumentar a carga computacional e diminuir a carga de memória. A dicotomia, no entanto, permaneceu: *palavras* seriam processadas no *Léxico*, enquanto que as *sentenças* seriam computadas na *Sintaxe*. Conexões entre estrutura/significado e alguns aspectos fonológicos seriam derivadas no Léxico, enquanto que outros aspectos fonológicos e outras relações entre estrutura/significado seriam derivados (posteriormente) na Sintaxe (Marantz, 1997).

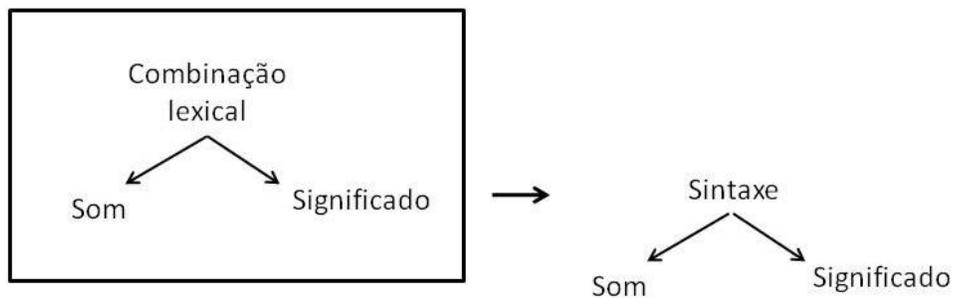


Figura 8: A dicotomia entre o Léxico e a Sintaxe.

Além de bastante redundante (fig.8), esta análise apresentou algumas dificuldades, por exemplo, como definir o conceito do que é realmente uma palavra (Marantz,1997)? Se pensarmos em termos fonológicos, dificilmente encontraremos limites fonológicos coincidindo com limites sintáticos (ex: /vocestudopraprova?/, Você estudou para a prova?). Adicionalmente, a idéia de que as palavras seriam as estruturas portadoras de significados idiossincráticos não se sustenta ao observarmos que o domínio entre correspondências especiais entre estrutura e significado não corresponde somente às palavras. Expressões idiomáticas muitas vezes são formadas por mais de um constituinte e recebem um significado idiossincrático particular (ex. “bateu as botas”). Vale ressaltar também que as palavras compostas unem mais de um item lexical e também apresentam larga idiossincrasia.

Procurando uma solução para estes problemas causados pela recorrente dicotomia entre Léxico e Sintaxe, a teoria da **morfologia distribuída** foi proposta. Distanciando-se das teorias anteriores, a Morfologia Distribuída (*Distributed Morphology*, doravante, **MD**) propõe a “explosão” do Léxico e a substituição deste por três listas não-computacionais distribuídas (Marantz, 1997). Estas listas são acionadas no curso da derivação e cada uma fornece um tipo de informação linguística: a lista 1, também chamada de “Léxico Estrito”, apresenta a informação gramatical formal e um conteúdo semântico mínimo (que será lido na Forma Lógica, ou *Logical Form*, adiante LF); a lista 2 fornece a informação fonológica; e a lista 3, a informação semântica que implica as informações extra-linguísticas.

A seguir, apresentamos mais detalhadamente algumas propriedades da Morfologia Distribuída que torna o modelo tão diferente dos demais.

4.2. As propriedades da Morfologia Distribuída

Como dito anteriormente, na MD não há mais a dicotomia entre Léxico e Sintaxe, *todas as composições são sintáticas*. Ou seja, a estrutura interna das palavras é criada a partir do mesmo mecanismo de construção da estrutura interna das sentenças. Esta é uma das propriedades que diferencia quase absolutamente a MD das teorias anteriores Lexicalistas e recebeu a nomeação “Estrutura Sintática Hierarquizada *All the way down*”. Outras duas propriedades fundamentais que diferenciam a MD das antigas teorias são a “Inserção Tardia” e a propriedade de “Subespecificação” (Halle e Marantz, 1994; Medeiros, 2008).

A propriedade nomeada “Inserção Tardia” refere-se à noção introduzida pela MD que propõe mecanismos diferentes para a produção de expressões “sintática-semanticamente” complexas e para as expressões fonológicas. Isto quer dizer que, enquanto nas teorias Lexicalistas os itens entram na computação com estrutura interna fechada (às operações sintáticas) e já com conteúdo fonológico, na MD as categorias sintáticas iniciais são puramente abstratas, não possuem traços fonológicos. Estes últimos, chamados por Itens do Vocabulário, serão apenas inseridos nos nós terminais após terminadas todas as operações sintáticas (e morfológicas).

Uma terceira propriedade bastante peculiar da MD é a “Subespecificação” que prediz que as expressões fonológicas não necessitam ser plenamente especificadas para serem inseridas nos nós terminais da derivação sintática. Ou seja:

- 1) os Itens do Vocabulário não são somente sequências de sons sem nenhuma especificação. Ao contrário, trazem informações sobre traços sintáticos, morfológicos e semânticos;
- 2) um nó sintático/morfológico (resultante de operações anteriores) pode apresentar mais informação do que aquela que contém o Item do Vocabulário que será inserido – ou seja, os Itens do Vocabulário podem ser *subespecificados*, não precisam conter todos os traços que o nó apresenta.

Por fim, dizer que a gramática tem uma Estrutura Sintática Hierarquizada *All the Way Down* significa propor que:

- 1) os nós terminais, nos quais os Itens do Vocabulário serão inseridos, se organizam em estruturas hierárquicas determinadas por princípios e operações da sintaxe;
- 2) esta estrutura determinada pela sintaxe está também dentro das palavras e, por isto, “*all the way down*” (ou seja, “todo o caminho abaixo” até às menores estruturas da língua).

Estas propriedades fazem parte do modelo de Gramática proposta pela MD que apresenta a participação de três listas distribuídas, fornecedoras das informações necessárias para o processamento das palavras e frases. Vejamos essas listas na seção seguinte.

4.3. As Listas Distribuídas

A participação de cada uma das listas ocorre separadamente em diferentes momentos do processamento da palavra:

- 1) A Lista 1, também chamada de Léxico “Estrito”, fornece as unidades atômicas com as quais a sintaxe opera. Estas unidades são os feixes de traços gramaticais ou morfossintáticos (como, por exemplo, traços de gênero, número, pessoa, etc.). Estes traços são determinados pela Gramática Universal (cada língua utiliza apenas um subconjunto dos traços disponibilizados) e, talvez, por alguns princípios próprios às línguas. Já as maneiras como os traços se agrupam em feixes parecem ser particulares às línguas. Esta lista também fornece previsões de pontos para a inserção das raízes (Medeiros, 2008).
- 2) A Lista 2 é responsável por fornecer os Itens do Vocabulário para os nós terminais da sintaxe/morfologia. Os Itens do Vocabulário, por sua vez, são as unidades mínimas de conexões de traços de som com traços sintático/semântico. Desta forma, os Itens do Vocabulário se assemelham com a noção de morfema

das outras teorias, pois são as peças mínimas de armazenagem com informação fonológica (Marantz, 1996).

- 3) A Lista 3 é chamada de Enciclopédia e relaciona informação linguística com extra-linguística. Nesta lista estão os significados idiossincráticos das palavras e expressões. Na verdade, de acordo com a MD, até palavras simples como “gato” recebem seu significado idiossincrático na Enciclopédia (Marantz, 1996). Afinal, o conhecimento de que “gato” refere-se ao significado “animal felino doméstico” é uma idiossincrasia da língua tal como “gatos pingados” em “foram apenas alguns *gatos pingados*”, refere-se a “poucos”. Ambos os conhecimentos (semânticos) não estão na estrutura da palavra/expressão. Consequentemente, precisam ser, de alguma forma, aprendidos e listados. Assim, o papel da Enciclopédia está em listar as conexões entre os nós da derivação sintática e os significados irrelevantes para o processamento morfo-fonológico.

Alguns esquemas visuais têm sido propostos para demonstrar a estrutura da Gramática prevista pela MD. Optamos por expor o esquema apresentado em Medeiros (2008) por nos parecer bastante condizente com a teoria, além de já apresentar os termos e os exemplos em português (fig.7).

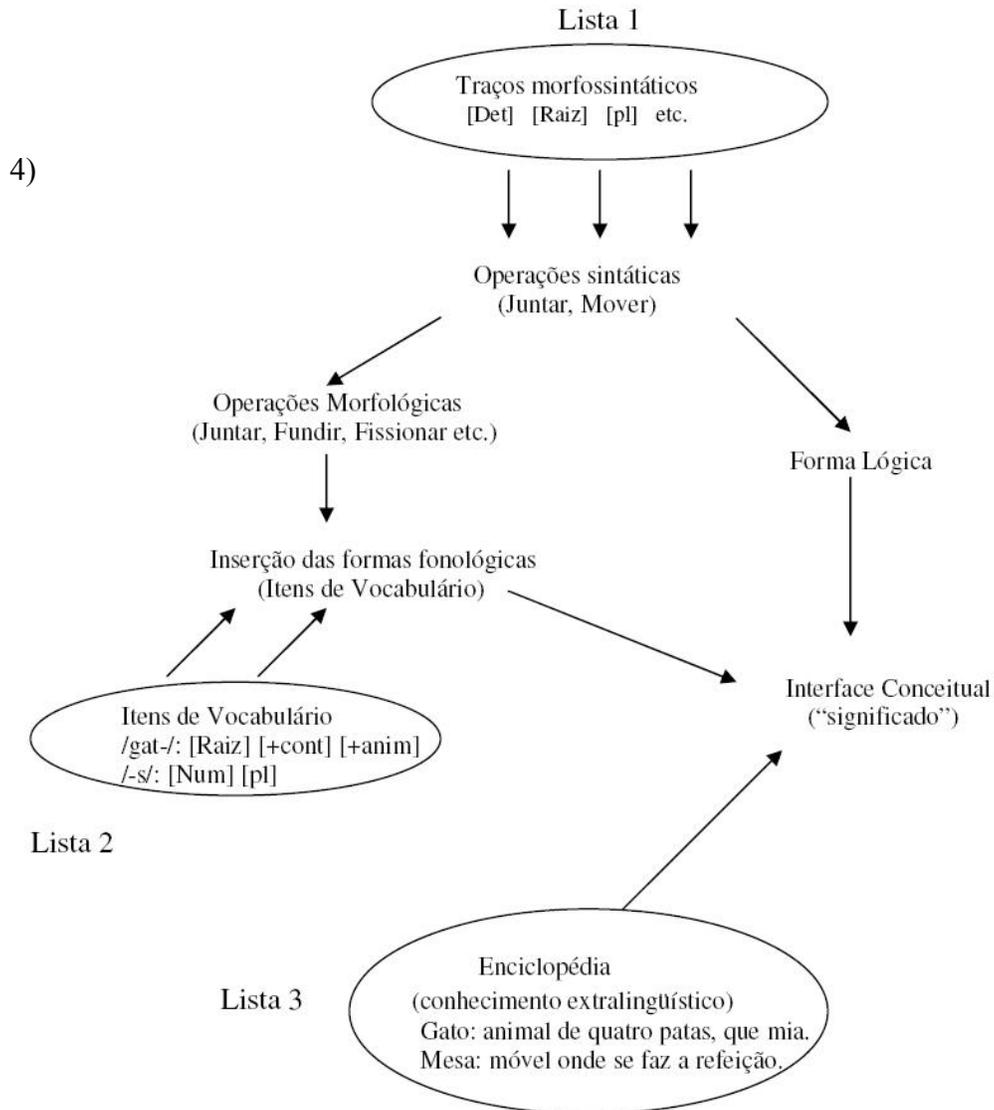


Figura 9: As listas da Morfologia Distribuída.

De acordo com o esquema anterior, podemos observar que a MD assume que o sistema computacional trabalha somente com os feixes atômicos iniciais fornecidos pela Lista 1. Isto significa dizer que: 1) os nós terminais da sintaxe não contêm traços fonológicos nem traços semânticos irrelevantes (os Itens do Vocabulário sofrem inserção tardia e os traços semânticos iniciais são os mínimos possíveis); 2) A sintaxe começa com constituintes simples não-bifurcados (Marantz, 1996).

Portanto, a computação inicial das palavras inicia-se com uma numeração, ou melhor, com uma pré-seleção dos feixes e marcação de posição para as raízes que serão usadas na derivação sintática. O inventário básico dos terminais sintáticos é dividido entre morfemas abstratos e raízes (Embick e Noyer, 2007):

- Morfemas-f (f de funcionais): também chamados de morfemas abstratos, são compostos exclusivamente de traços não-fonéticos, como traço de Passado [Pas], plural [pl];
- Morfemas-l (l de lexical) ou Raízes: as raízes de acordo com Marantz (a aparecer) são como signos *saussureanos*, uma vez que são identificadas por sua forma fonológica e por seu(s) significado(s), os quais estão inexoravelmente ligados. Diferentemente dos signos *saussureanos*, as raízes podem ter significados múltiplos, contextualmente determinados. Entretanto, elas não podem ter múltiplas formas fonológicas subjacentes (alomorfes supletivos). Essa diferença é consequência do fato de que os traços fonológicos são parte do sistema da língua, enquanto que os significados das raízes caem nos sistemas conceituais, extra-linguísticos, da mente. No entanto, para Harley (Medeiros, 2008), as raízes formam uma lista que é acessada juntamente com a lista de traços abstratos – logo, ambas contribuem para a formação de nós terminais. Adversa à idéia de que as raízes são como signo *saussureanos*, Harley acredita que as raízes correspondem aos *conceitos* de Fodor (1998), ou seja, estão relacionados aos *universais* semânticos e, portanto, não possuem matrizes fonológicas e nem são particulares de cada língua.

Deste modo, podemos ver que a natureza da raiz não está ainda de toda esclarecida. Porém, parece haver consenso na seguinte afirmação: raízes não podem aparecer sozinhas, elas precisam sempre ser categorizadas por um núcleo funcional (fig.8). Estes núcleos são os morfemas categorizadores: “n” (para nome), “v” (para verbo) ou “a” (para adjetivo), que

determinam as margens das fases (domínios cíclicos) e, por isto, são também chamados de *licenciadores*. Isto quer dizer que a concatenação com um morfema categorizador deflagra o *spell-out* da estrutura computada até este momento do histórico da derivação (Marantz, 2001; Medeiros, 2008).

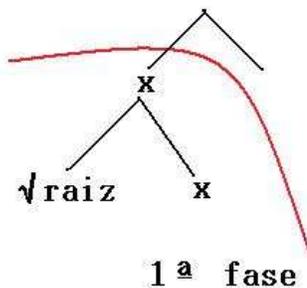


Figura 10: A primeira fase da derivação.

Ao concatenar a raiz com o morfema categorizador, a informação contida nos nós resultantes das operações sintáticas é mandada para LF e para o componente Morfológico do sistema computacional (*Morphological Structure*, MS). Neste componente, outro conjunto de operações se aplica sobre os nós sintáticos, criando novos nós, movendo alguns, copiando ou apagando traços, etc. Após completas as operações sintático/morfológicas, os Itens do Vocabulário competem por inserção nos nós resultantes na Forma Fonológica (mais conhecido por *Phonological Form*, adiante PF).

Em seguida, qualquer escolha de traço, de concatenação, de movimento não-forçado, de escolha do Item do Vocabulário, ou seja, qualquer escolha “não-forçada” no curso da derivação serve de *input* para a interpretação semântica na Enciclopédia. Isto é, distante das teorias anteriores, na Morfologia Distribuída, LF não é sozinho o *input* para a interpretação semântica. Ao contrário, todas as escolhas feitas no curso da derivação até a chegada na terceira lista servirão para determinar o significado idiossincrático correto da estrutura processada.

A derivação como um todo é interpretada, mas os aspectos específicos da derivação que são interpretados na Enciclopédia são precisamente qualquer e todas as escolhas “não-forçadas” feitas. Ainda de acordo com a MD, o único sistema capaz de fazer a conexão

entre as informações do Léxico Estrito, do Vocabulário e da Enciclopédia é a Gramática como um todo (Marantz, 1996).

Embick e Noyer (2007) propõem um esquema simplificado da Gramática que exporemos aqui a título de resumo sobre as listas distribuídas e os estágios da derivação (fig. 9):

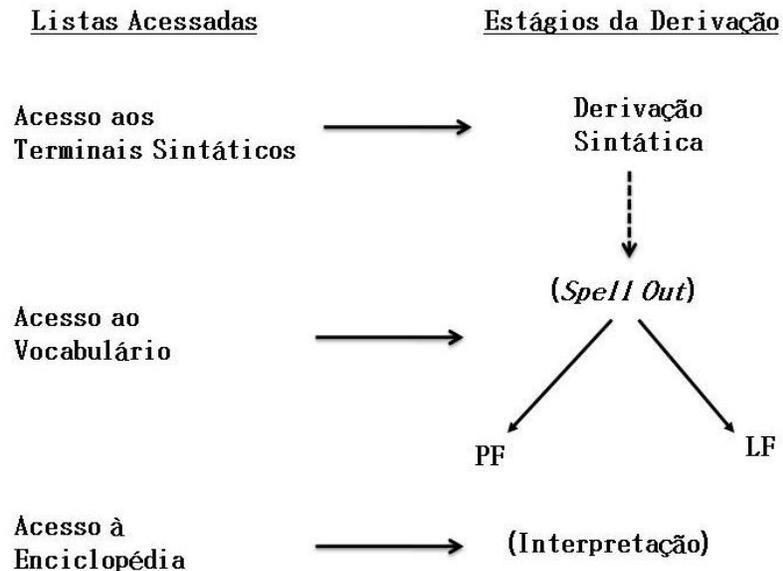


Figura 11: As listas distribuídas e os estágios de derivação.

4.4. A idiosincrasia semântica

A idéia de que somente a palavra é local de variados tipos de idiosincrasia está na base do pensamento linguístico dominante, já que é usual observar entre as línguas palavras complexas apresentando opacidade semântica. Contudo, a MD demonstra que não só as palavras complexas apresentam idiosincrasia, mas também as expressões idiomáticas e, até mesmo, as palavras simples. Pois as palavras, assim como as sentenças, são as interpretações dos resultados gerados pelo sistema computacional da Gramática. Mesmo uma palavra simples, ou seja, uma estrutura morfológica simples é uma estrutura sintática. Como foi observado anteriormente, as usuais categorias sintáticas de nome, verbo e

adjetivo, na verdade, são categorias morfológicas que emergem durante a derivação no contexto de certas posições funcionais.

Esta característica da Gramática pode ser mais facilmente visualizada em línguas flexionais como o português, pois não é possível pronunciar a raiz de uma palavra sozinha (*gat-). A palavra do português “gato”, por exemplo, é fruto da concatenação da raiz (√) com o morfema categorizador de nome (n). Línguas como o inglês, todavia, apresentam uma falsa impressão de que as palavras simples são as raízes (“cat”).

Observando o processo de uma palavra simples como “gato”, já sabemos que a Lista 1 fornece alguns traços estruturais como √ ou “n”, e alguns traços semânticos mínimos como “animado”, “singular” ou “objeto”. Estes traços são ditos mínimos no sentido de que não seriam capazes de distinguir a palavra “gato” da palavra “cachorro”. Os outros conhecimentos linguísticos sobre “gato” não são importantes para o sistema computacional. Por exemplo, o traço que distingue as palavras “cachorro” e “gato” não é de natureza estrutural. Tal diferença faz parte do conteúdo extra-linguístico e, por isso, é matéria do conhecimento da Enciclopédia. Este conhecimento é usado na interpretação semântica em LF, mas não em computações gramaticais no sistema computacional e nem envolvendo LF (Marantz, 1996).

Por outro lado, no componente fonológico, todos os Itens do Vocabulário competem por inserção em cada nó terminal. Ganha o Item do Vocabulário que possui mais traços compatíveis com o nó em questão. Desta forma, o uso da Enciclopédia, precisa envolver o conhecimento de quem venceu esta competição entre os Itens do Vocabulário, se /gato/ ou /cachorro/, para fornecer corretamente o significado relativo. No caso de “gato”, o significado que poderá ser encontrado na Enciclopédia será algo como “animal felino doméstico”.

Resumidamente, podemos dizer que as entradas enciclopédicas conectam as derivações das conexões ocorridas em LF e PF aos significados não-composicionais, ou seja, idiomáticos. Concluímos, então, que “gato” é um sintagma (*phrase*) por si só, pois trata da realização da raiz em um ambiente nominal sintático. Já o significado de “gato” é idiomático/não-composicional/enciclopédico, ou seja, o correspondente da escolha de um Item Vocabular particular em um ambiente sintático particular. Deste modo, podemos

concluir também que não são as palavras que recebem significados idiomáticos/especiais e sim as raízes, já que estas não apresentam seus significados completamente determinados pelos traços gramaticais (Marantz,1996).

A MD assume que existem, portanto, significados composicionais e significados não-composicionais (idiosincráticos). Significados composicionais são aqueles que são previsíveis a partir da descrição estrutural morfossintática. Já os significados não-composicionais são aqueles que são idiosincráticos. Estes últimos são listados e armazenados na Enciclopédia.

Uma vez que são as raízes que recebem os significados idiosincráticos a partir do ambiente sintático em que se encontram, Marantz (2001) demonstra que os significados não-composicionais são definidos dentro da primeira fase cíclica da formação da palavra; enquanto os significados composicionais são construídos após a primeira categorização da raiz. Ou seja, ao se criar uma palavra no domínio da raiz (*inner*) define-se o significado não-composicional idiosincrático da estrutura; após a segunda fase (*outter*) o significado passa a ser composicional (fig.10)⁶.

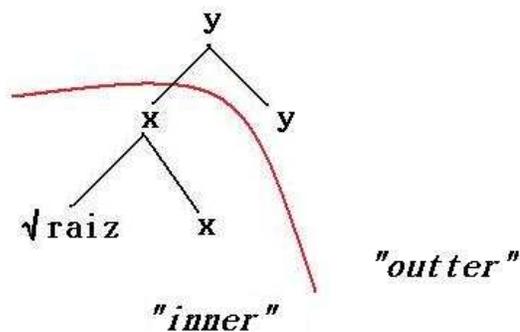


Figura 12: O domínio dentro (*inner*) e fora (*outter*) da raiz.

Na derivação exposta anteriormente (fig.10), o morfema categorizador “x” determina a margem do domínio cíclico. Portanto, a combinação de raiz e “x” é mandada para LF e PF para as interpretações semânticas e fonológicas. Em seguida, o significado da raiz no contexto de “x” é negociado, utilizando o conhecimento Enciclopédico. Quando um segundo núcleo (y) anexa-se à “x”, este núcleo terá como complemento uma estrutura que

⁶ Esta figura foi elaborada de acordo com Marantz (a aparecer).

já teve o significado e a pronúncia da raiz negociados, pois “x” já foi enviado para LF e PF em sua primeira fase.

Logo, estruturalmente, quando “y” se anexa fora do domínio de “x”, o núcleo “vê” os traços de “x” localmente, e não os traços, as propriedades ou a identidade da raiz concatenada à “x”. Isto ocorre porque as propriedades das raízes são “visíveis” apenas para o primeiro categorizador, pois é o domínio definido por “x” que irá associar a pronúncia da raiz ao seu significado idiossincrático. Por outro lado, todos os núcleos adicionados acima da primeira concatenação contribuirão composicionalmente para o significado da expressão resultante (Marantz, 2001; Medeiros, 2008).

A raiz $\sqrt{\text{filtr-}}$ do português⁷, por exemplo, pode se combinar com um categorizador nominal diretamente, resultando no nome “filtro” (fig.11), ou pode ser concatenada com o categorizador verbal diretamente, resultando no verbo “filtrar” (fig.12). A raiz apresentará seus significados na Enciclopédia de acordo com cada ambiente sintático em que se encontra⁸.

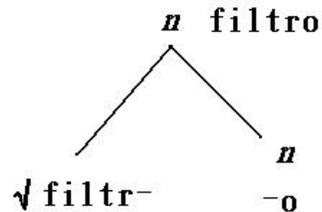


Figura 13: Raíz concatenada ao categorizador nominal.

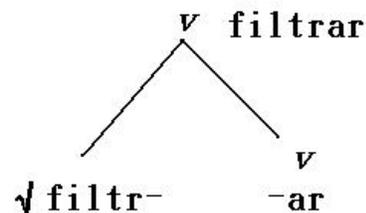


Figura 14: Raíz concatenada diretamente ao categorizador verbal.

⁷ Exemplo retirado de Medeiros (2008).

⁸ Como já sabemos, a inserção dos Itens do Vocabulário é tardia, ou seja, ainda não há forma fonológica neste momento. Os gráficos apresentam a forma fonológica apenas para facilitar a compreensão.

Por outro lado, no caso da palavra “filtragem” (fig.13) ocorre uma concatenação inicial da raiz com o categorizador verbal que define um domínio cíclico e negocia na Enciclopédia o significado da raiz neste ambiente sintático. A nominalização posterior vai contribuir composicionalmente, ou seja, dizendo que “filtragem” é o nome da eventualidade de “filtrar”.

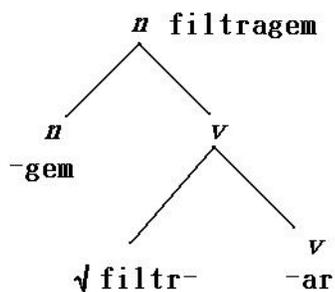


Figura 15: Raiz concatenada ao categorizador verbal e, posteriormente, ao categorizador nominal.

Esta proposta da MD empobrece a tradicional visão que argumenta que a morfologia derivacional e a morfologia flexional são dois processos diferentes. De acordo com teorias anteriores, a primeira estaria relacionada a aspectos lexicais idiossincráticos, enquanto que a segunda a aspectos paradigmáticos mais produtivos. A proposta da MD, no entanto, argumenta que os dois processos (flexão e derivação) são sintáticos, porém há uma distinção no local em que ocorrem. Algumas palavras são formadas no domínio da raiz, outras são formadas a partir de outras palavras. Derivações que ocorrem a partir da concatenação da raiz com morfema categorizador podem apresentar idiossincrasia semântica e fonológica. Em contraste, palavras derivadas de palavras já existentes são composicionais, ou seja, o significado da nova palavra é baseado no significado da base mais as características semânticas associadas ao(s) afixo(s) (Volpe, 2005).

A Morfologia Distribuída também prevê as estruturas que formam significados não-composicionais (idiossincráticos) fora do domínio da raiz, como por exemplo, as expressões idiomáticas. A formação de palavras compostas também faz parte de um processo que ocorre a partir da junção de itens após o *spell-out* e apresenta muitas vezes significados idiomáticos. Estruturas deste tipo, entretanto, são diferentes, pois apresentam dois significados: um composicional, previsível a partir de suas estruturas, e um idiomático,

chamado de significado não-composicional especial (“*special noncompositional meaning*” Marantz, 1997).

4.5. O significado não-composicional especial

De acordo com a Morfologia Distribuída, o local onde estão listados os significados não-composicionais é a Enciclopédia. Conseqüentemente, assim como há uma entrada lexical na Enciclopédia que diz que “gato” é um “animal felino doméstico”, há também uma entrada lexical para o composto “pão-duro” definindo que seu significado idiossincrático é “pessoa avarenta”. Naturalmente, há também entradas lexicais para expressões idiomáticas, como, por exemplo, para “chutar o balde” com o significado não-composicional de “desistir”.

Ainda conforme a MD, todos esses significados envolvem conhecimento Enciclopédico sobre representações linguísticas independentemente licenciadas e, por isto, devem ser tratados em igualdade. No entanto, expressões idiomáticas, assim como algumas palavras compostas, mais do que apresentar um significado não-composicional, precisam negar um efeito semântico canônico (Marantz, 1996). Como no caso de “chutar o balde”, que pode apresentar o significado composicional (“dar um pontapé no objeto balde”) ou o significado idiossincrático (“desistir”). Por esta razão, este último tipo de significado é chamado também de *significado não-composicional especial*.

Estudos sobre frases idiomáticas (Titone e Connine, 1999; Fogliata *et al.*, 2007) vêm demonstrando que no acesso a estas estruturas, ocorre a inibição do significado literal e o acesso do significado figurado a partir da memória semântica. Conseqüentemente, parece bastante claro que as expressões idiomáticas não são sequências de palavras impossíveis de serem decompostas e nem peças que não são processadas literalmente. Ademais, o fato dos constituintes de uma frase idiomática influenciarem sua interpretação não descarta a possibilidade de existir um significado “empacotado” (*package meaning*) associado a uma configuração de palavras bastante particular (Titone e Connine, 1999).

Marantz (1996) argumenta que em casos como esse, não se cria uma nova composição entre os elementos sintáticos apresentados. Uma expressão idiomática não

apresenta mudanças entre sujeito e paciente ou objeto e agente. O que talvez ocorra é uma supressão da referencialidade e a produção do significado não-composicional especial que é determinado pelo contexto sintático dentro do domínio local.

“the one syntactic representation is all you get, and this will be interpreted semantically, whatever that comes to, but semantic interpretation won’t be simply a translation into another syntactic-like representation. (...) Our claim is that the Encyclopedia does not map syntactic representations onto syntactic representations but only provides a key to the noncompositional meanings of syntactic atoms in syntactic environments. (...) The Encyclopedia doesn’t relate linguistic representations to other linguistic representations but precisely provides the kind of information that isn’t encoded in linguistic representations at all.” (Marantz, 1996, p.9)

Assim, apesar das estruturas além do domínio da raiz apresentarem significados composicionais, expressões idiomáticas suprimem a correspondência entre estrutura sintática e conceitual, providenciando em seu lugar uma estrutura conceitual não-composicional especial (Marantz, 1996).

O domínio local para estes significados especiais são definidos sintaticamente, muitas vezes tomando estruturas maiores do que palavras, outras vezes, menores. Algumas observações na literatura sobre o domínio da idiomatização seguem a hipótese de que o nó que projeta sintaticamente o “agente” serve como limite para os significados não-composicionais especiais, ou seja, a fronteira que limita o domínio de qualquer informação contextual em uma entrada enciclopédica (fig. 14).

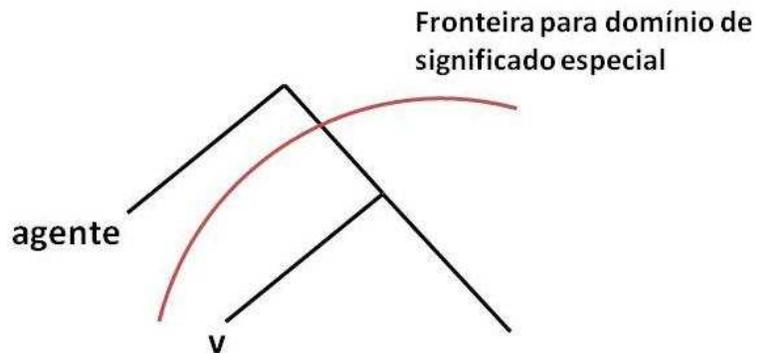


Figura 16: A fronteira para o domínio de significado especial.

Como se vê, o domínio agentivo é o limite para a reanálise do significado de uma estrutura após a passagem pelo sistema computacional e a inserção dos Itens do Vocabulário, isto é, no estágio de participação da Lista Enciclopédica. Desta maneira, os significados especiais são criados pelo ambiente sintático e não por uma simples listagem de expressões com significados “não-literais” como se previa anteriormente em outras teorias.

No presente estudo pretendemos observar as palavras compostas que também parecem apresentar tanto significados composicionais, como significados não-composicionais. Quando apresentam significados não-composicionais parecem ter comportamento parecido com as expressões idiomáticas: precisam suprir o significado composicional e providenciar o significado idiossincrático. De tal modo, procuramos em seguida detalhar melhor a estrutura dos compostos.

4.6. As palavras compostas

As palavras compostas são conhecidas por concatenarem duas ou mais palavras e, portanto, não são formadas a partir da raiz, mas a partir de outras palavras. Ou seja, as palavras compostas são formadas após o *spell-out* de duas ou mais palavras. Por conseguinte, seu significado muitas vezes é composicional e previsível a partir das estruturas formadoras. No entanto, é muito comum também observar compostos apresentando significados não-composicionais. Estes apresentam comportamento semelhante às expressões idiomáticas, pois além de serem formados após o *spell-out*, possuem um significado idiossincrático diferente do literal.

Por esta razão, é muito frequente, nos estudos sobre os compostos, os termos “transparente” e “opaco” na definição das relações semânticas entre as palavras constituintes e a palavra resultante. Compostos transparentes são aqueles que apresentam relação praticamente de um para um entre o significado das palavras constituintes e o significado do composto como um todo. Já os compostos opacos são aqueles que não apresentam a relação clara entre os significados; pelo contrário, mesmo quando há participação semântica de um dos constituintes, não se observa composição entre os

significados dos constituintes e o significado da palavra. Por exemplo, na palavra inglesa *butterfly* (borboleta), é possível que muitos digam que a semântica da palavra *fly* (voar) participa no significado da palavra. Contudo, é indiscutível o fato de que não há composição entre os significados dos dois constituintes para a construção do significado da palavra inteira; afinal, o significado da palavra borboleta não é fruto da composição das palavras manteiga (*butter*) e voar (*fly*).

Em japonês, uma palavra composta como *shinnen* (ex.1), em que *shin* significa “novo” e *nen* significa “ano”, apresenta claramente um significado transparente e composicional. No entanto, uma palavra como *seinen* (ex.2) apresenta um significado opaco e não-composicional, pois a combinação das palavras “azul” com “ano” origina a palavra “jovem”.

1) 新年 (*shinnen*)

novο – ano (novο ano)

2) 青年 (*seinen*)

azul – ano (jovem)

Alguns estudos anteriores, como o de Sandra (1990), afirmam que as palavras compostas com significado não-composicional provavelmente seriam armazenadas por inteiro e não decompostas. No entanto, como um procedimento de decomposição morfológica automática poderia distinguir compostos transparentes de compostos opacos? E, caso realmente fosse capaz, até que ponto este processo seria automático? (Libben, 1998). Seguindo as premissas da MD, acreditamos que tanto as palavras compostas com significado transparente, quanto as que apresentam significado opaco, são processadas em decomposição. Há, no entanto, uma diferença no processamento das duas: aquela que também recebe significado idiomático (e.g. *seinen*) irá apresentar uma reanálise na Enciclopédia no momento de concatenação dos dois núcleos.

Desta maneira, podemos imaginar que uma palavra como *shinnen* (fig.15) é formada da composição de duas partes: uma raiz $\sqrt{\text{shin-}}$ que se concatena a um morfema categorizador de adjetivo (*a*) e uma raiz $\sqrt{\text{-nen}}$ que se concatena a um morfema categorizador de nome (*n*). Quando as raízes se concatenam aos morfemas funcionais e são enviadas para LF e PF, têm seus significados negociados na Enciclopédia. Em seguida, quando as duas partes se concatenam, seus significados se compõem e resultam no significado total da estrutura.

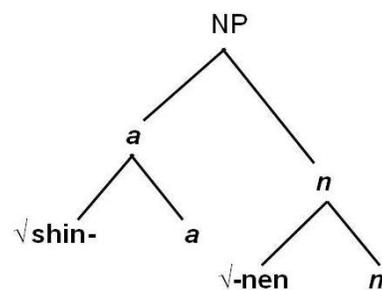


Figura 17: Palavra composta com significado composicional.

Uma palavra composta com significado idiossincrático como *seinen* também é formada de duas partes e, por isto, precisa da mesma maneira ser decomposta para ser processada. Contudo, no momento de concatenação das duas raízes (que já foram concatenadas aos seus morfemas categorizadores), o significado literal é inibido e ocorre a reanálise do significado na Enciclopédia (fig.16).

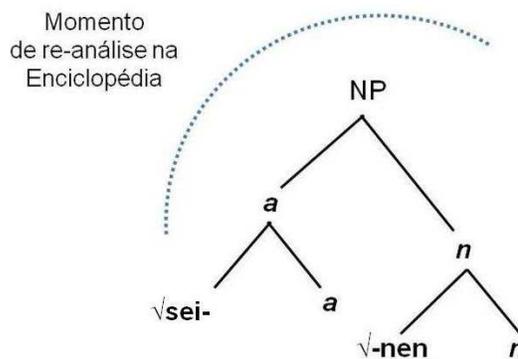


Figura 18: Palavra composta com significado não-composicional.

Assim, de acordo com a MD, prevemos que as duas estruturas, tanto a que apresenta significado transparente, como a que apresenta significado opaco, sofrem o processo de decomposição, pois o processamento morfológico inicial é automático e geral. Porém, o que parece ainda não estar claro é de que forma o processo de reanálise semântica é “engatilhado”. Uma hipótese sugere que esta reanálise pode ser iniciada a partir de fatores contextuais, uma vez que, neste momento, a estrutura já sofreu o *spell-out* e, assim, apresenta participação da Interface Conceptual.

Além disso, pistas prosódicas também parecem ser capazes de “engatilhar” este processo, pois há evidências do uso da prosódia para diferenciar palavras compostas de sintagmas posicionais simples. Por exemplo, em inglês, compostos em geral são acentuados na primeira palavra, enquanto que em sintagmas posicionais simples o acento vem na segunda palavra. Assim /bláckbóard/ apresenta necessariamente a estrutura de um sintagma posicional e significa um quadro que é preto (*black* = preto, *board* = quadro). Por outro lado, /bláckboard/ apresenta a estrutura de um composto e, muitas vezes, o objeto que representa possui a cor verde ou branca. Assim, sem a pronúncia ou a pontuação como guias, algumas cadeias de palavras podem ser lidas tanto como um sintagma como uma palavra composta, pois é a prosódia que define a estrutura (Pinker, 1994).

É preciso observar, no entanto, se esta característica é presente e regular também na língua japonesa e em outras línguas em geral. Por esta razão, a realização de estudos experimentais parece cada vez mais crucial para a obtenção de respostas mais consistentes sobre as hipóteses e teorias propostas.

Apresentaremos a seguir os experimentos realizados neste estudo.

5. Os experimentos

5.1. Apresentação dos experimentos

O objetivo do presente estudo era o de investigar o fator da transparência semântica e a decomposição morfológica das palavras compostas em japonês através da análise de dois experimentos de *priming*. Como foi explicitado anteriormente, ainda há controvérsias em relação à natureza do conteúdo morfológico: enquanto algumas teorias defendem que este conteúdo é empacotado e processado no nível das palavras, outras teorias acreditam que o conteúdo morfológico são os átomos do processamento lingüístico. As últimas defendem que a palavra não é a menor unidade lingüística, pelo contrário, esta precisa ser decomposta em pedaços menores. Ademais, também demonstramos que muitos estudos argumentam que os fatores semânticos participam do processo de decomposição, enquanto outros estudos defendem que a decomposição inicial no reconhecimento de uma palavra é puramente morfológica.

Através da observação de estudos que trabalham na linha Decomposicional (Longtin, Segui e Hallé, 2003; Rastle, Davis e New, 2004; Fiorentino e Fund-Reznicek, 2009) e seguindo as premissas da teoria da Morfologia Distribuída, acreditamos que, de fato, há um processo de decomposição morfológica no início do processamento lingüístico. Este processo é automático e mecânico, portanto, só participam os traços semânticos mínimos; isto quer dizer que a transparência semântica das palavras não é um fator relevante nesse estágio do processamento lingüístico. Para melhor analisar esta questão, optamos por contrastar, em dois experimentos de *priming*, palavras compostas semanticamente transparentes e opacas da língua japonesa. Na realização do experimento, tínhamos como finalidade os seguintes pontos: 1) observar se encontraríamos efeitos de facilitação na apresentação de palavras compostas posterior a um de seus constituintes morfológicos; 2) examinar se os tempos de resposta para as palavras transparentes e opacas diferenciarão significativamente entre si; 3) tentar analisar o curso temporal do reconhecimento de palavras, contrastando diferentes tempos de permanência da palavra *prime*.

Visto que estamos lidando com palavras compostas com dois constituintes, ou seja, duas raízes concatenadas, a análise se estendeu tanto para o primeiro constituinte, quanto para o segundo. Adicionalmente, uma vez que o tipo de condição controle utilizado em um experimento de *priming* ainda é motivo de dúvidas (Forster, Mohan e Hector, 2003), determinamos utilizar os dois tipos de controle em discussão: um constituinte não relacionado e uma condição em branco. Para alguns estudiosos o uso do constituinte ou de uma palavra não relacionada pode causar a inibição no acesso de uma palavra apresentada posteriormente. Deste modo, os tempos encontrados no grupo controle neste caso são muito altos e, logo, não formam um conjunto ideal para comparação. Por outro lado, outros estudiosos argumentam que a condição em branco também não é o controle ideal, pois o reconhecimento de uma palavra sem uma antecedente não é obviamente igual ao reconhecimento de uma palavra após outra e, desta maneira, a comparação entre os dois conjuntos não pode servir de análise.

Em relação ao curso temporal do reconhecimento visual das palavras, optamos por utilizar um SOA mais longo de 150ms e um SOA mais curto de 60ms. Nossa hipótese sugeria que encontraríamos efeitos de facilitação nos pares morfologicamente relacionados, já que a estrutura morfológica das palavras é acessada inicialmente. Contudo, no experimento com SOA mais longo, a semântica dos constituintes apresentados como *prime* seria reconhecida e acarretaria diferenças nos resultados entre as palavras transparentes e opacas. Já no segundo experimento, por apresentar os constituintes *prime* por um tempo bastante curto, acreditávamos que não seria possível o acesso aos traços semânticos e, conseqüentemente, as médias das palavras transparentes e opacas seriam similares.

5.2. A metodologia

5.2.1. Material

Inicialmente foram organizados dois questionários para definir as palavras semanticamente transparentes e opacas de acordo com a compreensão de um falante nativo de japonês. Cada questionário apresentava 100 palavras das quais, de acordo com a nossa intuição, 50 palavras eram transparentes e 50 eram opacas. Os dois questionários

apresentaram um total de 200 palavras, 100 palavras supostamente transparentes e 100 palavras supostamente opacas. As palavras foram listadas aleatoriamente em folhas de papel impressa com fonte MS 明朝, tamanho 12.

A familiaridade de todas as palavras foi observada a partir do volume dois do banco de dados da NTT⁹ (Amano e Kondo, 1999): as palavras opacas apresentavam média de familiaridade 6.01 e as palavras transparentes, média 5.93. Adicionalmente, todas as palavras apresentavam ideogramas que fazem parte dos sete primeiros anos de ensino no Japão, ou seja, todos os ideogramas faziam parte do grupo de *kanji* ensinados no Ensino Fundamental. Nos questionários era requerido que os participantes definissem o nível de composição entre as palavras compostas inteiras e seus constituintes em uma escala de 1 a 7 (1 para menor composição e 7 para maior composição). Para a correta compreensão da tarefa do questionário, procuramos exemplificar o nível de composição através do seguinte exemplo:

A palavra 「変形」 possui dois ideogramas: o primeiro possui o significado de mudança e o segundo de forma. Através do significado dos dois constituintes é compreendido o significado da palavra composta que é “mudança da forma”. Portanto, dentro da escala de pontuação, esta palavra parece apresentar o nível mais alto de composição, ou seja, o valor 7.

Os questionários foram respondidos por 40 falantes nativos de japonês (20 para cada questionário) entre eles professores das Universidade de Tama e da Universidade de Tecnologia de Tóquio, além de funcionários do escritório da Associação de Descendentes de japonês no Exterior da JICA (*Japan International Center Abroad*) de Yokohama. A partir do questionário, foram separadas 36 palavras opacas que apresentaram a média de pontuação menor que 3, e 36 palavras transparentes que apresentaram a média de pontuação igual ou maior a 6 (veja ao final o apêndice). Estas 72 palavras compostas foram

⁹ O banco de dados da NTT (*Nippon Telegraph and Telephone Corporation*) consiste de sete volumes. Os volumes de 1 a 6 contém dados a partir de coletas com sujeitos sobre a familiaridade e a plausibilidade ortográfica de 80.000 palavras do dicionário de Japonês *Shinmeikai*, assim como os índices de familiaridade e complexidade de 6.847 caracteres da lista *Japanese Industrial Standard*. O volume 7 contém dados a partir da coleta de 14 anos do jornal *Asahi* e apresenta os índices de frequência de palavras e caracteres.

utilizadas como palavra alvo nos experimentos. Para manter a mesma proporção, foram adicionadas 72 não-palavras à lista de estímulos. As não-palavras eram formadas por dois ideogramas existentes na língua Japonesa e foram retiradas da coleta realizada por Joyce (1999). Posteriormente, escolhemos 144 ideogramas para serem utilizados como constituinte não relacionado nos pares de um dos grupos de controle. Nenhum ideograma era apresentado mais de uma vez, ou seja, não houve apresentação de *kanji* repetido. Adicionalmente, visando diminuir os efeitos ortográficos, as palavras *prime* e alvo foram apresentadas com diferentes fontes de letra. O constituinte *prime* foi apresentado com a fonte MS ゴシック e a palavra alvo foi apresentada com a fonte MS 明朝, ambos com tamanho 60.

Nos experimentos, a apresentação das palavras e não-palavras alvo era antecedida pelo estímulo *prime* que foi dividido entre quatro condições: 1) o primeiro constituinte; 2) o segundo constituinte; 3) um constituinte não relacionado; 4) uma condição em branco. Em suma, uma palavra composta como 閉店(*heiten*) poderia apresentar quatro tipos de *prime*: 1) o primeiro constituinte (閉), 2) o segundo constituinte (店), 3) um constituinte não relacionado (用), ou, 4) uma tela em branco. Este material foi dividido em uma plataforma quadrado latino, ou seja, todos os sujeitos viam todas as condições, mas nunca nenhum item mais de uma vez. Cada sujeito observava apenas uma das versões e, portanto, via nove itens de cada condição sem ver nenhum item repetido.

A figura a seguir procura exemplificar o *design* do experimento:

Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4
閉 → 閉店	店 → 閉店	喜 → 閉店	□ → 閉店
学 → 休学	受 → 休学	□ → 休学	休 → 休学
塩 → 大切	□ → 大切	大 → 大切	切 → 大切
□ → 自由	自 → 自由	由 → 自由	害 → 自由

Figura 19: Exemplificação do experimento organizado em uma plataforma quadrado-latino.

Em relação à forma fonológica das palavras, vimos anteriormente que os ideogramas japoneses apresentam muitas vezes mais de uma leitura. Contudo, este fator não foi controlado em nosso estudo, uma vez que seria muito difícil reunir uma quantidade suficiente de palavras com índice alto de familiaridade, média suficiente sobre transparência e apenas uma forma de leitura. Posteriormente, o padrão estrutural também é um aspecto interessante das palavras compostas japonesas. Como foi explicitado, há diferentes tipos de padrão como verbo – verbo, verbo – nome, nome – nome, etc. No entanto, através do nosso estudo, observamos que as palavras opacas parecem perder o padrão estrutural. Um exemplo é a palavra 人參 (*ninjin*) que compreende a soma das palavras 人 (pessoa) e 參 (ir, visitar) resultando a palavra “cenoura” que, por sua vez, não apresenta o padrão estrutural nome – verbo. Esta observação parece estar de acordo com a nossa hipótese de que as palavras compostas opacas se assemelham às frases idiomáticas ao demonstrar supressão de referencialidade. Por conseguinte, os diferentes tipos de estrutura dentro das palavras compostas não foram controlados.

5.2.2. Aparatos: O experimento foi elaborado utilizando o software SuperLab (versão 4.0), e foi realizado em um Laptop IBM (X60s, CoreDuo 1.66 GHz/1.49GB; 74.5GB HDD; 12.1”) com uma caixa de resposta conectada.



Figura 20: Caixa de resposta do SuperLab.

5.2.3. Variáveis independentes e dependentes

As variáveis independentes eram o grau de semântica das palavras (transparente ou opaca) e a condição do *prime* (primeiro constituinte, segundo constituinte, constituinte não relacionado ou a condição em branco). Por outro lado, as variáveis dependentes eram os tempos de resposta da decisão lexical, assim como a acuidade das respostas.

5.2.4. Procedimentos

Os experimentos foram realizados no Laboratório Akama no Departamento de Ciências do Sistema Humano da Universidade de Tecnologia de Tóquio. Os sujeitos eram primeiramente requeridos a assinar um termo de aceitação, onde estava escrito não somente que os dados seriam utilizados exclusivamente para estudos de origem cognitiva e linguística, como também afirmava que a privacidade dos sujeitos seria mantida em sigilo. Em seguida, os sujeitos eram direcionados a ler atentamente as instruções que foram impressas em duas folhas A4. As instruções esclareciam que o sujeito veria no centro da tela, primeiramente, uma cruz [+] e, logo em seguida, apareceria rapidamente um *kanji* ou uma tela em branco. Por fim, apareceria uma palavra com dois ideogramas. Neste momento, o sujeito deveria julgar se aquela era uma palavra em japonês ou não, apertando o botão verde com o dedo indicador da mão direita para SIM e o botão vermelho com o dedo indicador da mão esquerda para NÃO. Era também esclarecido que não apareceriam palavras de cunho científico ou palavras arcaicas, pelo contrário, as palavras seriam usuais e de amplo conhecimento. Deste modo, os sujeitos deveriam ter plena confiança em seus julgamentos que deveriam ser rápidos e precisos, uma vez que após 1,5 segundos a palavra desapareceria da tela. Adicionalmente, os sujeitos também foram esclarecidos de que no caso de algum erro, não precisavam se preocupar e sim, continuar o experimento com mais atenção. Após a leitura das instruções e sancionadas qualquer dúvida adicional, uma sessão de prática era realizada contendo 10 palavras. Em seguida, caso o experimentador observasse que as instruções foram corretamente compreendidas, o experimento era iniciado.

Os procedimentos do primeiro e do segundo experimento eram praticamente iguais, a única diferença era o tempo de permanência da palavra *prime* na tela. Enquanto no primeiro experimento a palavra *prime* era apresentada por 150ms, no segundo experimento, este tempo de permanência era de apenas 60ms. No entanto, os sujeitos não possuíam conhecimento sobre esta diferença, e o material dos experimentos, assim como as instruções direcionadas para os participantes eram as mesmas. fixação

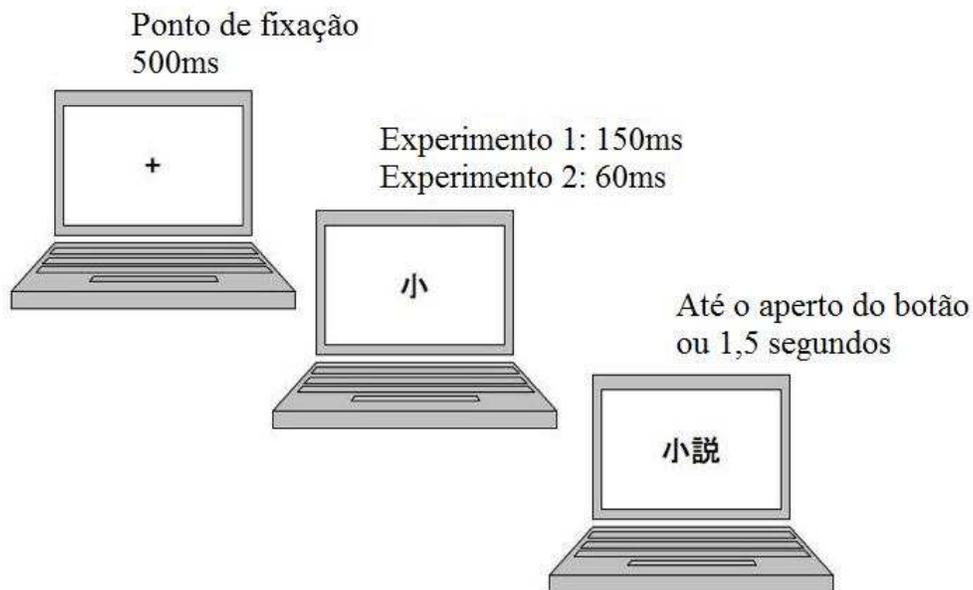


Figura 21: Exemplificação dos procedimentos dos experimentos.

5.3. Experimento 1

Participaram do primeiro experimento 51 alunos (39 homens e 12 mulheres) com idade entre 18 a 27 anos (média 20.18) da graduação e da pós-graduação da Universidade de Tecnologia de Tóquio no Japão, recebendo um pagamento em dinheiro pela participação. Os dados de três sujeitos precisaram ser retirados da análise por ultrapassarem a margem de 10% de erros e a média dos outros sujeitos. Assim, somente os dados de 48 sujeitos foram utilizados para a análise.

As respostas erradas foram excluídas da análise sobre os tempos de resposta, ou seja, apenas as respostas SIM para as palavras é que foram consideradas. Os tempos de

resposta assim como a acuidade das respostas foram analisados estatisticamente. Os tempos de respostas que ultrapassaram 2.5 da média de um participante foram ajustadas para o valor da média mais 2.5.

5.3.1. Resultados dos tempos de respostas

Embora a análise de variância não tenha demonstrado interação entre o tipo de palavra e a condição do *prime* ($p=0.11$, $F=1.41$) e nem no fator do tipo de palavra ($p=0.27$, $F=0.29$), as condições do *prime* demonstraram diferenças extremamente significantes ($p<0.0001$, $F=8.30$).

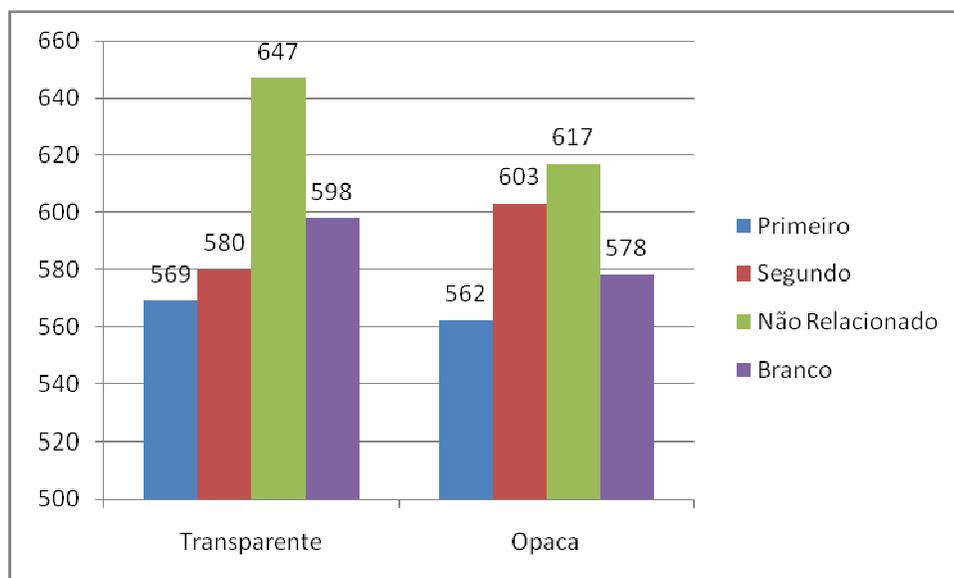


Gráfico 1: Média dos tempos para as palavras transparentes e opacas no Experimento1.

Inicialmente, as palavras transparentes demonstraram efeitos de facilitação tanto no encontro com o primeiro constituinte, como com o segundo constituinte. Este resultado foi observado através da análise do teste-t que demonstrou diferenças significativas entre as palavras transparentes precedidas pelo primeiro constituinte e o grupo com constituinte não relacionado ($p<0.0001$, $t=8.174$), e também na comparação com o *prime* em branco ($p=0.0008$, $t=3.352$). As palavras transparentes precedidas pelo segundo constituinte também demonstraram tal facilitação tanto quando comparadas ao grupo com constituinte não relacionado ($p<0.0001$, $t=7.095$), como quando comparadas ao grupo com o *prime* em branco ($p=0.03$, $t=2.134$). Adicionalmente, o teste t também indicou não haver diferença

significativa entre as condições de primeiro e segundo constituintes nas palavras transparentes ($p=0.24$, $t=1.1$). Ou seja, a partir da análise do teste t podemos ver um efeito de facilitação nas palavras transparentes tanto na apresentação do primeiro, como do segundo constituinte, se comparados com as duas condições controle.

Entretanto, as palavras opacas não apresentaram o mesmo comportamento. As palavras precedidas pelo *prime* de primeiro constituinte, por exemplo, não demonstraram diferenças significativas com aquelas que apresentaram o *prime* em branco ($p=0.10$, $t=1.626$). Isto quer dizer que não foi observado efeito de facilitação robusto nem com a apresentação do primeiro constituinte. Em adição, as palavras antecedidas pelo segundo constituinte demonstraram tempos significativamente mais longo do que aquelas com o *prime* em branco ($p=0.0073$, $t=2.69$) e não apresentaram diferenças significativas com os tempos das palavras antecedidas por um constituinte não relacionado ($p=0.19$, $t=1.30$). Consequentemente, as palavras opacas apresentadas posteriormente ao segundo constituinte parecem demonstrar inibição e não facilitação. Por fim, ao contrário das palavras transparentes, as palavras opacas apresentaram diferenças significativas entre os tempos das palavras antecedidas pelo primeiro e segundo constituintes ($p<0.0001$, $t=3.96$).

Em relação aos grupos controle, observamos uma diferença extremamente significativa entre os tempos de resposta para as palavras antecedidas por um constituinte não relacionado e aquelas antecedidas por uma tela em branco, tanto nas palavras transparentes ($p<0.0001$, $t=5.48$), quanto nas palavras opacas ($p<0.0001$, $t=4.05$). Isto quer dizer que as palavras com antecedente não relacionado demoraram significativamente mais do que aquelas com a condição em branco, demonstrando provavelmente um efeito inibitório.

5.3.2. Resultados dos índices de resposta

Os índices de resposta apresentaram um resultado coerente com a análise dos tempos de resposta. As palavras opacas apresentaram significativamente mais erros do que as palavras transparentes ($p=0.001$, $X^2= 10.42$).

	SIM	NÃO
Transparente	1668	60
Opaca	1627	101

Tabela 1: Índices de resposta para as palavras transparentes e opacas no Experimento 1.

Na análise entre as condições, também foi observada diferença significativa entre as palavras transparentes e opacas ($p=0.007$, $X^2=12.03$). Através dos índices é possível observar que as palavras opacas precedidas pelo primeiro e segundo constituinte apresentaram significativamente mais erros do que as palavras transparentes nestas mesmas condições.

	Primeiro	Segundo	Não-relacionado	Branco
Transparente	5	11	35	9
Opaca	18	34	32	17

Tabela 2: Índices de erros nas respostas entre as diferentes condições no Experimento 1.

5.3.3. Discussão

Embora os dados não sejam robustos o suficiente para demonstrar interação em uma análise de variância, foi possível, através deste primeiro experimento, observar diferenças entre os tempos de resposta para as palavras transparentes e opacas. Acreditamos que este resultado foi obtido em parte devido ao SOA longo utilizado. Uma vez que o estímulo *prime* foi visualizado por mais tempo, os traços semânticos provavelmente foram amplamente acessados e, portanto, ajudaram no acesso da palavra transparente tanto com o encontro do primeiro constituinte como com o encontro do segundo. Já nas palavras opacas, não há participação dos significados dos constituintes no significado da palavra inteira, conseqüentemente, o *prime* dos constituintes não demonstrou efeitos de facilitação no encontro posterior aos compostos.

Os índices de resposta também corroboram para esta interpretação, uma vez que as palavras opacas apresentaram índices maiores de erros com o encontro do primeiro e do segundo constituintes. Desta forma, os resultados apontam para a facilitação de erros na

decisão lexical quando há a apresentação do constituinte e o posterior acesso à palavra composta opaca.

5.4. Experimento 2

Participaram do segundo experimento 46 alunos (35 homens e 11 mulheres) com idade entre 18 a 33 anos (média 21.52) da graduação e da pós-graduação da Universidade de Tecnologia de Tóquio no Japão, recebendo uma remuneração em dinheiro pela participação. Os dados de dois sujeitos precisaram ser retirados da análise por ultrapassarem a margem de 10% de erros e por apresentarem tempos de reação mais altos do que média dos outros sujeitos. Assim, somente os dados de 44 sujeitos foram utilizados para a análise.

Assim como o primeiro experimento, as respostas erradas foram excluídas da análise sobre os tempos de resposta, ou seja, apenas as respostas SIM para as palavras é que foram consideradas. Os tempos de resposta assim como a acuidade das respostas foram analisados estatisticamente. Os tempos de respostas que ultrapassaram 2.5 da média de um participante foram ajustados para a média mais 2.5 deste valor.

5.4.1. Resultados dos tempos de respostas

Conforme esperado, neste segundo experimento, as palavras transparentes e opacas apresentaram padrões de resposta semelhantes. Uma análise de variância (*Two-way ANOVA*) entre sujeitos com as médias dos tempos de resposta demonstrou não haver interação entre a condição do *prime* e o tipo de palavra ($p=0.547$; $F=0.71$). Além disso, no contraste entre palavras transparentes e opacas também não foram encontradas diferenças significativas ($p=0.528$; $F=0.40$). Por outro lado, as condições do *prime* demonstraram diferenças extremamente significantes ($p<0.0001$, $F=0.40$) entre si.

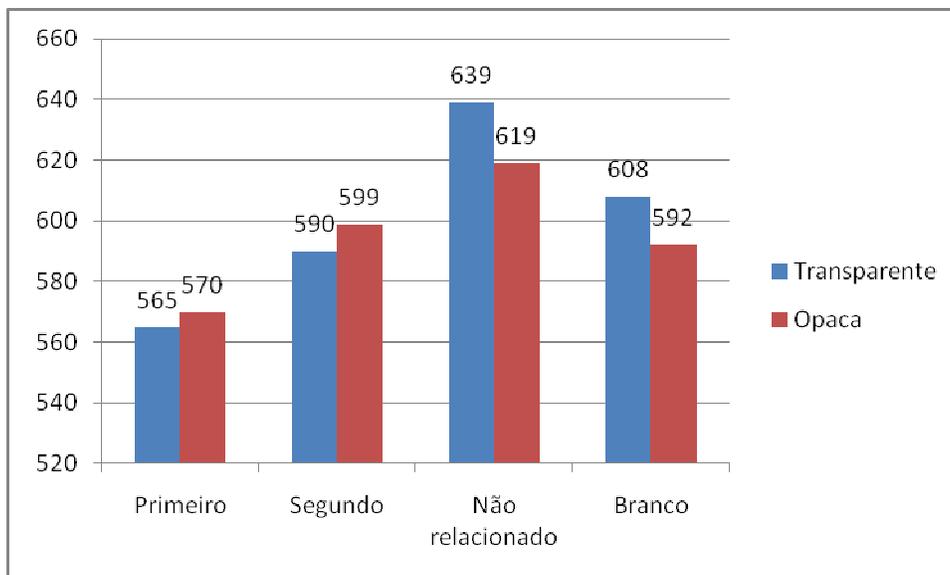


Gráfico 2: Médias dos tempos para as palavras transparentes e opacas no Experimento 2.

Para observar mais detalhadamente as diferenças dentro das condições do *prime*, efetuamos um teste-t. A análise do teste-t indicou diferenças significativas entre os tempos de resposta para as palavras antecedidas pelo primeiro constituinte e aquelas antecedidas por um constituinte não relacionado, tanto nas palavras transparentes ($p < 0.0001$, $t = 7.55$), quanto nas palavras opacas ($p < 0.0001$, $t = 4.51$). Este contraste também foi observado entre o grupo com primeiro constituinte e o grupo antecedido por uma tela em branco, tanto nas palavras transparentes ($p < 0.0001$, $t = 4.69$), quanto nas palavras opacas ($p = 0.03$, $t = 2.060$). Ou seja, o resultado indica um efeito de facilitação na decisão lexical das palavras transparentes e opacas com a apresentação anterior do primeiro constituinte.

Todavia, o mesmo não foi observado ao contrastar as palavras antecedidas pelo segundo constituinte. Neste caso, encontramos diferença significativa apenas entre as palavras transparentes antecedidas pelo segundo constituinte e aquelas antecedidas por um constituinte não relacionado ($p < 0.001$, $t = 4.793$). Não foram observadas diferenças significativas entre os pares com segundo constituinte e aqueles com a condição em branco (Transparente: $p = 0.06$, $t = 1.882$; Opaca: $p = 0.52$; $t = 0.6343$). As palavras opacas antecedidas pelo segundo constituinte também não apresentaram diferença significativa quando comparadas aos pares com constituinte não relacionado ($p = 0.0575$, $t = 1.903$).

Ademais, quando comparamos as palavras antecedidas pelo primeiro constituinte e aquelas antecedidas pelo segundo constituinte observamos que o primeiro grupo apresentou

tempos significativamente mais rápidos do que o segundo grupo, tanto nas palavras transparentes ($p=0.006$, $t=2.711$), quanto nas palavras opacas ($p=0.0072$; $t=2.694$). Isto quer dizer que as palavras antecedidas pelo primeiro constituinte apresentaram efeitos de facilitação e foram significativamente mais rápidas do que aquelas antecedidas pelo segundo constituinte que, conseqüentemente, não apresentaram efeitos de facilitação.

Por outro lado, também contrastamos os grupos controle e encontramos, assim como o primeiro experimento, tempos de reação significativamente mais longos para as palavras com o *prime* de constituinte não relacionado do que aquelas com a condição em branco, tanto nas palavras transparentes ($p=0.002$, $t=3.033$), quanto nas palavras opacas ($p=0.01$, $t=0.634$). Este resultado fortalece a hipótese de que as palavras antecedidas por um constituinte não relacionado causam efeito inibitório no acesso posterior às palavras.

5.4.2. Resultados dos índices de respostas

Os índices de resposta também foram observados e organizados em uma tabela de contingência. Um teste de Qui-quadrado demonstrou não haver diferenças significativas entre os índices das palavras transparentes e opacas ($X^2= 3.76$; $p=0.052$).

	SIM	NÃO
Transparente	1509	75
Opaca	1483	101

Tabela 3: Índices de resposta para as palavras transparentes e opacas no Experimento 2.

Na análise entre as condições da palavra *prime*, também não foram observadas diferenças significativas em relação ao número de erros entre os tipos de condição ($X^2=6.078$; $p=0.10$).

	Primeiro	Segundo	Não-relacionado	Branco
Transparente	13	12	35	15
Opaca	19	31	38	13

Tabela 4: Índices de erros nas respostas entre as diferentes condições no Experimento 2.

5.4.3. Discussão

As médias dos tempos demonstraram padrões semelhantes entre as palavras transparentes e opacas e foram observados efeitos de facilitação no encontro anterior com o primeiro constituinte. No entanto, os tempos de resposta para as palavras que foram precedidas pelo segundo constituinte foram muito parecidos com os tempos encontrados para as palavras antecedidas por uma tela em branco, isto é, não apresentaram efeitos de facilitação robustos. Este resultado talvez aponte para uma facilitação natural encontrada no reconhecimento do primeiro constituinte ortográfico, já que a leitura horizontal é efetuada da esquerda para a direita (fig.20).

閉 店
閉店 閉店

Figura 22: Exemplificação idealizada sobre a visualização dos experimentos.

Após encontrar o primeiro constituinte, a palavra composta é visualizada e apresenta facilitação, uma vez que o constituinte a ser processado primeiro acabou de ser acessado anteriormente. Contudo, no caso do segundo constituinte como *prime*, é necessário processar a palavra composta a partir do primeiro constituinte e não a partir do constituinte que acabou de ser acessado. Por conseguinte, os tempos de resposta para as palavras que foram precedidas pelo primeiro constituinte são significativamente mais rápidos, enquanto o mesmo não ocorre com a apresentação do segundo constituinte.

Este resultado parece indicar a influência de fatores ortográficos. Todavia, ainda que os valores não tenham sido robustos o suficiente para demonstrar diferenças significativas, é possível observar que as médias para as palavras com o segundo constituinte como *prime* são mais baixas do que aquelas com constituinte não relacionado. Isto indica que o efeito encontrado não foi simplesmente ortográfico, pois caso fosse, as médias para os dois grupos citados anteriormente deveriam ser semelhantes.

Adicionalmente, a tabela de contingência em conjunto com uma análise de Qui-quadrado não demonstrou diferenças significativas nos índices de erros entre as palavras transparentes e opacas. Acreditamos que o curto período de permanência do prime não foi o suficiente para um acesso amplo dos traços semânticos e, portanto, a decisão lexical não foi influenciada no grupo das palavras opacas.

Em relação aos grupos controles, observamos novamente um efeito inibitório com a apresentação de um constituinte não relacionado. Desta forma, nos perguntamos se o uso de um constituinte não relacionado é o grupo controle ideal para se observar efeitos genuínos de facilitação. Por outro lado, a partir do uso deste grupo de controle foi possível realizar uma análise mais apurada quando não encontramos diferenças significativas com a condição em branco.

6. Conclusões

Sabemos que no processo de compreensão de uma palavra, é necessário ligar uma forma arbitrária a um significado. Contudo, a maneira e o momento em que este significado emerge ainda é fruto de ricas questões e diversas teorias, como foi apresentado ao longo deste trabalho. Com o crescente número de pesquisas nas últimas décadas demonstrando a natureza computacional da mente, é natural que entre as hipóteses sugeridas se teorize uma estrutura imersa por detrás da passagem de um som ou um símbolo para um significado. A busca e o encontro desta estrutura linguística poderiam trazer ricas respostas sobre a linguagem humana que permanecem como mistérios até hoje.

No presente estudo, procuramos responder algumas questões sobre a estrutura da linguagem através da observação da interface entre a sintaxe e semântica em dois experimentos de *priming* com palavras compostas da língua japonesa. Guiados pelos pressupostos da teoria da Morfologia Distribuída, prevemos a existência de um mecanismo linguístico que trabalha com unidades menores do que uma palavra e, por conseguinte, defendemos a decomposição morfológica no acesso lexical. Posteriormente, também argumentamos que no primeiro estágio do processamento há apenas uma estrutura sintática/morfológica com traços semânticos mínimos. Deste modo, os traços semânticos mais amplos que definem o significado de uma palavra participam apenas em estágios posteriores e não nos estágios iniciais.

Os resultados de ambos os experimentos realizados em nosso estudo demonstraram efeitos de facilitação na apresentação de palavras compostas após um de seus constituintes. Caso a teoria Lexicalista mais radical estivesse correta, o acesso aos constituintes, apresentados separadamente, seria independente do acesso às palavras e, conseqüentemente, não seria possível encontrar efeitos de *priming*. Os resultados encontrados, entretanto, parecem apontar para a participação do conteúdo morfológico no processamento das palavras compostas em japonês.

Adicionalmente, os resultados também favorecem uma das propostas da Morfologia Distribuída que prediz a participação dos traços semânticos apenas em momentos posteriores do acesso lexical. Foi observado no primeiro experimento com SOA longo

(150ms), efeitos de facilitação nas palavras transparentes tanto com a apresentação do primeiro constituinte, como com a apresentação do segundo. Porém, o mesmo não foi observado nas palavras opacas. O resultado parece indicar que em 150ms os traços semânticos principais de um constituinte parecem ser acessados e, em decorrência, podem ajudar o acesso das palavras transparentes, mas não o acesso das palavras opacas. Já no experimento com SOA mais curto (60ms), encontramos padrões similares de tempos de respostas nas palavras transparentes e opacas. Ambas as condições demonstraram efeitos de facilitação com a apresentação do primeiro constituinte, mas este comportamento não foi observado com a apresentação do segundo constituinte. Acreditamos que este resultado demonstre um efeito natural no reconhecimento de uma palavra escrita, que por ser processada da esquerda para a direita, oferece maior vantagem para o reconhecimento repetido do primeiro constituinte.

Ademais, os índices de erro apresentaram resultados consistentes com os padrões demonstrados nos índices de respostas: diferenças significativas entre as palavras transparentes e opacas somente no Experimento 1. Acreditamos que este resultado indique que no Experimento 1, com SOA longo, os traços semânticos dos constituintes como *prime* foram acessados, e portanto, facilitaram a ocorrência de erros nas palavras opacas. Todavia, no Experimento 2, devido ao curto espaço de tempo de permanência do *prime*, não foi possível acessar amplamente os traços semânticos e, em consequência, a decisão lexical das palavras opacas não foi influenciada.

Entretanto, em busca de uma observação mais pura dos fatores morfológicos sem a participação dos fatores semânticos, prevemos utilizar em experimentos posteriores medidas abaixo de 50ms para o tempo de permanência da palavra *prime*. De acordo com o estudo de Perea e Gordon (1997) a influência de fatores semânticos também pode ser vista em experimentos de *priming* com SOA de 67ms e 50ms. Em contraste, o uso de medidas mais baixas (como 33ms) parece ser mais produtivo na observação de fatores puramente morfológicos. Além disso, para evitar também a participação e a influência de fatores ortográficos, que parecem ter influenciado os resultados dos experimentos, em especial do segundo, seria aconselhável o emprego da modalidade oral.

O presente trabalho apresentou também dados que incitam questionamentos em relação ao grupo controle em um paradigma de *priming*. Observamos que as palavras compostas apresentadas após um constituinte não relacionado demonstraram tempos de resposta bastante longos. Logo, com o uso deste grupo como controle, haverá uma facilidade muito maior de se encontrar um efeito de facilitação, que, no entanto, pode ser ilusória se comparada com uma condição em branco. Este último parece ser mais neutro, pois permite recolher as médias decorrentes do reconhecimento de uma palavra, sem interferência forte da apresentação anterior de outra palavra. Por outro lado, em nosso estudo, a presença do grupo controle com constituinte não relacionado nos ajudou a analisar as médias que, embora não demonstrassem facilitação se comparadas à condição em branco, não eram altas o suficiente para demonstrar inibição. Desta forma, na organização do grupo controle em um experimento de *priming*, é necessário se atentar para os variados tipos de condição que oferecem resultados diferentes e que podem vir a ser o *baseline* ideal de acordo com o *design* e o material do experimento.

Certamente outras questões ainda podem surgir a partir dos procedimentos e dos resultados do presente estudo. Uma destas questões diz respeito às propriedades fonológicas que apresentam extrema relevância no processo de reconhecimento das unidades morfológicas e o acesso ao significado. Neste trabalho, porém, não controlamos as características fonológicas por algumas razões, em especial devido à grande falta de correspondência entre o som representado por um constituinte sozinho e aquele representado pelo mesmo constituinte dentro de uma palavra composta na língua japonesa.

Em adição, a grande diversidade de padrões estruturais das palavras compostas oferece vasto material para estudo, mas que, infelizmente, não vem sendo muito buscada nas últimas décadas. As relações de semelhança entre compostos e expressões idiomáticas apresentam também um campo interessante para maior investigação. Os compostos, assim como as expressões idiomáticas, são expressões restringidas pelo contexto e que possuem altos índices de arbitrariedade no significado (Kavka, 2009). Além disso, é possível identificar, com muita facilidade, a supressão da referencialidade sintática nos dois tipos de estrutura.

De acordo com alguns estudos (Cacciari e Tabossi, 1988; Titone e Connine, 1999; Fogliata *et al.*, 2007), a estrutura de uma expressão idiomática é processada e o significado literal é acessado em um primeiro momento. Todavia, em seguida, o significado figurado é buscado e o significado literal é inibido. Análises como esta parecem apontar para a teoria da Morfologia Distribuída, a qual prediz um momento de reanálise para as estruturas com significado não-composicional especial. Contudo, acreditamos que seja necessária a realização de outros estudos utilizando técnicas experimentais que busquem analisar a semelhança entre compostos e expressões idiomáticas, assim como investigar mais profundamente a interface entre a sintaxe e a semântica.

Esperamos, no entanto, que este estudo não somente colabore com o campo dos estudos das línguas orientais no Brasil, mas também que exerça contribuição nos estudos experimentais sobre processamento morfológico que defendem um estágio puramente sintático anterior a participação de fatores semânticos.

7. Referências

AMANO, S.; KONDO, T.(Eds) **Nihongo no goitokusei** (Propriedades Lexicais do japonês) NTT Database series. Tokyo: Sanseido, 1999.

BUTTERWORTH, B. Lexical representation. In. _____. **Language production: Volume 2: Development, Writing and Other Language Processes**. London: Academic Press, 1983. Cap. 11, p.257-294.

CACCIARI, C.; TABOSSI, P. The Comprehension of Idioms. **Journal of Memory and Language**, 27, p.668-683, 1988.

CARAMAZZA, A.; LAUDANNA, A.; ROMANI, C. Lexical access and inflectional morphology, **Cognition**, 28, p.297-332, 1988.

CHIALANT, D.; CARAMAZZA, A. Where is morphology and how is it processed? The case of written word recognition, In: Feldman, L. B. (ed.): **Morphological Aspects of Language Processing**. Hillsdale, NJ: Erlbaum Press. 1995. Cap.3, p.55-76.

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

DI SCIULLO, A.M.; WILLIAMS, E. **On the definition of word**. Cambridge: MIT Press, 1987.

EMBICK, D.; NOYER, R. Distributed Morphology and the Syntax-Morphology Interface. In Ramchand, G.; Reiss, C. (eds.). **The Oxford Handbook of Linguistic Interfaces**. Oxford: Oxford University Press, 2007, p. 289-324.

ENDO, Y. **Locality and Information Structure: A cartographic approach to Japanese**. Amsterdam, Philadelphia: John Benjamin Publishing Company, 2007.

FIORENTINO, R.; FUND-REZNICEK, E. Masked morphological priming of compound constituents. **The Mental Lexicon**. 4(2), p.159-193, 2009.

FIORENTINO, R.; POEPEL, D. Compound words and structure in the lexicon. **Language and Cognitive Processes**, 22, 6, 953-1000, 2007.

FODOR, Jerry. **Concepts: where cognitive science went wrong**. New York: Oxford University Press, 1998.

FOGLIATA, A. *et al.* The time course of idiom processing. **Neuropsychologia**, 45, p. 3215-3222, 2007.

FORSTER, K. I.; MOHAN, K.; HECTOR, J. The mechanics of masked priming. In S. Kinoshita; S. J. Lupker (eds.): **Masked priming: The state of the art**. Hove, UK: Psychology Press, 2003, Cap.1, p. 3-37.

FRANCA, A. I. *et al.* A neurofisiologia do acesso lexical: palavras em português. **Veredas on-line – Psicolinguística**, 2, p.34-49. 2008.

GARCIA, D. C. de **Elementos estruturais no acesso lexical: o reconhecimento de palavras multimorfêmicas no Português Brasileiro**. 2009. Dissertação (Mestrado em Linguística), Universidade Federal do Rio de Janeiro, Rio de Janeiro.

GOMES, J. N. **A direcionalidade no relacionamento semântico: um estudo de ERP**. 2009. Dissertação (Mestrado em Linguística), Universidade Federal do Rio de Janeiro, Rio de Janeiro.

HALLE, M.; MARANTZ, A. Distributed Morphology and the Pieces of Inflection. In: Hale, K.; Keyser, S.J. (ed.): **The View from Building 20**. Cambridge, Massachusetts: MIT Press, 1993, p.111-176.

HARLEY, H.; NOYER, R. State-of-the-article: Distributed Morphology. **Glott International**, 4.4: p.3-9. 1999.

HUANG, S. H.; CHEN, J.; REN-YI, L. **Eye movements during reading of Chinese V-V compounds**. Proceedings of the 13th International Conference on the Processing of East Asian Languages, 9-11 October, Beijing, China, 2009.

JOYCE, T.; OHTA, N. The morphology of two-kanji compound words: data from a word-formation classification survey. **Tsukuba Psychological Research**, 22, p.45-61, 1999.

JOYCE, T. **Lexical access and the mental lexicon for two-kanji compound words: A priming paradigm study**. Proceedings of the 2nd International Conference on Cognitive Sciences and 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference, 27-30 July, Tokyo, Japan, p. 511-514. 1999.

JOYCE, T. Constituent-morpheme priming: Implications from the morphology of two-kanji compound words. **Japanese Psychological Research**, vol.4, n.2, p.79-90. 2002.

KATO, M. A. Teoria sintática: de uma perspectiva de “-ismos” para uma perspectiva de programas. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, vol.13, n.2, São Paulo, 1997.

KAGEYAMA, T. Word Formation. In. Tsujimura, N. (ed.): **The Handbook of Japanese Linguistics**. Malden, MA: Blackwell Publishers, 2002, Cap.10, p.297-325.

KAVKA, S. Compound and idiomatology. In: Lieber, R.; Stekauer, P. (eds): **The Oxford Handbook of Compounding**. Oxford: Oxford University Press, 2009, Cap.1, p.19-33.

KESS, J. F.; MIYAMOTO, T. **The Japanese Mental Lexicon: Psycholinguistic studies of kana and kanji processing**. Amsterdam, Philadelphia: John Benjamins Publishing Company, 1999.

LIBBEN, G. Semantic transparency in the processing of compounds: consequences for Representation, Processing and Impairment. **Brain and Language**, 61, 1, p.30-44. 1998.

LIBBEN, G.; JAREMA, G. (eds.): **The representation and processing of compound words**. Oxford: Oxford University Press, 2005.

LONGTIN, C.M.; SEGUI, J.; HALLÉ, P.A. Morphological priming without morphological relationship. **Language and Cognitive Processes**. vol.18, p.313-334, 2003.

MAIA, M., LEMLE, M. FRANÇA, A.I. Efeito stroop e rastreamento ocular no processamento de palavras. **Ciência e Cognição**. Vol.12: 02-17, 2007.

MARANTZ, A. **Cat as a phrasal idiom**. MIT, 1996.

_____. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In: Dimitriadis, A.; Siegel, L., et al. (eds.): **University of Pennsylvania Papers in Linguistics**, Vol. 4.2, Proceedings of the 21 st. Annual Penn Linguistics Colloquium, p. 201-225. 1997.

_____. **Words and Things**. MIT, 2001.

_____. **Phases and Words**. (a aparecer).

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

MEDEIROS, A. B. **Traços morfossintáticos e subespecificação morfológica na gramática do português: um estudo das formas participais**. 2008. Tese (Doutorado em Linguística), Universidade Federal do Rio de Janeiro, Rio de Janeiro.

MEYER, D. E.; SCHVANEVELDT, R. W. Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. **Journal of Experimental Psychology**. Vol. 90(2), p.227-234, 1971.

MORTON, J.; SASANUMA, S. Lexical access in Japanese. In: Henderson, L. (ed.): **Orthographies and Reading: Perspectives from Cognitive Psychology, Neuropsychology, and Linguistics**. Hillsdale, NJ: Lawrence Erlbaum Associates, 1984, p.25-42.

MORTON, J. *et al.* The organization of the lexicon in Japanese: single and compound kanji. **British Journal of Psychology**, 83, p.517-531, 1992.

NAKAYAMA, T. Sentence Processing. In. Tsujimura, N. (ed.): **The Handbook of Japanese Linguistics**. Malden, MA: Blackwell Publishers, 2002, Cap.10, p.398-424.

NIEMI, J.; LAINE, M.; TUOMINEN, J. Cognitive morphology in Finnish: Foundations of a new model. **Language and Cognitive Processes**, 9, 3, p.423-446, 1994.

PEREA, M.; GORTON, A. Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. **Cognition**, 62, p.223-240, 1997.

PINKER, S. **The Language Instinct. How the mind creates language**. New York: William Morrow and Company, 1994.

_____. **Words and Rules: The Ingredients of Language**. New York: Basic Books. HarperCollins, 1999.

PYLKKÄNEN, L.; OKANO, K. **The nature of abstract orthographic codes: Evidence from masked priming and magnetoencephalography**. (a aparecer).

RASTLE, K. *et al.* Morphological and semantic effects in visual word recognition: A time course study. **Language and Cognitive Processes**, 15, p.507-538, 2000.

RASTLE, K., DAVIS, M. H., NEW, B. The broth in my brother's brothel: morpho-orthographic segmentation in visual word recognition. **Psychonomic Bulletin and Review**, 11 (6), p.1090-1098, 2004.

ROGERS, H. **Writing Systems, a linguistic approach**. Oxford: Blackwell Publishing, 2005.

SANDRA, D. On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. **Quarterly Journal of Experimental Psychology**, 42A, 529-567, 1990.

SAKUMA, N.; ITOH, M.; SASANUMA, S. Puraimingu Paradaimu ni Yoru Kanji Tango no Ninchi Yunitto no Kento [Recognition Units of Kanji Words: Priming Effects on Kanji Recognition]. **Shinrigaku Kenkyū**, 60, p.1-8, 1989.

SHIBATANI, M. **The languages of Japan**. Cambridge: Cambridge University press, 1990.

SHREUDER, R.; BAAYEN, H. Modelling Morphological Processing. In: Feldman, L.B. (ed.): **Morphological Aspects of Language Processing**. Hillsdale: Lawrence Erlbaum Associates Publishers, 1995, p.131-154.

SHOOLMAN, N.; ANDREWS, S. Racehorses, reindeers and sparrows: Using masked priming to investigate morphological influences on word identification. In: Kinoshita, S.; Lupker, S. (eds.), **Masked priming: The state of the art**. New York, Hove: Psychology Press, 2003.

TAFT, M.; FORSTER, K. I. Lexical storage and retrieval of prefixed words. **Journal of Verbal Learn Verbal Behavior.**, 14(6), p. 638-647, 1975.

TAMAOKA, K.; HATSUZUKA, H. The effects of morphological semantics on the processing of Japanese two-kanji compound words. **Reading and Writing: an interdisciplinary journal**, 10, p. 293-322, 1998.

TITONE, D. A.; CONNINE, C. M. On the compositional and nondecompositional nature of idiomatic expressions. **Journal of Pragmatics**, 31, p.1655-1674, 1999.

UENO, K. **Eigo-go no kenkyū**. Tokyō: Kenkyūsha, 1980.

VOLPE, M. J. **Japanese Morphology and its theoretical consequences: Derivational morphology in Distributed Morphology** Ph.D. thesis, Stony Brook University, 2005

YAMASHITA, H. The effects of word-order and case marking information on the processing of Japanese. **Journal of Psycholinguistic Research**. v.26, n.2, p.163-188, 1997.

ZHANG, T.; VAN HEUVEN, W. J. B. **The role of semantic transparency on Chinese compound processing: evidence for morphemic decomposition**. Proceedings of the 13th International Conference on the Processing of East Asian Languages, 9-11 October, Beijing, China, 2009.

ZHOU, X. *et al.* Morphology, orthography, and phonology in reading Chinese compound words. **Language and Cognitive Processes**, 14, 525-565, 1999.

ZWITSERLOOD, P. Processing and representation of Dutch compounds: Effects of semantic transparency. **Language and Cognitive Processes**, 9(3), 341-368. 1994.

Dicionário virtual Jim Breen's WWWJDIC

<http://www.csse.monash.edu.au/~jwb/cgi-bin/wwwjdic.cgi?1C>

APÊNDICE - Palavras opacas

Composto	Transcrição	Tipo	Familiaridade	Média do questionário	Significado do 1o constituinte	Significado do 2o constituinte	Significado do composto
駄目	dame	Opaca	5.969	1.5	envio a cavalo	olho	inutilizável
大切	taisetsu	Opaca	6.438	1.55	grande	cortar	importante
経済	keizai	Opaca	6.281	1.75	longitude	terminar	economia
支度	shitake	Opaca	5.688	1.8	suporte	nível	preparação
普段	fudan	Opaca	5.844	1.8	universal	degrau	comum
人參	ninjin	Opaca	5.5	1.9	peessoa	perplexidade, ir, visitar	cenoura
包丁	hōchō	Opaca	5.906	2.1	embrulho	rua, cidade	faca
全然	zenzen	Opaca	6.312	2.15	tudo	se, neste caso	nenhum, tudo
帽子	bōshi	Opaca	5.875	2.2	chapéu	criança	chapéu
交番	kōban	Opaca	6.062	2.2	mistura	número	posto de polícia
丈夫	jōbu	Opaca	5.531	2.4	comprimento	marido, homem	bem, sadio
貿易	bōeki	Opaca	6.125	2.4	troca	fácil	comércio
料理	ryōri	Opaca	6.438	2.4	material	razão	comida
自由	jiyū	Opaca	6.438	2.4	o próprio	razão	liberdade
風俗	fūzoku	Opaca	5.719	2.45	vento	comum	costumes
利息	risoku	Opaca	5.812	2.45	vantagem	respiração	juros
首相	shushō	Opaca	6	2.5	pescoço	mútuo	Primeiro Ministro
床屋	tokoya	Opaca	5.719	2.5	cama	telhado, casa	barberaria
味方	mikata	Opaca	6.156	2.5	Sabor	direção	amigo
約束	yakusoku	Opaca	6.438	2.55	promessa	pacote	promessa
文化	bunka	Opaca	6.25	2.55	sentença	troca	cultura
景気	keiki	Opaca	6.156	2.55	cenário	ar, atmosfera, espírito	condição, estado
小説	shōsetsu	Opaca	6.219	2.55	pequeno	teoria	romance
猶予	yūyo	Opaca	5.125	2.6	adicionalmente	previamente	adiamento
会社	kaisha	Opaca	6.469	2.65	encontro	sociedade	empresa
厚生	kōsei	Opaca	5.031	2.65	denso, rico	vida	bem-estar
柔道	jūdō	Opaca	6.312	2.65	suave	caminho	Judô
記念	kinen	Opaca	6.188	2.7	escrita	desejo, idéia	comemoração, memória
都合	tsugō	Opaca	6.156	2.75	capital	combinar, unir	circunstância
正月	shōgatsu	Opaca	6.469	2.75	certo	lua, mês	Ano Novo
機関	kikan	Opaca	5.344	2.8	mecanismo	conexão	organização, órgão
本当	hontō	Opaca	6.562	2.8	livro, base	certo, apropriado	verdade
案内	annai	Opaca	6.375	2.9	plano, sugestão	dentro	guia
将来	shōrai	Opaca	6.438	2.9	líder, comando	vir	futuro
軽率	keisotsu	Opaca	5.656	2.92	leve	proporção	descuido
信号	shingō	Opaca	6.438	2.95	fê, confiança	número, item	sinal, semáforo

APÊNDICE - Palavras transparentes

Composto	Transcrição	Tipo	Familiaridade	Média do questionário	Significado do 1o constituinte	Significado do 2o constituinte	Significado do composto
納税	nōzei	Transparente	5.719	6.5	pagar	imposto	pagamento de imposto
作詞	sakushi	Transparente	5.938	6.5	escrever, fazer	palavras, poesia	escrita de letras(de música)
新設	shinsetsu	Transparente	5.688	6.5	novo	estabelecimento	novo estabelecimento
表紙	hyōshi	Transparente	6.094	6.5	superfície	papel	capa
白米	hakumai	Transparente	5.844	6.5	branco	arroz	arroz polido, arroz branco
終戦	shūsen	Transparente	5.688	6.5	término	guerra	fim da guerra
失恋	shitsuren	Transparente	6.125	6.55	erro, perda	romance, amor	amor não correspondido
未知	michi	Transparente	5.656	6.6	ainda não	conhecido	desconhecido
昼食	chūshoku	Transparente	6.375	6.6	tarde	comida	refeição da tarde
少数	shōsū	Transparente	6.031	6.65	pouco	número	minoridade
休学	kyūgaku	Transparente	5.688	6.65	descanso	colégio	ausência da escola
飲酒	inshu	Transparente	6.094	6.65	beber	alcoól	beber alcoól
牛乳	gyūnyū	Transparente	6.562	6.7	vaca	leite	leite de vaca
外泊	gaihaku	Transparente	5.844	6.7	fora	pernoite	pernoitar fora
公認	kōnin	Transparente	5.969	6.7	público, oficial	reconhecimento	reconhecimento oficial, público
即死	sokushi	Transparente	5.875	6.7	instante	morte	morte imediata
重病	jūbyō	Transparente	5.656	6.75	pesado, grave	doença	doença grave
投球	tōkyū	Transparente	5.906	6.75	arremesso	bola	lançamento de bola
早朝	sōchō	Transparente	6.156	6.75	cedo	manhã	de manhã cedo
夜勤	yakin	Transparente	5.969	6.75	noite	trabalho	trabalho noturno
進路	shinro	Transparente	6.062	6.75	progresso	rota	rota, curso
登山	tozan	Transparente	6.125	6.75	escalar	montanha	escalar montanha
寢室	shinshitsu	Transparente	5.656	6.75	dormir	quarto	quarto (de dormir)
再婚	saikon	Transparente	5.812	6.8	de novo	casamento	casar novamente
育児	ikuji	Transparente	5.906	6.8	cuidar	criança	assistência a infância
着陸	chakuriku	Transparente	5.875	6.8	chegar	território, continente	aterrissagem
急増	kyūzō	Transparente	5.625	6.8	rápido	aumento	aumento rápido
辞職	jishoku	Transparente	6.062	6.85	resignação	trabalho	resignação
乗馬	jōba	Transparente	5.75	6.85	montar, subir	cavalo	montado a cavalo
返金	henkin	Transparente	5.812	6.85	devolução	dinheiro	reembolso
閉店	heiten	Transparente	5.969	6.85	fechamento	loja	fechamento de uma loja
冷水	reisui	Transparente	5.656	6.85	gelado, frio	água	água fria
洗顔	sengan	Transparente	5.875	6.9	lavar	rosto	lavar face
脱毛	datsumō	Transparente	5.688	6.9	remoção	cabelo	remoção de cabelo
入国	nyūkoku	Transparente	5.719	6.9	entrada	país	entrada no país
高熱	kōnetsu	Transparente	5.969	6.9	alto	febre	febre alta

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)