

# A RESOLUÇÃO DA REFERÊNCIA EM CASOS DE AMBIGUIDADE ESTRUTURAL: A INFLUÊNCIA DA INFORMAÇÃO VISUAL

Maria Luiza Cunha Lima

Cláudia Brandão Vieira

**Resumo:** Este trabalho analisa como a informação visual pode afetar a identificação de referentes temporariamente ambíguos em português brasileiro. O experimento aqui detalhado é uma adaptação do trabalho de Tanenhaus *et al.* (1995) a respeito da influência do contexto visual no processamento de frases temporariamente ambíguas do inglês. Esses autores comprovaram, por meio do estudo de padrões de movimento ocular, que a informação disponível visualmente atua no início do processamento da língua falada, facilitando a identificação de referentes. Tais resultados permitem que se questione o quadro teórico previsto pelos modelos de processamento serial (Frazier & Fodor, 1978; Frazier, 1979), em que o *parser* utilizaria inicialmente apenas informações gramaticais durante o processamento de sentenças. Neste trabalho, foi realizado um experimento de rastreamento ocular que explorava o *paradigma do mundo visual* (Cooper, 1974). No experimento, estímulos auditivos solicitavam que os participantes movessem objetos em diferentes cenários. Parte dos estímulos possuíam ambiguidades temporárias, devido à possibilidade de se interpretar um sintagma preposicional como modificador de um sintagma nominal ou como um complemento de um verbo. Os movimentos oculares dos participantes foram monitorados durante a tarefa. Os resultados sugerem que a informação visual é utilizada pelos ouvintes durante o processamento incremental da linguagem.

**Palavras-chave:** Compreensão da língua falada; Ambiguidade; Efeitos de contexto; Referência.

**Abstract:** The purpose of this paper is to analyze how visual information affects the establishing of temporary ambiguous referents in Brazilian Portuguese (PB). Based on the approach presented by Tanenhaus, Spivey, Eberhard & Sedivy (1995) who analyzed the influence of visual context in the processing of temporary ambiguous sentences. The study of eye movement patterns allowed Tanenhaus et al. (1995) to conclude that visual information available during the earliest moments of language processing facilitates the identification of referents. These results seemed to challenge the theoretical framework predicted by serial processing models (Frazier & Fodor, 1978; Frazier, 1979), which states that only grammatical

---

· Professora Doutora da Faculdade de Letras e do Programa de Pós-Graduação em Estudos Linguísticos da Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brasil. E-mail: marialuiza.cunhalima@gmail.com.

·Doutoranda do Programa de Pós-graduação em Estudos Linguísticos da Universidade Federal de Minas Gerais (UFMG). Belo Horizonte, Minas Gerais, Brasil. E-mail: claubrandaov@gmail.com.

information would be used by the parser during sentence processing. In order to evaluate the impact of visual information on the resolution of temporary ambiguities in PB, we conducted an eye-tracker experiment which explored the visual world paradigm (Cooper, 1974). During this experiment, participants were asked to move objects around in different displays. Some of the auditory stimuli were temporarily ambiguous since the prepositional phrase could be interpreted as a noun phrase modifier or a verb complement. Participants had their eye movements monitored during the task. Results show that visual information is used by listeners during incremental language processing.

**Keywords:** Reference, Spoken language comprehension; Ambiguity; Context effects.

## Introdução

A identificação de referentes em um contexto discursivo envolve julgamentos pragmáticos sobre quais entidades disponíveis no mundo real e na memória dos interlocutores adaptam-se tanto à descrição do sintagma nominal quanto às restrições sintáticas que organizam essas descrições. Atualmente, inúmeros trabalhos têm analisado os mecanismos de escolha das expressões referenciais adequadas a certos contextos sintáticos e discursivos. O extenso debate sobre as condições de emprego de expressões nominais *versus* pronomes (Cf. GELORMINI-LEZAMA, 2012) é um claro exemplo disso.

No presente trabalho, voltamo-nos para a interação entre a construção local da representação sintática e a utilização de informações contextuais durante a compreensão de referências. Para tanto, analisamos como a informação visual pode afetar a identificação de referentes em português brasileiro (PB). O experimento aqui detalhado é uma adaptação para o PB do trabalho de Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy (1995) e de Spivey-Knowlton *et al.* (2002)<sup>1</sup> a respeito da influência do contexto visual no processamento de frases temporariamente ambíguas do inglês. Esses autores comprovaram, por meio do estudo de padrões de movimento ocular, que a informação disponível visualmente atua no início do processamento da língua falada, facilitando a identificação de referentes.

Um dos principais objetivos dos estudos sobre processamento da linguagem é conhecer a maneira como os seres humanos relacionam as palavras em uma frase e, com isso, compreendem algo que foi dito. O mecanismo humano de processamento de frases - chamado de *parser* - é responsável por identificar as relações existentes entre palavras de uma

---

<sup>1</sup> Embora trabalhos mais recentes discutam a influência do contexto visual no processamento linguístico (Cf. HUETTIG *et al.*, 2011), neste trabalho, trataremos especificamente de Tanenhaus *et al.* (1995) e de Spivey-Knowlton *et al.* (2002).

sentença. Esse mecanismo determina como as palavras relacionam-se umas com as outras, permitindo entendimento de diferentes proposições.

Existe um amplo debate sobre qual tipo de informação o *parser* utiliza primordialmente durante o processamento. De modo geral, os modelos sobre processamento de frases podem ser divididos em dois diferentes grupos. Em um primeiro grupo, encontram-se os modelos de processamento serial regidos inicialmente por informações estritamente sintáticas (FRAZIER & FODOR, 1978; FRAZIER & RAYNER, 1982; FERREIRA & CLIFTON, 1986; FRAZIER & CLIFTON, 1996). Esses trabalhos propõem que preferências na interpretação de frases ambíguas podem ser explicadas por meio de um processador que utiliza, a princípio, uma única possibilidade de organização estrutural e informações exclusivamente sintáticas. Com isso, informações do nível pragmático só entrariam na computação das relações sintáticas em um momento posterior, de revisão e integração.

Em um segundo grupo, encontram-se os modelos de processamento paralelo que podem ser regidos inicialmente por informações de diferentes origens, não apenas por informações sintáticas (ALTMANN & STEEDMAN, 1988; TRUESWELL *et al.* 1993; MACDONALD, PEARLMUTTER & SEIDENBERG, 1994; SPIVEY-KNOWLTON & SEDIVY, 1995; SPIVEY-KNOWLTON *et al.*, 2002). Para esses autores, o *parser* considera simultaneamente diferentes formas de se organizar estruturalmente as frases ambíguas até que informações de múltiplas origens (e.g., sintáticas, lexicais, contextuais) interajam entre si, restringindo certas alternativas de interpretação de frases.

Traxler & Tooley (2007, p.60) afirmam que o processamento de sentenças responde a dois grandes princípios: iminência e processamento incremental. Ou seja, uma determinada cadeia de palavras será compreendida tanto quanto for possível e assim que possível. Tal afirmação está relacionada ao fato de que os ouvintes interpretam rapidamente palavras a que têm acesso, antes mesmo de ouvirem uma sentença completa.

Como a compreensão da língua falada ocorre incrementalmente, durante a articulação de cada palavra, ouvintes lidam com diferentes níveis de ambiguidade até que se estabeleça um significado completo. A ambiguidade temporária é inerente à língua falada, pois as sequências de palavras reorganizam-se a cada momento em mais de uma forma. No entanto, ao final da articulação, é possível que o significado completo de uma mensagem torne-se único, sem qualquer ambiguidade. Por exemplo, ao ouvir o fragmento de frase como:

(1) Coloque a colher [sobre a xícara]...

Na maior parte das vezes, um nativo do PB compreenderá momentaneamente o sintagma preposicional (SP) *sobre a xícara* como um locativo que completa o verbo *colocar*, que expressa o lugar para onde a colher deve ser movida. No entanto, também há possibilidade de o falante compreender o sintagma *sobre a xícara* como um modificador do sintagma nominal (SN) definido *a colher*. Em geral, este segundo tipo de interpretação é direcionado pela existência de informações contextuais que influenciam a interpretação. Assim, durante a interpretação incremental, o fragmento de frase pode ser compreendido das seguintes maneiras:

- (2) Coloque a colher (no lugar) sobre a xícara...
- (3) Coloque a colher que está sobre a xícara...

Porém, ao ouvir a frase completa:

- (4) Coloque a colher sobre a xícara no pires.

No exemplo (4), após a articulação do SP *no pires*, surgem outras possíveis interpretações, pois a sentença é estruturalmente ambígua. Uma das possíveis interpretações preserva o SP *sobre a xícara* com o sentido locativo. Neste caso, o SP *no pires* poder ser interpretado como um modificador do SN definido *a xícara*, e o sentido da frase será:

- (5) Coloque a colher sobre [a xícara que está no pires]

Outra possibilidade mantém o caráter modificador do SP *sobre a xícara* e transforma *no pires* em um argumento do verbo *colocar*. Neste caso, o sentido da frase será:

- (6) Coloque [no pires] [a colher que está sobre a xícara]

Neste trabalho, tratamos especificamente das ambiguidades temporárias geradas pela articulação incremental de palavras. Investigamos se, ao ouvir fragmentos de frases como (1),

a ocorrência de um contexto visual no qual há dois candidatos a referentes (e.g.: um contexto em que há duas colheres) é capaz de influenciar a maior parte dos ouvintes a interpretar imediatamente o SP *sobre a xícara* como um modificador sintagma nominal definido *a colher*. Em outras palavras, gostaríamos de saber se a informação visual, estando disponível, influencia os estágios iniciais do processamento.

Segundo Crain e Steedman (1985, p.360), muitas das ambiguidades estruturais mais comuns envolvem a opção por uma estrutura sintática na qual um sintagma ambíguo pode tanto modificar um sintagma nominal definido quanto completar o sentido de um verbo. Para os autores, as preferências na interpretação de sintagmas ambíguos são originadas por fatores discursivos não locais que interagem com pressuposições linguísticas. Tendo em vista que um sintagma nominal definido pressupõe a existência de uma única entidade, é possível que a interpretação de um sintagma nominal modificado seja facilitada pela existência de outros possíveis candidatos a referentes.

Quando o contexto referencial visual possui mais de um candidato a referente do sintagma nominal definido (e.g., há no campo visual uma colher sobre uma xícara e uma colher sobre um guardanapo), a interpretação de que o SP exerce a função de modificador pode ser preferida pelos falantes, pois existe uma motivação para se diferenciar os possíveis referentes. Todavia, quando o contexto referencial visual não requer modificações para que se estabeleça a unicidade do sintagma nominal definido, os falantes assumirão que o SP modificador é redundante e, por isso, a interpretação do SP em função de complemento do verbo pode ocorrer mais frequentemente.

O debate sobre o tipo de informação utilizada nos estágios iniciais da compreensão de sentenças motivou Tanenhaus *et al.* (1995) e Spivey-Knowlton *et al.* (2002) a criarem o experimento adaptado nesta pesquisa. Para esses autores, grande parte das evidências sobre o uso privilegiado de informações sintáticas pelo *parser* foi obtida em estudos baseados em paradigmas de leitura. Esses estudos investigavam se um contexto referencial estabelecido pelo discurso poderia influenciar certas preferências na interpretação sintática. Em geral, o contexto discursivo era introduzido em um pequeno parágrafo e os movimentos oculares durante a leitura de sentenças que continham ambiguidades temporárias eram examinados. Os resultados nunca foram conclusivos, alguns estudos mostravam a influência do contexto discursivo e outros não. Certas preferências sintáticas se mostravam mais fortes do que outras, como é o caso dos sintagmas preposicionais interpretados como complementos locativos em verbos com esse tipo de exigência (e.g., pôr, colocar).

Tanenhaus *et al.* (1995) e Spivey-Knowlton *et al.*(2002) questionavam se a falta de resultados conclusivos estava relacionada ao fato de o contexto discursivo ser introduzido por um texto escrito. Por ser representado na memória, o contexto discursivo poderia não estar imediatamente acessível durante a leitura da sentença temporariamente ambígua. Além disso, no momento em que a região ambígua da sentença era lida, os participantes poderiam considerar que o contexto discursivo era irrelevante. Tais questionamentos foram primordiais para a realização de um experimento psicolinguístico que manipulasse um relevante contexto visual.

Com base na proposta de Crain e Steedman (1985), os pesquisadores conduziram um estudo em que o contexto referencial visual poderia influenciar os participantes a compreender um sintagma ambíguo como um modificador do sintagma nominal definido.

(7) Put the apple on the towel in the box<sup>2</sup>

No exemplo (7), o sintagma *on the towel* é temporariamente ambíguo, pois pode ser interpretado como um modificador do SN definido *the apple* ou como um complemento do verbo *put*. Para verificar as preferências dos participantes na resolução desse tipo de ambiguidade, o experimento explorava o método do *paradigma do mundo visual*. Esse método utiliza mecanismos de rastreamento ocular para mostrar em quais objetos os participantes estão olhando ao ouvir uma sentença falada.

Durante o experimento, objetos reais eram posicionados em uma mesa e, logo em seguida, os participantes ouviam instruções para realizar tarefas de movimentação dos objetos. Os pesquisadores manipulavam as características dos cenários para verificar o efeito de diferentes tipos de contextos referenciais sobre a interpretação de sentenças temporariamente ambíguas. Metade dos cenários possuía apenas um possível referente. Os cenários restantes possuíam dois referentes concorrentes.

Os resultados apresentaram diferentes parâmetros de fixação nos dois tipos de cenários. O sintagma ambíguo *on the towel* era rapidamente interpretado como um complemento locativo quando havia apenas um referente visual. Porém, diante da ocorrência de referentes visuais concorrentes, o sintagma ambíguo era geralmente interpretado como um modificador.

---

<sup>2</sup> Ponha a maçã sobre a toalha na caixa (Tradução nossa)

Em suma, os resultados demonstraram que o contexto referencial pode influenciar rapidamente a forma como um ouvinte interpreta um sintagma ambíguo. Assim, foi encontrada uma forte evidência contra os modelos seriais de *parser*, nos quais as decisões sintáticas iniciais são guiadas somente por informações estritamente sintáticas (FRAZIER, 1979; FRAZIER & CLIFTON, 1996), contribuindo para o arcabouço teórico de modelos de processamento baseados em restrições (MACDONALD, PEARLMUTTER, & SEIDENBERG, 1994; SPIVEY-KNOWLTON & SEDIVY, 1995; EBERHARD SEDIVY, 1995, TRUESWELL et al., 1993) que defendem a interação entre contexto e processamento sintático.

Em nossa adaptação do estudo conduzido por Tanenhaus *et al.* (1995) e Spivey-Knowlton *et al.* (2002), utilizamos a técnica do rastreamento ocular para compreender a maneira como falantes do português brasileiro compreendem frases ambíguas, quando estão diante de um contexto visual que privilegie interpretações menos usuais. Exploramos o paradigma do mundo visual para conhecer o curso temporal da inclusão de informações visualmente disponíveis na resolução das estruturas sintáticas ambíguas e consequente seleção de referentes.

## **Experimento**

Irwin (2004, p.105) afirma que os movimentos oculares aparentam ser uma variável dependente ideal para estudos de processos cognitivos, pois remetem a um comportamento naturalmente humano: pessoas tendem a olhar para coisas sobre as quais desejam obter mais informações. O uso de técnicas de rastreamento ocular permite a obtenção de pistas cada vez mais refinadas sobre a maneira como se desenvolve o processo de compreensão.

Na década de 1970, Cooper (1974) mostrou que, diante de um cenário referencial relevante para a compreensão, ouvintes realizam movimentos oculares que são praticamente sincronizados com instruções faladas. Vinte anos depois, Tanenhaus *et al.*(1995) utilizaram uma metodologia similar a de Cooper (1974), dando início a uma série de estudos psicolinguísticos que exploravam as relações sistemáticas entre contexto visual, movimentos oculares e língua falada. Esse paradigma é atualmente conhecido como *paradigma do mundo visual*.

Os experimentos de compreensão de linguagem que utilizam o *paradigma do mundo visual* são definidos pelas seguintes características (HUETTIG *et al.*, 2011):

- a. Um grupo de participante ouve uma instrução falada;
- b. Durante a instrução eles também veem um cenário (ou contexto referencial visual) relevante para a interpretação;
- c. Os movimentos oculares são gravados por um rastreador ocular para uma análise posterior. Neste tipo de experimento, cada fixação reflete a interpretação incremental que o sujeito faz da frase ao longo do tempo.

Ao ouvir o início de uma palavra alvo, um ouvinte demora aproximadamente 100 a 150 milésimos de segundo para programar um movimento e começar a mover o olho, e, adicionalmente, 50 milésimos de segundo para que a informação visual atinja ao córtex cerebral, após do início da fixação (TRAXLER, 2012, p.373). Em movimentos oculares, a atuação de informações relevantes para identificação de referentes pode ser observada extremamente cedo. Sedivy (2003) encontrou evidências desse fenômeno apenas 200 milésimos de segundo após o início de uma palavra alvo. Tais medidas são relevantes para a análise a ser apresentada na seção *Resultados*, pois verificamos qual objeto referente foi fixado 250<sup>3</sup> milésimos de segundo depois do início da palavra alvo. Dentre os modelos que explicam as preferências na interpretação de sintagmas preposicionais como complementos verbais ou modificadores nominais (ALTMANN & STEEDMAN, 1988; FERREIRA E CLIFTON, 1986), a proposta de Crain e Steedman (1985) foi bastante relevante para o desenho de nosso experimento. Sob a perspectiva destes autores, quando o contexto visual expõe mais de um possível referente, os ouvintes tendem a analisar os sintagmas preposicionais como modificadores de nomes, dado que a modificação é necessária para se estabelecer o referente único. O experimento visa, então, investigar se a mudança na informação visual apresentada influencia como os participantes interpretam o que ouvem. Para isso, criamos diferentes contextos visuais e diferentes instruções verbais que passaremos a detalhar abaixo.

---

<sup>3</sup> Por razões que serão detalhadas na seção *Resultados*, optamos por um parâmetro conservador em relação ao tempo mínimo para a janela de fixação.



## **Materiais e métodos**

O experimento consistiu na apresentação de instruções para mover objetos em um contexto referencial (cenário) relevante. O tipo de instrução apresentada aos participantes foi utilizado como variável independente. As instruções alternavam entre uma sentença com ambiguidade temporária (Cf. 7) e uma sentença sem ambiguidade temporária (Cf. 8) <sup>4</sup>.

(7) Coloque a colher sobre a xícara no pires

(8) Coloque a colher que está sobre a xícara no pires

As frases sem ambiguidade temporária foram utilizadas como uma forma de controle, já que não esperávamos que os participantes hesitassem ao identificar o objeto citado. Ao contrário das sentenças com ambiguidades temporárias, que poderiam causar hesitações nos participantes durante a interpretação inicial da identidade do referente, a depender dos cenários com que se deparavam.

Os cenários do experimento eram compostos por objetos de brinquedo situados sobre uma mesa. A fim de facilitar a posterior análise dos dados, a mesa foi dividida em quatro quadrantes (áreas de interesse) nos quais foram posicionados os diferentes objetos referentes. Os tipos de contextos referenciais também foram utilizados como variável independente. Vinte e quatro variações de cenários que alternavam entre dois tipos de contextos referenciais foram criadas para as passagens experimentais. Esses contextos continham:

I. Contexto de 1 referente principal: um referente principal (e.g., colher posicionada sobre uma xícara), um falso locativo (e.g., xícara sozinha), um objeto distrator (e.g., ovo) e um locativo final (e.g., pires). (Cf. Figura 1)

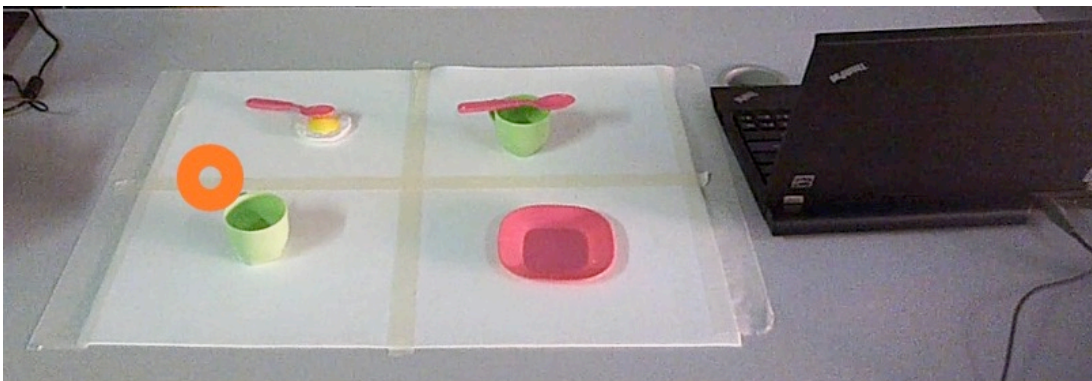
---

<sup>4</sup> É importante notar que as instruções apresentadas durante o experimento eram previamente gravadas. Ou seja, não havia para os diferentes tipos contextos visuais (contextos de 1 referente; contexto de 2 referentes) variação prosódica nas instruções ambíguas e não ambíguas, não podendo, portanto, a prosódia ser considerada um fator para influenciar o tipo de interpretação.



**Figura 1:** Contexto de 1 referente principal

II. Contexto de 2 referentes concorrentes: um referente principal (e.g., colher posicionada sobre uma xícara), um falso locativo (e.g., outra xícara), um falso referente (e.g., colher posicionada sobre um ovo) e um locativo final (e.g., pires). (Cf. Figura 2)



**Figura 2:** Contexto de 2 referentes concorrentes

Se o contraste entre múltiplos referentes fosse relevante para a interpretação do sintagma *sobre a xícara* como um modificador do sintagma nominal *a colher*, esperávamos que no contexto de 1 referente principal houvesse mais fixações no falso locativo (e.g.: xícara sozinha). No contexto de 2 referentes concorrentes, esperávamos que houvesse mais fixações no referente principal (e.g., colher posicionada sobre uma xícara).

Portanto, todos os itens experimentais eram apresentados em quatro diferentes condições:

<b>Condição</b>	<b>Descrição</b>
<b>Condição 1</b>	A instrução é temporariamente ambígua. O referente principal (e.g., colher posicionada sobre uma xícara) é apresentado em um cenário onde há apenas um possível referente (e.g.: há apenas uma colher no cenário).
<b>Condição 2</b>	A instrução é temporariamente ambígua. O referente principal (e.g., colher posicionada sobre uma xícara) é apresentado em um cenário onde há dois possíveis referentes (e.g.: há duas colheres no cenário).
<b>Condição 3</b>	A instrução não é temporariamente ambígua. O referente principal (e.g., colher posicionada sobre uma xícara) é apresentado em um cenário onde há apenas um possível referente (e.g.: há apenas uma colher no cenário).
<b>Condição 4</b>	A instrução não é temporariamente ambígua. O referente principal (e.g., colher posicionada sobre uma xícara) é apresentado em um cenário onde há dois possíveis referentes (e.g.: há duas colheres no cenário).

**Tabela 1:** Descrição das condições utilizadas no experimento

Alguns controles específicos foram criados para evitar qualquer tipo de viés visual ou acústico. O tamanho dos objetos que compunham cada cenário era semelhante, prevenindo que a saliência visual de alguns objetos alterasse os resultados. Antes da versão final do experimento, um pré-teste foi aplicado para testar a prototipicidade dos objetos, isto é, se os objetos eram de fato reconhecidos pelos participantes como sendo representações prototípicas dos nomes citados. Esse pré-teste consistia numa tarefa de nomeação, em que os participantes deveriam dizer o nome de alguns objetos mostrados pelo pesquisador. Todos os objetos foram unanimemente identificados com os nomes esperados.

Para evitar que a saliência visual influenciasse as primeiras fixações, os participantes tiveram um curto tempo para se familiarizar e reconhecer os objetos de cada cenário, antes das instruções orais começarem. A cada cenário, a posição do *referente principal*, do *falso locativo* e do *locativo final* era modificada, evitando qualquer padrão na organização dos objetos. Todas as instruções foram gravadas previamente para garantir a qualidade do sinal acústico e a uniformidade da prosódia utilizada. Além disso, as mesmas gravações foram utilizadas nos diferentes contextos, excluindo qualquer viés devido à duração ou à prosódia da fala. Os estímulos experimentais e distratores foram apresentados aleatoriamente e os áudios do experimento foram organizados no programa DMDX.

Cada participante ouvia 12 sentenças (instruções) experimentais e visualizava 12 cenários (contextos) experimentais. Ou seja, os participantes eram expostos três vezes a cada uma das condições detalhadas acima. As sentenças experimentais eram apresentadas em

blocos com um item experimental e dois itens distratores. Apesar dos blocos serem apresentados aleatoriamente, em cada bloco, a sentença experimental era sempre a primeira a ser articulada em cada um deles. Como as instruções demandavam a movimentação de objetos, a ordem em que as instruções eram apresentadas dentro de cada bloco foi controlada para evitar que o cenário tivesse sido modificado antes que a sentença experimental fosse apresentada. Ao fim de cada bloco, o cenário era completamente modificado.

As sentenças experimentais possuíam as seguintes estruturas:

Sentença com ambiguidade temporária					
Ponha a	colher	sobre a	xícara	no	pires
	<b>Primeiro referente</b>		<b>Palavra alvo</b>		<b>Locativo</b>
Sentença sem ambiguidade temporária					
Ponha a	colher	que está sobre a	xícara	no	pires
	<b>Primeiro referente</b>		<b>Palavra alvo</b>		<b>Locativo</b>

**Quadro 1:** Estrutura utilizada nas sentenças experimentais

Controlamos o número de sílabas nas regiões do *primeiro referente*, da *palavra alvo* e do *locativo*. O *primeiro referente* possuía duas sílabas; a *palavra alvo* possuía três sílabas; o *locativo* possuía duas sílabas. Em nossa análise, a *palavra alvo* constitui o ponto em que falante inicia a resolução da ambiguidade temporária. A partir do início dessa palavra, utilizamos um intervalo de 250ms para verificar qual objeto atraía mais olhares dos participantes: o referente principal (e.g., colher posicionada sobre xícara) ou o falso locativo (e.g., xícara sozinha).

Adicionalmente, durante o experimento, os participantes também eram expostos a 12 blocos com três instruções distratoras e seus respectivos cenários, totalizando 60 itens distratores no experimento completo (24 instruções eram articuladas em grupos com itens experimentais; 36 instruções eram articuladas em grupos com apenas itens distratores). Cerca de 12% das distratoras possuíam sintagmas nominais modificados, porém não diretamente relacionados às manipulações realizadas nas sentenças experimentais.

## Participantes

8 alunos (6 do sexo feminino e 2 do sexo masculino) da Universidade Federal de Minas Gerais participaram voluntariamente do experimento em troca de créditos acadêmicos. Todos os participantes da pesquisa eram adultos, falantes nativos de PB, apresentando visão normal e tinham de 21 a 33 anos de idade (idade média de 24 anos).

## **Procedimentos**

Os áudios do experimento foram rodados em um computador pessoal por meio do programa DMDX. Simultaneamente, acoplado a um notebook, usamos os óculos para rastreamento ocular SMI I VIEW ETG, binocular, com precisão temporal de 30 Hz e precisão espacial de  $0.1^{\circ 5}$ .

Inicialmente, o experimentador informava a cada participante da pesquisa que, após ouvir algumas instruções gravadas, os participantes deveriam movimentar alguns objetos de brinquedo em um cenário demarcado (Cf. Figuras 1 e 2). Era solicitado que a tarefa fosse realizada no ritmo mais natural possível, e que os participantes seguissem sua intuição em relação ao que deveria ser feito.

Depois de informadas as instruções e de solucionadas as eventuais dúvidas do participante, iniciava-se a sessão de calibragem do equipamento, na qual o sujeito da pesquisa deveria fixar o olhar em um ponto fixo posicionado a 1,50m de distância. Tal procedimento se faz necessário para garantir que o rastreador encontre o olho do participante, e também para que seja capaz de acompanhar com precisão o trajeto do olhar no espaço em que estavam localizados os objetos.

Os objetos de cada cenário eram posicionados sobre o espaço predeterminado com quadrantes demarcados, como ilustrado nas figuras 1 e 2. Os participantes observavam essa organização, ganhando alguns segundos para reconhecer os objetos. Após cada instrução (sentença), experimental ou distratora, os participantes tinham o tempo necessário para movimentar os objetos (cerca de 1 minuto). Terminada a tarefa, o experimentado iniciava a execução do próximo estímulo. Um total de 24 áudios com instruções triplas eram tocados diante de 24 diferentes cenários com objetos. As passagens críticas sempre eram iniciadas com uma instrução experimental, ambígua ou não ambígua, como as exemplificadas em (7) e (8). Foram criadas quatro listas nas quais a apresentação das doze instruções experimentais alternava entre as quatro condições exemplificadas na tabela 1. Os participantes ouviam pelo

---

<sup>5</sup> Para informações detalhadas sobre o aparelho acesse <http://www.smivision.com/>

menos uma forma (ambígua ou não ambígua) de cada instrução experimental. A duração de uma seção experimental completa era de aproximadamente vinte e cinco minutos, variando conforme o ritmo do participante. Abaixo há um exemplo de um grupo de 3 instruções<sup>6</sup>:

(9) Coloque a colher sobre a xícara no pires (Instrução experimental)

Coloque o ovo próximo da colher no pires (Instrução distratora)

Coloque as xícaras onde estava o ovo (Instrução distratora)

## Resultados

Por meio do *software* Be Gaze (versão 3.2.28), os dados de rastreamento ocular foram analisados, quadro a quadro, a cada 5ms. Os dados de rastreamento ocular, áudio e vídeo dos experimentos também foram sincronizados neste programa. Analisamos apenas os dados de participantes que realizaram as tarefas corretamente, de acordo com a interpretação sustentada pelo contexto visual. Em um total de 96 dados experimentais, 14 foram eliminados por erros de interpretação e 1 por problemas no vídeo (totalizando 15,62% dos dados). Nas instruções experimentais, os inícios do *primeiro referente* (e.g., colher), da *palavra alvo* (e.g., xícara) e do *locativo* (e.g., pires) foram marcados com ajuda do *software* Praat (versão 5.3.39). A partir do início do *primeiro referente*, registramos os períodos de fixações em cada objeto, começando no momento em que o olhar pousava na área de interesse demarcada no cenário. Nossa análise baseou-se no objeto referente fixado após a articulação da *palavra alvo* (e.g.: xícara). A escolha de tal trecho deve-se ao fato de que a *palavra alvo* contém a principal informação para a interpretação incremental. Após a articulação da *palavra alvo*, os participantes precisavam optar por uma interpretação que considerasse o sintagma preposicional um complemento do verbo ou um adjunto do sintagma nominal.

Ao analisar as fixações, definimos um intervalo de 250ms para o processamento da referência, a partir do início da *palavra alvo*. Conforme citado na seção *Experimento*, um ouvinte demora aproximadamente 200ms para iniciar os movimentos de uma sacada. Além disso, Sedivy (2003, p.17) comprovou que a identificação de referentes temporariamente ambíguos ocorre em aproximadamente 200ms a partir do início da palavra desambiguadora.

---

<sup>6</sup> A lista completa das instruções experimentais encontra-se no Anexo 1.

Como os tempos de fixações foram registrados a partir do momento em que os participantes pousavam os olhos nas áreas de interesse, optamos por uma análise conservadora, na qual acrescentamos 50ms para compensar o início da sacada. Consideramos que a interpretação feita incrementalmente impactava nos objetos fixados. Portanto, o objeto fixado 250ms após início da *palavra alvo* refletia a maneira como os participantes estavam interpretando o sintagma preposicional.

A variável dependente analisada foi a ocorrência de fixações no *referente principal*, no *falso locativo* e no *locativo final* 250ms após o início da *palavra alvo*<sup>7</sup>. Isto é, se a interpretação inicial dos sujeitos foi a de identificar o sintagma preposicional como modificador (fixações no *referente principal*) ou como locativo (fixações no *falso locativo*). Como a análise partiu de uma classificação binária dos dados (presença ou ausência de fixação na janela de interesse), utilizamos o teste de chi-quadrado<sup>8</sup> para a análise estatística.

Os dados sugerem um efeito geral do contexto referencial visual. Em uma comparação entre o número de fixações no *referente principal* (e.g.: colher sobre uma xícara) e o número de fixações no *falso locativo* (e.g.: xícara sozinha), todas as condições favoreceram a fixação no *referente principal* [condição 1 referente em frase ambígua - ( $\chi^2 = 11.5$ ,  $df=4$ ,  $p < 0.05$ ); condição 2 referentes em frase ambígua ( $\chi^2 = 18.1111$ ,  $df=4$ ,  $p < 0.001$ ); condição 1 referente em frase não ambígua ( $\chi^2 = 35.5833$ ,  $df=4$ ,  $p < 0.0001$ ); condição 2 referentes em frase não ambígua ( $\chi^2 = 26.3478$ ,  $df=4$ ,  $p < 0.0001$ )]. O gráfico 1 mostra as proporções de fixações por condição aos 250ms<sup>9</sup>.

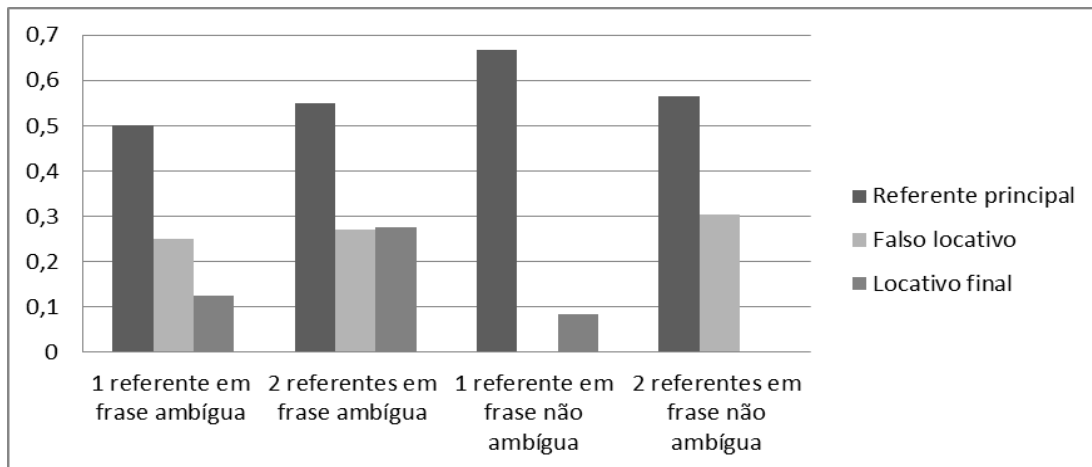
Ao comparar os dados das quatro condições entre si, nenhuma condição apresentou um comportamento significativamente diferente, facilitando a identificação do *referente principal* ( $\chi^2 = 2.194$ ,  $df=3$ ,  $p > 0.05$ ). Lembremos que o referente principal é, por exemplo, a colher a ser movida, sendo que algumas vezes o cenário oferecia dois objetos concorrentes que correspondiam à palavra *colher*. O fato de os sujeitos majoritariamente ignorarem a colher alternativa nas condições em que ela estava presente, comportando-se de forma semelhante a quando ela está ausente, é forte indicativo de que o cenário favoreceu, em certa medida, uma leitura do sintagma preposicional *sobre a xícara* como um modificador do sintagma nominal *colher*.

---

<sup>7</sup> Objetos fixados aos 250ms só eram considerados na análise se assim permanecessem por no mínimo 200ms. Quando nenhum objeto era fixado aos 250 ms, o primeiro objeto fixado após tal período era considerado.

<sup>8</sup> O teste chi-quadrado ( $\chi^2$ ) de Pearson é um teste estatístico criado para avaliar se há associação entre duas variáveis categóricas. Esse teste estatístico é baseado na comparação entre a frequência que uma categoria é observada e a frequência que essa categoria esperada para certo evento. (Cf. FIELD et al., 2012).

<sup>9</sup> Apesar de a análise ser categorial, o gráfico exhibe porcentagens para maior facilidade de visualização dos resultados.



**Gráfico 1:** Proporções de fixações aos 250ms.

A comparação entre as condições 1 referente em frase ambígua e 2 referentes em frase ambígua mostrou tendências parecidas de fixações. Os dados mostram ainda que não houve preferência pelo *falso locativo* em nenhuma condição - ( $\chi^2 = 1.9032$ ,  $df=3$ ,  $p > 0.05$ ). Segundo Craig em Steedman (1985), a condição com apenas um possível referente deveria favorecer a fixação no falso locativo. No entanto, tal favorecimento não se verificou.

## Discussão

Nossos dados indicam uma preferência pela estrutura modificadora durante a compreensão do sintagma preposicional, independentemente do número de possíveis referentes presentes no contexto visual. Portanto, os resultados do presente experimento diferem dos resultados de Tanenhaus *et al.* (1995) e Spivey-Knowlton *et al.* (2002), nos quais a preferência pela estrutura modificadora se manifestava apenas em contextos com dois possíveis referentes. Uma possível explicação para tal diferença seria a preferência em português brasileiro pela leitura modificadora de sintagmas prepositivos. Porém, acreditamos que essa hipótese é pouco provável, tendo em vista a natureza locativa das preposições e das locuções prepositivas utilizadas em nossos itens experimentais (Cf. Anexo).

Acreditamos que as diferenças entre os nossos resultados e os resultados de Tanenhaus *et al.* (1995) e Spivey-Knowlton *et al.* (2002) devem-se principalmente à metodologia de



análise dos padrões de movimento ocular. Como o experimento original foi feito há quase vinte anos, é possível que, na época, a capacidade dos recursos tecnológicos para análise de dados era bastante limitada. Já está comprovado na literatura psicolinguística que a resolução de referências ambíguas ocorre bastante cedo, logo após o início da palavra desambiguadora (Eberhard *et al.* 1995; Sedivy, 2003). Nos resultados relatados por Spivey-Knowlton *et al.* (2002), em condições que favoreciam a interpretação do sintagma preposicional como modificador do nome (sentença ambígua e cenário com 2 referentes), foi encontrado um intervalo médio de 500ms entre o fim da *palavra alvo* (e.g., xícara) e a fixação no *referente principal* (e.g., colher posicionada sobre a xícara). Entretanto, em contextos semelhantes, encontramos um número significativo de fixações nos *referentes principais* em pouco mais de 200ms após o início das palavras alvos.

Outra questão com a qual lidamos em nosso experimento, mas que não está relatada nos artigos em que nos baseamos, diz respeito à compreensão incorreta das tarefas. Eliminamos os dados dos itens em que os participantes não se comportavam como o esperado, porém, no trabalho original, a forma como os pesquisadores lidaram com esse tipo de desvio não está especificada. Uma diferença notável entre os dois experimentos é o fato de a estrutura linguística dos itens experimentais em português ser completamente ambígua, o que pode ter ocasionado uma série de comportamentos anômalos durante o pós-processamento. Tal questão necessita de aprofundamento, porém constitui uma possibilidade para se explicar os resultados divergentes.

Acreditamos que nossos resultados sugerem uma influência bastante precoce do contexto, e, por isso, em todas as condições (frases ambíguas com um referente; frases ambíguas com dois referentes; frases não ambíguas com um ou dois referentes), o *referente principal* foi significativamente mais fixado do que o *falso locativo*. Ou seja, após o início das *palavras alvo* (e.g., xícara) os participantes de nosso experimento conseguiam identificar rapidamente e corretamente os *referentes principais* (e.g., colher posicionada sobre uma xícara). Assim, em certa medida, os resultados indicam que a informação visual sobre possíveis referentes foi utilizada de alguma forma pelos ouvintes durante o processamento incremental da linguagem.

Portanto, os resultados sugerem que a forma como ouvintes inicialmente interpretam sintagmas temporariamente ambíguos é influenciada de maneira geral pelo contexto referencial visual. Porém, diante da ausência de resultados que poderiam nos trazer evidências sobre as motivações para a integração dessas informações contextuais, lançamos para investigações futuras uma análise mais detalhada das fixações durante uma janela temporal

diferente, e, ainda, a aferição do tempo total de fixação em cada objeto utilizado nos itens experimentais.

## Referências

ALTMANN, Gerry. T. M.; STEEDMAN, Mark. *Interaction with context during human sentence processing*. Cognition, 30, 1988.

COOPER, Roger M. *The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing*. Cognitive Psychology, 6, 1974.

CRAIN, Stephen; STEEDMAN, Mark. On not being led up the garden path: The use of context by the psychological parser. In: Dowty, David; Karttunen, Lauri; Zwicky, Arnold. *Natural language parsing: Psychological, computational, and theoretical perspectives*. Cambridge: Cambridge University Press, 1985.

EBERHARD, Katheleen. M.; SPIVEY-KNOWLTON, Michael J.; SEDIVY, Julie C.; TANENHAUS, Michael K. *Eye movements as a window into real-time spoken language comprehension in natural contexts*. Journal of Psycholinguistic Research, 24, 1995.

FERREIRA, Fernanda F.; CLIFTON, Charles. *The independence of syntactic processing*. Journal of Memory and Language, v.25, 1986.

FIELD, Andy P.; MILES, Jeremy; FIELD, Zoë . *Discovering Statistics Using R*. London: Sage, 2012.

FRAZIER, Lyn; FODOR, Janet D. *The Sausage Machine: A new two-stage parsing model*. Cognition, v.6, 1978.

\_\_\_\_\_, Lyn. *On comprehending sentences: Syntactic parsing strategies*. Tese de Doutorado. University of Connecticut , 1979.

\_\_\_\_\_, Lyn; RAYNER, Keith. *Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences*. Cognitive Psychology, v.14, 1982.

\_\_\_\_\_, Lyn; CLIFTON, Charles. *Construal*. Cambridge: MIT Press, 1996.

GELORMINI-LEZAMA, Carlos. *A review of the repeated name penalty: implications for null subject languages*. Revista Lingüística. Volume 8, número 2, 2012.

HUETTIG, Falk; ROMMERS, Joost; MEYER, Antje S. *Using the visual world paradigm to study language processing: A review and critical evaluation*. Acta Psychologica, 137, 2011.

IRWIN, David E. Fixation location and fixation duration as indices of cognitive processing. In: HENDERSON, John M.; FERREIRA, Fernanda F. (Eds.), *The integration of language, vision, and action: Eye movements and the visual world*, New York: Psychology Press, 2004.

MACDONALD, Maryellen C.; PEARLMUTTER, Neal J.; SEIDENBERG, Mark S. *Lexical nature of syntactic ambiguity resolution*. Psychological Review, 101, 1994.

SEDIVY, Julie C. *Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations*. Journal of Psycholinguistic Research. v.32, nº1, 2003.

SPIVEY-KNOWLTON, Michael J.; SEDIVY, Julie C. *Resolving attachment ambiguities with multiple constraints*. *Cognition*, 55, 1995.

SPIVEY-KNOWLTON, Michael J.; TANENHAUS, Michael K.; EBERHARD, Katherine M.; SEDIVY, Julie C. *Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution*. *Cognitive Psychology*, 45, 2002.

TANENHAUS, Michael K.; SPIVEY-KNOWLTON, Michael J.; EBERHARD, Katherine M.; SEDIVY, Julie C. *Integration of visual and linguistic information in spoken language comprehension*. *Science*, 268, 1995.

TRAXLER, Matthew J.; TOOLEY, Kristen M. *Lexical mediation and context effects in parsing*. *Brain Research*, 1146, 2007.

\_\_\_\_\_, Matthew J. *Introduction to Psycholinguistics: Understanding Language Science*. Chichester: Wiley-Blackwell, 2012.

TRUESWELL, John C.; TANENHAUS, Michael K.; KELLO, Christopher. *Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1993.

Artigo recebido em: 29/08/2015

Artigo aceito em: 02/12/2015

Artigo publicado em: 28/12/2015

## **Anexo**

### **Lista das sentenças experimentais com ambiguidades temporárias**

1. Ponha a faca sobre a panela no fogão.
2. Coloque a nota junto da cadeira no sofá.
3. Ponha o milho perto do revólver na mesa.

4. Coloque a colher sobre a xícara no pires.
5. Ponha a chave junto do martelo na caixa.
6. Coloque a pera perto da banana no pote.
7. Ponha o garfo sobre a poltrona na cama.
8. Coloque o brinco junto do secador no balde.
9. Ponha o buquê perto da lanterna na jarra.
10. Coloque o pente sobre a sacola na bolsa.
11. Ponha a taça perto do esmalte na tábua.
12. Coloque o frasco junto do pimentão no prato.

**Lista das sentenças experimentais sem ambiguidades temporárias**

1. Ponha a faca que está sobre a panela no fogão.
2. Coloque a nota que está junto da cadeira no sofá.
3. Ponha o milho que está perto do revólver na mesa.
4. Coloque a colher que está sobre a xícara no pires.
5. Ponha a chave que está junto do martelo na caixa.
6. Coloque a pera que está perto da banana no pote.
7. Ponha o garfo que está sobre a poltrona na cama.
8. Coloque o brinco que está junto do secador no balde.
9. Ponha o buquê que está perto da lanterna na jarra.
10. Coloque o pente que está sobre a sacola na bolsa.
11. Ponha a taça que está perto do esmalte na tábua.
12. Coloque o frasco que está junto do pimentão no prato.