

DURANTE LEITURA, ESTÍMULOS VISUAIS, COMO A IMAGEM DO(A) AUTOR(A), PODEM PREJUDICAR O DESEMPENHO DE LEITORES**DURING READING, VISUAL STIMULI, AS THE IMAGE OF THE AUTHOR, MAY IMPAIR IN THE PERFORMANCE OF READERS**Wagner Ferreira Lima¹

RESUMO: Durante a leitura, estímulos visuais, como a imagem fotográfica do(a) respectivo(a) autor(a), podem influenciar a compreensão textual de leitores juvenis? Essa questão motivou um estudo sobre significados outros, presentes na comunicação, além do significado do falante. Assim, um grupo de alunos de 9. ano escolar (n = 26) foi solicitado a ler textos em dois níveis de multimodalidade (unimodal (só texto) vs bimodal (texto + foto)); e a executar testes de inferências específicos (escolha de alternativas) acerca do assunto lido (testes de inferência). No final, cada participante informou, em uma escala ordinal (1 a 10), seu nível de certeza quanto à escolha feita por ele (testes de avaliação). Os dados encontrados apontam para a existência de efeito de bimodalidade sobre as inferências acerca dos textos (p = 0.000). E isso foi reforçado pelos resultados dos testes de avaliação. Contudo, uma análise mais minuciosa, considerando o tema dos textos como variável independente (texto *a* “As redes sociais” vs texto *b* “Gravidez na adolescência”), mostrou que o efeito observado foi de interferência e não de facilitação, como era esperado. Isso foi evidenciado especialmente na situação de leitura do texto *a* (p = 0.01). Em suma, entre leitores juvenis, as fotos não funcionaram como pistas extraverbais, nem foram relevantes para a compreensão leitora. Portanto, conclui-se que a expressividade do(a) autor(a) contribui exclusivamente com a recepção da mensagem, e não com a compreensão do significado do texto.

PALAVRAS-CHAVE: Educação. Compreensão leitora. Multimodalidade.

ABSTRACT: During reading, can visual stimuli, such as the photographic image of the respective author, influence the textual understanding of young readers? This question motivated a study on other meanings, present in communication, in addition to the speaker's meaning. Thus, a group of students in the 9. school year (n = 26) were asked to read texts in two levels of multimodality (unimodal (text only) vs bimodal (text + photo)); and to perform specific inferences tests (choice of alternatives) on the subject read (inference tests). In the end, each participant reported, on an ordinal scale (1 to 10), their level of certainty regarding their choice (assessment tests). The data found point to the existence of a bimodality effect on the inferences about the texts (p = 0.000). This was reinforced by the results of the evaluation tests. However, a more detailed analysis, considering the theme of the texts as an independent variable (text *a* “Social networks” vs text *b* “Teenage pregnancy”), showed that the observed effect was of interference and not of facilitation, as expected. This was evidenced especially in the situation of reading the text *a* (p = 0.01). In short, among juvenile readers, the photos did not function as extraverbal cues, nor were they relevant to reading comprehension. Therefore, it is concluded that the author's expressiveness contributes exclusively to the reception of the message, and not to the understanding of the meaning of the text.

KEYWORDS: Education. Reading comprehension. Multimodality.

1 Introdução

Gesticulações facial e corporal são componentes essenciais da comunicação face a face. De um lado, elas atuam como dicas extraverbais a indicarem como os interlocutores devem se

¹ Prof. Associado do Departamento de Letras Vernáculas e Clássicas, da Universidade Estadual de Londrina (UEL). E-mail: wflima@uel.br <<http://orcid.org/0000-0003-0505-4668>>

engajar na comunicação. Por exemplo, um orador que se apresenta com amistosidade e entusiasmo tem mais chances de cativar seu público do que aquele que se comunica com rispidez e austeridade. De outro, elas servem para o esclarecimento cognitivo da comunicação (como os gestos manuais) (KELLY; ÖZYÜREK; MARIS, 2010; WILLIAMS; HAGOORT, 2007). Isso pode ser o caso de alguém descrevendo um objeto redondo. Ele pode fazer gestos circulares com as mãos ao falar desse objeto. É por isso que na ausência de gesticulação manual a comunicação perde sua naturalidade (GARROD; PICKERING, 2009). Em vista disso, perguntamos se durante a leitura expressões faciais são capazes de exercer algum dos efeitos acima aludidos. Essa é uma questão pertinente, pois textos midiáticos, tais como propagandas, blogs, posts etc., são um misto de ortografia e imagens visuais. Textos multimodais são muito atraentes aos jovens, porque são altamente informativos e situados. Por isso, o tema da multimodalidade na leitura merece a atenção por parte dos educadores que lidam com adolescentes.

A fim de oferecer uma resposta a essa questão, realizamos uma pesquisa com um grupo de alunos solicitando a leitura de textos em duas condições, com a imagem visual dos autores e sem a imagem deles. Postulamos que, como na oralidade, tais imagens poderiam atuar como pistas psicossociais e influenciar a compreensão textual. Recentemente, Bouko (2020) mostrou como emoções são criadas em comunicação multimodal; mas essa pesquisa não dá conta de esclarecer como a compreensão pode ser aprimorada ou facilitada pelas atitudes retratadas em fotos pareadas com os textos (no nosso caso, retratos do(a) autor(a) do texto). Desse modo, o desafio nosso era prover evidências de que os retratos do(a) autor(a) poderiam facilitar a compreensão leitora ao estabelecerem uma (aguardada) interação afetiva com os leitores.

Encontramos que, diferente do esperado, o referido pareamento produziu efeitos de interferência, e não de facilitação sobre a leitura. Esses resultados requerem, assim, um entendimento de por que esse fenômeno teve lugar. Para a realização desse esclarecimento, empregamos dois construtos pouco explorados nos estudos da linguagem: a noção filosófica de autoexpressão e a noção cognitiva de imageria mental multimodal. A combinação deles providenciou a um instrumental teórico ótimo para lidarmos com o problema de comunicação multimodal.

Como vamos esclarecer na seção 2, o texto escrito contém uma dimensão expressiva que mostra o estado psicológico em que está seu autor durante a enunciação. Na seção 3, pretendemos elucidar como essa expressividade reflete aspectos da comunicação oral; e é capturada pelo leitor devido a processos de imageria mental multimodal. Descrevemos, nas seções 4, a plausibilidade do estudo que propomos; e em 5 e 6, a metodologia da pesquisa e os resultados obtidos, respectivamente. Análise e discussão dos dados serão apresentadas na seção 7, onde propomos uma explicação para os resultados encontrados. Considerações finais, seção 8, encerram a apresentação apontando os aspectos positivos e negativos desta pesquisa.

2 Expressividade do texto

Quando falamos em pistas afetivas na escrita, uma primeira questão que precisa ser esclarecida é quanto ao aspecto intersubjetivo do texto. Aqui assumimos que, mesmo desprovida das contingências da oralidade, a escrita contém uma significação interpessoal como a do diálogo. Essa significação ocorre no nível do que podemos chamar de expressividade do texto. Uma pessoa que escreve sobre um assunto qualquer está fazendo mais do que descrever como o mundo é. Ela está também expressando a si mesma. O filósofo da linguagem M. Green (2007) denomina esse fenômeno de autoexpressão (*self-expression*).

Consideremos o texto *Gravidez na adolescência* (Figura 2). A primeira impressão que temos dele é que se trata de um texto de opinião. Embora ignoremos sua autoria (o texto não é assinado), concluímos que ele foi escrito por uma pessoa relativamente jovem, consciente da realidade e preocupada com o bem-estar dos adolescentes. Além disso, em vista do tom ríspido

do discurso, compreendemos que o(a) autor(a) está indignado(a) com os atos impensados de alguns jovens. Tomadas juntas, essas informações permitem conhecimento acerca do(a) autor(a). Isso é assim porque, segundo Green (2008), o texto *mostra* o estado psicológico em que o enunciador está no momento da enunciação. “Eu sugiro que em autoexpressão nós designadamente mostramos o que está dentro”.² (GREEN, 2010, p. 67).

Vale lembrar que, para esse pensador, mostrar é uma forma de sinalizar, no sentido da biologia da comunicação. “Um *signal* é uma dica que foi designada por sua capacidade de transmitir a informação que ela transmite. O desenho em questão pode ser o resultado de seleção natural, seleção artificial, ou intenção consciente”³ (GREEN, 2009, p. 150. Grifo no original). Trabalhos sobre seleção de sinal sustentam que um sinal pode ser o resultado da evolução de uma *dica*, esta entendida como sendo “qualquer característica de uma entidade que transmite informação (incluindo informação equivocada)”⁴ (GREEN, 2009, p. 149)

Para ficar mais clara a diferença entre mostrar como dica e mostrar como sinal, notemos o seguinte exemplo: Em uma circunstância de constrangimento, o rubor no rosto de uma pessoa mostra o seu embaraço. Mas, ainda que a mudança na pele seja uma evidência do estado de pudor (dica), a ruborização cutânea não expressa o estado afetivo de pudor. Uma exigência para que isso seja o caso é que o mostrar seja *designado*, pela natureza (seleção de sinais) ou artificialmente (convenção), para gerar esse efeito.

Alguém poderia objetar que tudo quanto é mostrado deve ser evidente *per se* (uma crítica ao texto do exemplo). Uma distinção entre “mostrar” e “significar” pode dissolver essa questão. Uma pessoa que diz: “Estou com raiva de fulano!”, ou ainda: “Pelo menos uma vez por ano a gente poderia ter o direito de passar com caminhão em cima de alguém”; está significando sua raiva. No primeiro caso, ela assim o faz dizendo literalmente; no segundo, mediante uma implicatura.

Outra coisa, contudo, é essa mesma pessoa se comportar de maneira enfurecida (exibindo uma carranca, ofendendo o colega etc.). Aqui o estado de raiva se mostra, e um observador apropriado é capaz de perceber esse agente como estando nesse estado. Ou seja, pelo menos nos casos em que uma pessoa expresse emoções primárias universais (raiva, alegria, nojo etc.), o observador é justificado a interpretar tais expressões como sinalizando aquilo que elas mostram *prima facie*.

Acontece, no entanto, que na maioria das vezes estados psicológicos não são claros. No domínio dos afetos, por exemplo, as fronteiras entre culpa e remorso são menos nítidas do que entre raiva, nojo e alegria, que são emoções caricaturais (GAZZANIGA; HEATHERTON, 2005). Além disso, temos o agravante de que hábitos, convenções e normas sociais são fatores a condicionar a percepção dos sujeitos. A aquisição do vocabulário mental é um exemplo disso. Em condições normais, um infante é capaz de naturalmente notar estados mentais primordiais, como as intenções e emoções primárias do outro (TOMASELLO et al., 2005). Contudo, à medida que ele vai aprendendo a falar desses estados, dizendo “achar”, “pensar”, “sentir”, “querer” etc.; ele melhora a percepção e a discriminação deles (NOÉ, 2011).

O texto mencionado acima é, então, para ser visto como um sinal no sentido de Green (2007). Ele é capaz de transmitir aquilo para o que ele foi designado, o estado mental do seu enunciador (preocupação e cuidado com os adolescentes). Por essa razão, os leitores devem ser entendidos como naturalmente sensíveis a essas informações. Mesmo assim, salvo em casos de excepcionalidade, é esperado que o nível de desempenho deles varie conforme as características psicossociais de cada sujeito. Desse modo, daqui por diante, quando nós nos referirmos à

²[...] I suggest that in self-expression we designedly show what’s within.”

³ “A *signal* is a cue that was designed for its ability to convey the information it does. The design in question may be in result of natural selection, artificial selection, or conscious intention”.

⁴ “any feature of an entity that convey information (including misinformation)”.

“expressividade do texto” estamos significando o fato de um texto expressar o estado mental do(a) autor(a), e isso se mostrar aos leitores.

3 Expressividade e percepção

Uma outra propriedade do sinal é que ele mostra algo a um receptor apropriado. Ser um “receptor apropriado” significa dizer encontrar-se em condições (biológicas, psicológicas e sociais) de perceber e interpretar o sinal. Por exemplo, podemos dizer que somos receptores apropriados de sinais sociais (como contato visual), porque a evolução nos proveu dessa capacidade. Todavia, não somos receptores apropriados de alguns sinais sociais emitidos por nossos *pets* (como feromônios); uma vez que eles evoluíram em um nicho socioecológico onde o odor é relevante de comportamento. Essas alegações vão ficar mais claras com algumas notas sobre o processo perceptual.

Atualmente, a percepção é concebida como algo mais que a mera geração de perceptos, esses eliciados por estímulos sensoriais correspondentes de modalidade. Por exemplo, uma chaleira fervente apitando é um evento multimodal que inclui aspectos visuais e auditivos associados. Mas, tradicionalmente, a percepção desse evento é explicada em termos de associação de canais unissensoriais independentes: a visão processando aspectos visuais e a audição os aspectos sonoros – a integração acontecendo em nível de processamento mais alto no cérebro (NANAY; 2009; KAYSER; LOGOTHETIS, 2007; MACALUSO, 2006). Além disso, cada estímulo é visto como eliciando seu canal sensorial correspondente: estímulos visuais, a visão; estímulos acústicos, a audição; estímulos químicos, a olfação; etc.

A principal mudança conceitual nesse campo advém da descoberta de que processos perceptuais são mais abrangentes que a percepção *per se*. De acordo com essa visão, processo perceptual inclui, além da percepção como tal, a imagéria mental e a intermodalidade interna. Nanay (2017), um dos atuais defensores dessa tese, define imagéria mental como o “processamento perceptual que não é desencadeado por estimulação sensorial correspondente, em uma dada modalidade de sentido”⁵ (p. 1016); p.ex., a imagéria mental da visão de um objeto desencadeada pelos estímulos acústicos oriundos desse objeto.

Voltando ao exemplo da chaleira, uma pessoa ouvindo o apito vindo da cozinha, mas privada da visão da fonte (oclusão visual) (NANAY, 2009), vai automaticamente associar a representação auditiva desse utensílio à imagem visual dele. O que aconteceu aqui é que o estímulo auditivo (o apito) acionou no cérebro do percipiente a imagem visual da fonte (a chaleira). Casos desse tipo são também atestados no fenômeno da substituição perceptual – picadas na pele em sujeitos cegos, simulando a frequência da luz capturada por uma câmera à sua frente, eliciam imagéria visual dos objetos registrados por esse dispositivo (NANAY, 2017).

No atual estado da arte, a imagéria mental pode ser multimodal e resultar de uma ilusão criada pela mente humana. Uma evidência clássica disso é o chamado efeito McGurk (ROSENBLUM, 2008), em que a percepção de um estímulo acústico da fala é enviesada em direção à visão da articulação que alternativamente chega a ele. Na realidade, imagéria mental decorre de como nosso cérebro opera para integrar informações e dar sentido à realidade (CALVERT; THESEN, 2004). Isso é o fenômeno da imagéria mental multimodal, que “é imagéria mental em uma modalidade de sentido induzida por estimulação sensorial em outra modalidade de sentido”⁶ (NANAY, 2017, p. 1018).

Outra evidência disso é que áreas sensoriais acionadas por estímulos não correspondentes (visão causada pela audição) ficam ativas durante o processamento, e a fenomenologia decorrente é relatada pelo percipiente (SCHAEFER, 2014). Scott et al. (2013),

⁵ “[...] perceptual processing that is not triggered by corresponding sensory stimulation in a given sense modality”.

⁶ “[...] is mental imagery in one sense modality induced by sensory stimulation in another sense modality”.

p.ex., estudando o papel da imageria mental vocal na percepção de sons ambíguos da fala, descobriram que fala interna subvocal determina captura perceptual de sons linguísticos.

Imageria mental multimodal frequentemente gera sentido de presença do percepto. Em um teste sobre imageria mental visual, arquitetos com olhos vendados conseguiram realizar tarefas motoras baseando-se na imaginação de um espaço real que eles tiveram de simular internamente. Os relatos verbais introspectivos dos participantes revelaram que a imageria mental visual deles era similar às características do espaço real que serviu como modelo (PILLAI et al., 2013). Essa similaridade entre imageria interna e realidade correspondente sugere a existência de um sentido de presença contido nas imagens mentais.

Neste relato, assumimos que imageria mental multimodal está na base de outros fenômenos além da percepção, seja de estímulos exteroceptivos (visuais, sonoros etc.) ou interoceptivos (táteis, fisiológicos etc.). Ele participa também da cognição social (reconhecimento do outro e identificação social) e da cognição pragmática (inferências e percepção de autoexpressão). É assim que, em nível verbal escrito, o conhecimento trazido pelo mostrar do texto é possível. As sentenças encadeadas na superfície textual acionam significados fatuais associados a elas; e tais significados atuam como estímulos internos que ativam representações psicossociais referentes a seus enunciadores.

Com efeito, uma vez que nosso cérebro funciona para tornar as experiências inteligíveis, é esperado que as representações textuais e as representações de seu(sua) autor(a) sejam congruentes entre si. Essa integração gera coerência interna, que é crítica para a conservação da estabilidade de nossas crenças e, *ipso facto*, para a facilitação da nossa compreensão leitora. Entretanto, quando tais representações são incongruentes entre si, o sentido de incoerência decorrente pode prejudicar o entendimento sobre o texto. A primeira possibilidade era a esperada pela pesquisa. Mas a segunda, como veremos, foi a que aconteceu.

4 Plausibilidade dessa possibilidade de interpretação textual

Uma questão aberta até aqui é se essa análise que estamos propondo é plausível e mesmo factível. Consideramos que sim, e temos duas boas razões para sustentar que isso seja o caso. Uma das quais se apoia nos próprios princípios de funcionamento da linguagem, já consagrados pelos vários estudos teóricos, empíricos e experimentais: A linguagem, seja oral ou escrita, tem o potencial de modelar nossas crenças sobre o mundo. A outra repousa na verificação experimental de que incongruência entre componentes de compostos multimodais pode interferir no entendimento desses compostos com um todo (efeito *stroop*; aumento de amplitude de N400 etc.). Vejamos rapidamente cada argumento.

Desde que Sapir e Whorf alegaram que a língua falada por uma comunidade determina como seus membros vão conceitualizar o mundo, um número de trabalhos vem buscando evidências para essa suposição (BORODITISKY, 2011). No domínio experimental (para citarmos só um caso atual), pesquisas têm encontrado que a percepção e a ação das pessoas são realmente mediadas pelas categorias da língua que elas falam. Esse é o caso, p.ex., da pesquisa sobre como as emoções são percebidas, realizada por Lindquist e Gendron (2013). Elas descobriram que a linguagem cumpre um papel constitutivo na percepção emocional, na medida em que as palavras servem como contexto para os perceptos visuais afetivos. Isso seria assim porque as palavras fundamentam as instâncias altamente variáveis de uma categoria emocional.

Nesse sentido, se a cognição é ancorada também na expressividade do texto (o qual envolve ele mesmo imageria mental multimodal), logo esperamos que os significados da autoexpressão interpenetrem a significação referencial do texto. Voltando a *Gravidez na adolescência*, os leitores vão entender os argumentos aduzidos pelo enunciador como sendo função de quem quer conscientizá-los dos perigos da adolescência, e impedi-los de cair na armadilha da gravidez indesejada. Isso é o conhecimento interpessoal permitido pela autoexpressão. Disso resulta que uma incoerência nesse nível da autoexpressão vai interferir

negativamente na compreensão leitora, ao prejudicar o entendimento do referente textual. Isso é o que suspeitamos que aconteça quando os leitores notam a incongruência entre o que o(a) autor(a) mostra pelo texto e o que ele(a) o faz através da foto.

No tocante às possíveis incongruências entre as imagens de face geradas pelo texto e pela foto, vale destacar o papel que os gestos faciais cumprem para a validade da comunicação. Por exemplo, Mehu (2015, p. 4) afirma que “ao avaliar a confiabilidade de um sinal, observadores usam outras dicas ou índices presentes no sinal os quais são difíceis de ser manipulados ou voluntariamente controlados, por exemplo, componentes emocionais”⁷. Diferente da linguagem escrita, a qual pode dissimular os reais propósitos da interação, gesticulação facial e corporal é mais difícil de falsificar; tal que é esperado que os leitores usem esses dados extralinguísticos como pistas sobre a honestidade da comunicação.

Se é verdade que a compatibilidade intermodal entre textos e fotos pode facilitar a compreensão, é também verdade que a não correspondência entre esses dois componentes pode causar interferências negativas na leitura. De um lado, a psicologia cognitiva já evidenciou que, quando se rompe a associação automática entre duas modalidades habitualmente pareadas (p.ex., palavra “amarelo” escrita na cor amarela), o processamento fica mais demandante. Isso é um caso típico de efeito *stroop* clássico (para uma revisão, veja MacLeod (1991)). De outro, dados neurocientíficos de eletrofisiologia apontam que a violação de expectativa semântica (e não só) provoca um aumento da amplitude de N400 – um sinal de EEG⁸ evocado por eventos (ERP⁹) que é modulado por quebra de expectativas semântico-contextuais (para uma revisão, veja Kutas e Federmeier (2000)). Por exemplo, a ocorrência não esperada de palavras em um contexto comunicativo particular.

Porque esses eventos de incongruência de correspondência multimodal geram quebra de expectativas, eles demandam por extensão mais recursos de função executiva, como memória de trabalho (DIAMOND, 2013). Esse fato pode explicar os dados que encontramos em nossa pesquisa.

5 Metodologia

Como já dito, o que motivou essa pesquisa foi responder à questão de se elementos extralinguísticos presentes na situação de leitura, como a imagem fotográfica do(a) autor(a), que potencialmente gera significados psicossociais, poderiam funcionar como pistas; e, em caso positivo, se eles poderiam contribuir com a compreensão leitora. A seguir, apresentamos os metodologia empregada para cumprir esse propósito.

Amostra: Os dados desta análise foram obtidos de uma pesquisa sobre compreensão leitora entre um grupo de alunos de 9. ano escolar ($n = 26$, idade média = 14.16, desvio padrão = 0.76). Todos os 26 alunos participaram dos testes, com o consentimento de seus responsáveis legais. Um termo de Consentimento Livre e Esclarecido foi, assim, assinado pelo participante e pelo seu responsável.

Estímulos: 2 textos opinativos que tratam, cada qual, de temas conhecidos dos adolescentes (*As redes sociais* (texto *a*) e *Gravidez na adolescência* (texto *b*))¹⁰. Eles foram criados de maneira controlada, seguindo os mesmos critérios de elaboração (Figuras 1 e 2). Ambos os textos continham um parágrafo crítico, o qual exibia uma lacuna a ser preenchida por uma das opções oferecidas na sequência (“a”, “b”, “c” ou “d”). Esse parágrafo foi elaborado com o objetivo de testar se os leitores eram capazes de executar uma tarefa de inferência de forma esperada.

⁷ “[...] when evaluating the trustworthiness of a signal, perceivers use other cues or indices present in the signal that are difficult to manipulate or control voluntarily, for example emotional components.”

⁸ EEG é sigla para eletroencefalografia.

⁹ ERP é sigla em inglês para “potencial cerebral evocado por eventos”.

¹⁰ Vale lembrar que um terceiro texto também foi usado na pesquisa (“Maioridade penal”) como estímulo de pré-teste.

Inferência foi tomada na pesquisa como significando capacidade de tirar conclusões a partir de argumentos. Mas fazer inferências durante a leitura requer esforço cognitivo, pois trata-se de uma forma de função executiva para cujo sucesso é importante que o(a) leitor(a) iniba estímulos distratores e retenha só aquilo que é funcional (DIAMOND, 2013). Era esperado então que, lendo atentamente os textos, os alunos fossem capazes de escolher a opção que era uma decorrência lógica das informações anteriores. Em certo sentido, todas as opções eram possíveis, se a compreensão leitora fosse destituída de recursos de atenção. Após uma leitura minuciosa dos textos, nós definimos aquela que, para cada caso, fosse a mais apropriada (gabarito). E consideramos esses gabaritos como referência para julgarmos os desvios.

Figura 1

As redes sociais

Mesmo trazendo alguns bads, a internet vem mudando algumas práticas de convívio com a galera, especialmente quando aliada aos dispositivos móveis. A razão de tudo isso é que ela traz importantes lacres sociais, como estreitar os laços de amizades entre os migos.

Pensando no contato com a family atualmente é um desafio para muitos jovens manter-se em contato constante com seus parentes através do telefone ou até mesmo por e-mail.

É a oportunidade então de perguntarmos como a web aprimora o contato com o pessoal nos dias de hoje?

A nosso ver, pela disponibilidade da web especialmente em smartphones, que torna a comunicação possível de qualquer lugar e em qualquer momento. Essa disponibilidade oportuniza o compartilhamento simultâneo de experiências, conectando o usuário a seus miguxos e family.

É dessa maneira que atualmente as pessoas _____, ao fazerem uso do celular.

(a) ...estão se conectando mais e mais...
 (b) ...estão vencendo as distâncias físicas...
 (c) ...estão aprendendo a interagir na rede...
 (d) ...estão ganhando visibilidade social...

Concluimos que a disponibilidade da web em dispositivos móveis é a condição perfeita para a vida social. As pessoas vão cada vez mais se beneficiar disso, para estreitarem seus laços com os outros.

Figura 2

Gravidez na adolescência

A adolescência é uma das fases mais prazerosa da vida. É a época do aflorar da sexualidade. Infelizmente, ela também pode ser trágica. A iniciação sexual sem o devido cuidado pode acarretar um desastre, que tem acabado com os sonhos de muitos jovens: a gravidez indesejável.

Uma gravidez não planejada pode aniquilar os projetos de vida de um adolescente. Segundo pesquisas, a gravidez na adolescência está associada com o abandono escolar, a perpetuação da pobreza e a exclusão social.

Por que, mesmo com toda informação existente, isso ainda acontece?

Tenho certeza de que a causa é a droga chamada paixão, que entorpece a pessoa e é mais forte entre os adolescentes. O apaixonado vira um idiota. Ele acha que está numa relação afetiva segura e que pode se entregar ao outro sem medo; achando que nada de ruim pode acontecer a ele.

É assim que muitos jovens passionais acreditam que _____, estupidamente ignorando os métodos anticoncepcionais.

(a) ...o amor não tem limites...
 (b) ...a vida é para ser vivida...
 (c) ...estão livres de riscos...
 (d) ...no sexo vale tudo...

Fica então a dica! Na hora de namorar, seja consciente e responsável de seus atos. Assim, você está se livrando de ser pego de surpresa por uma desastrosa gestação precoce.

As imagens fotográficas, por sua vez, também foram produzidas em condições similares. Elas foram criadas para serem afetivamente congruentes com os textos (valência positiva = retrato do rapaz + redes sociais; valência negativa = retrato da moça + gravidez na adolescência); tal que elas pudessem servir como pistas a facilitarem a compreensão textual. *Procedimentos*: Foram solicitados aos participantes ler dois textos, um em cada nível de multimodalidade: “unimodal” (só texto) e “bimodal” (texto + foto); e escolher entre quatro alternativas (“a”, “b”, “c” e “d”) aquela que fosse mais apropriada ao preenchimento da lacuna do parágrafo crítico (Figuras 1 e 2). Em cada caso foi estabelecido um gabarito: alternativa “b” para *As redes...* e “c” para *Gravidez...* Após cada teste, os participantes assinalavam o grau de certeza da escolha feita, em uma escala ordinal (tipo Likert) variando de 1 a 10.

É importante salientar que, como os sujeitos participaram em ambas as condições (“unimodal” e “bimodal”), elegemos a condição “só texto” como controle da condição “texto + foto”. Assim, os alunos foram controles deles mesmos. P.ex., um aluno que lia um texto (*a*) na condição “unimodal”, lia o outro (*b*) na “condição bimodal”; e vice-versa.

6 Resultados

Cada um dos 26 participantes executou dois testes, perfazendo um total de 52 observações (26 X 2 = 52 observações). Nós utilizamos um método estatístico para elucidarmos se as diferenças entre os níveis eram significativas, a estatística do qui-quadrado (χ^2) (UGONI; WALKER, 1995). Essa estatística é frequentemente empregada quando as variáveis (dependente e independente) são qualitativas (nominal) e o número de observações é relativamente pequeno. Ela permite revelar se a diferença entre duas distribuições, ou uma apenas, é significativa ao nível de 5% de erro (p -valor $\leq 0,05$). Basicamente, ela compara frequências observadas com frequências esperadas, e determina se as diferenças obtidas se devem ao acaso ou não.

A análise usando essa estatística nos devolveu as seguintes diferenças:

Tabela 1
Distribuições observadas entre níveis

Níveis	Alternativas				Total
	A	B	C	D	
Unimodal	10	7	8	1	26
Bimodal	4	9	8	5	26
Total	14	16	16	6	52

Os números obtidos pelo teste de hipótese indicam que a diferença entre as distribuições é significativa (*teste* = 2601., *p*-valor = 0). Isso significa que os níveis de leitura considerados são dependentes entre si, mas não informa que tipo de dependência existe entre eles. Quanto aos testes de avaliação (formulados para mensurar o nível de certeza das respostas escolhidas), os dados trazem informações, senão significativas ($n = 26$, *teste* = 8., *p*-valor = 0.078), pelo menos sugestivas. Vejamos a tabela abaixo:

Tabela 2
Números dos testes de avaliação

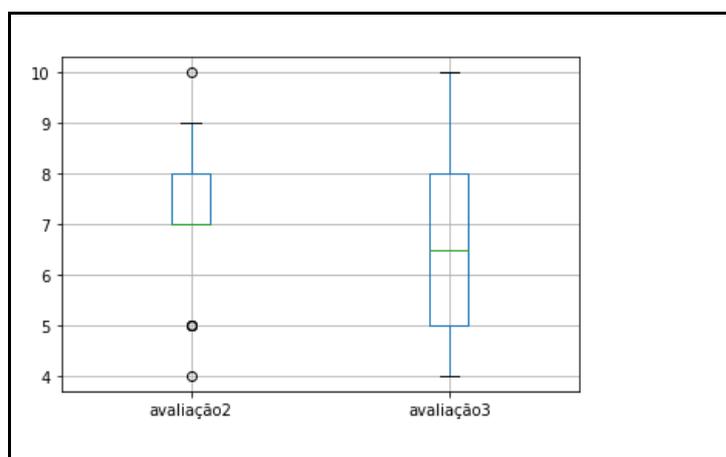
	Níveis	Mínimo	Máximo	Mediana
avaliação2	<i>Unimodal</i>	4	10	7.
avaliação3	<i>Bimodal</i>	4	10	6.5

Notemos que as medianas¹¹ são levemente diferentes, o valor mais alto sendo o da avaliação relativa ao texto *a*. Isso sugere um nível de certeza um pouco melhor que o verificado com o texto *b*. (Voltaremos a essa interpretação na seção seguinte, quando os dados acima vão ser confrontados com o gráfico de caixa, Figura 3 abaixo, o qual exhibe a variação dos escores dos participantes.)

7 Análise e discussão dos dados

Esta pesquisa lançou a hipótese de que a multimodalidade podia influenciar a compreensão leitora, devido à relevância de pistas extraverbais durante a leitura. Os resultados obtidos confirmaram que esse é caso, *mas* em termos unicamente de significância de diferença (Tabela 1). Quando observamos o gráfico de caixa da Figura 3, passamos a ter uma ideia mais clara dos fatores que motivaram essa diferença. Notemos a diferença no tamanho das caixas no gráfico.

Figura 3
Pontos mostrados acima e abaixo da “caixa” em avaliação2 são os *outliers*.



A diferença encontrada entre esses testes aponta que o grau de certeza foi maior no nível modal (avalição2). Isso porque o escore mais estendido em avaliação3 (entre 5 e 8) sugere uma correlação entre a distribuição das respostas e o grau de incerteza dos participantes.¹² Isso quer dizer que os sujeitos estavam menos seguros para tomar suas decisões e, *ipso facto*, eles flutuaram mais em suas respostas. Essa incerteza provavelmente decorreu de conflitos socioafetivos, provocados pelo pareamento incongruente entre as fotos dos autores e seus respectivos textos. (De outro modo, o que justificaria os sujeitos variarem suas respostas tal como eles fizeram em condição bimodal?)

Esses achados estão em linha com o que é encontrado nas pesquisas sobre incongruência em correspondência multimodal (CHEN; SPENCE, 2017). Variação maior de respostas em testes de avaliação em geral correlaciona com grau de incerteza nas tomadas de decisão, em situações de incongruência entre as modalidades. Esse comportamento é *mutatis mutandis* similar ao que ocorre no efeito *stroop*, bem como na modulação de N400. Como já dissemos, quando existe incoerência de alguma ordem no sinal, o processamento é cognitivamente mais demandante. E isso se reflete em interrupção do fluxo cognitivo.

¹¹ Mediana é uma medida de tendência central frequentemente usada em análise estatística de amostras pequenas e com distribuições não-paramétricas.

¹² Isso não equivale a dizer, contudo, que eles são estaticamente significativos (note-se $p = 0.07$).

A seguir, serão feitos alguns esclarecimentos sobre essa interpretação. Especialmente pretendemos clarear se esses estímulos visuais serviram como dicas extraverbais, e se eles contribuíram com a compreensão leitora.

7.1 A nulidade dos estímulos visuais como pistas para a leitura

Como os dados acima não deixam claro qual o efeito da multimodalidade sobre a leitura, realizamos uma análise mais minuciosa, desta vez cruzando as condições “temas do texto” (texto *a* vs texto *b*) e “multimodalidade” (unimodal vs bimodal). Para a análise estatística, usamos uma vez mais o método do qui quadrado (χ^2), que nos devolveu os dados exibidos na Tabela 3. Podemos observar que esses dados indicam que a compreensão leitora foi mais dispersiva em nível bimodal. Contudo, a significância dessa distribuição foi encontrada unicamente na situação do texto *a*, “As redes sociais” (teste = 17.076, $p = 0.01$). Quanto à situação do texto *b*, “Gravidez na adolescência”, mesmo sendo não significativa em termos estatísticos (teste = 10.307, $p = 0.17$), ela segue a tendência de dispersão verificada na situação do texto *a*.

Tabela 3

Distribuições das frequências de respostas entre condições. O sobrescrito (“asterisco”) indica o gabarito em cada caso.

Alternativas	“As redes sociais”		“Gravidez na adolescência”	
	Unimodal	Bimodal	Unimodal	Bimodal
A	9	2	1	3
B	3*	5*	3	3
C	1	4	8*	4*
D	0	2	1	3

Como o único fator de diferenciação eram as imagens fotográficas acompanhando os textos, concluímos que os estímulos visuais foram determinantes para esses resultados. No entanto, eles atuaram no sentido de interferir na escolha das alternativas e não no de facilitar a escolha delas. A causa mais provável desse efeito é que, ainda que fossem em alguma medida relevantes, as pistas extraverbais mais confundiram que esclareceram. (Daí as informações plotadas na Figura 3 acima.) Essa interpretação é consistente com os resultados dos testes de avaliação comentados na seção anterior. Existe um conflito entre o mostrar do texto e o mostrar da foto, o qual interfere negativamente na compreensão leitora.

O discurso dos textos é uma instância de “mostrar-que” (*showing-that*), um tipo de mostrar que Green (2010; 2009; 2008; 2007) opôs a outros dois: “mostrar- α ” (*showing- α*) e “mostrar-como” (*showing-how*). Mostrar-que, segundo Green (2009), é a autoexpressão que permite conhecimento proposicional (p.e., “Eu poderia, por meio de cálculos extensivos, mostrar *que* há um buraco negro no centro da Via Láctea”¹³ (GREEN, 2009, p. 141)). A categoria de mostrar- α inclui, por sua vez, as formas de expressão propriamente física que evidenciam o estado corporal e afetivo do indivíduo (p.e., “Eu mostro meu machucado, e por esse meio permito aos outros ver esse machucado”¹⁴ (GREEN, 2009, p. 141)). Por fim, Green esclarece o que entende por mostrar-como em termos de autoexpressões que permitem conhecimento experiencial ou empatia. “Mostrar-como pode fornecer conhecimento qualitativo para aqueles com a capacidade para empatia”¹⁵ (GREEN, 2009, p. 142). Por exemplo, “a

¹³ “[...] I might, by means of extensive calculations, show *that* there is a black hole in the center of the Milk Way”.

¹⁴ “I show my bruise, and thereby enable others to see that bruise”.

¹⁵ “Showing-how can provide qualitative knowledge for those with appropriate sensory capacities”.

trepidação em minha voz poderia lhe permitir conhecer como minha ansiedade se parece se você fosse suficientemente empático”¹⁶ (GREEN, 2009, p. 142).

Em um estudo¹⁷, considerando quais sentidos psicossociais os jovens reconhecem nos autores, encontramos que os adolescentes na idade dos nossos participantes exibem um padrão de autoexpressão. Eles buscam “amparo” dos adultos e “amizade” dos colegas, além da “expertise” de fontes confiáveis de informação. Por isso, desde que o objetivo (extralinguístico) dos autores seja a persuasão (como parece ser o caso aqui), não é novidade que eles apelem para o desejo dos leitores por amparo, amizade e expertise e, *ipso facto*, emitam sinais que sejam designados para isso. Mas isso parece entrar em choque com o que o retrato dos autores mostra-α. Esclarece Green (2010, p. 67) que “expressão facial, frequentemente agrupada com entonação e outros comportamentos não-verbais, mostram estados de mente, em particular emoções, na forma de mostrar-α”¹⁸. Os autores das imagens fotográficas usadas nesta pesquisa mostram a atitude deles ao torná-la um objeto perceptível¹⁹.

Em suma, as imagens fotográficas enriqueceram a situação de leitura, por possibilitarem inferências acerca das atitudes dos autores. Mas os participantes não conseguiram estabelecer uma coerência entre elas, tal que conduzisse a uma compreensão apropriada dos textos. O resultado foi que a interpretação ficou relativamente solta. Notemos, quanto a isso (Tabela 3), que com o texto *a* a acurácia das inferências melhora levemente (gabaritos: 3 > 5); já com o texto *b* o efeito foi o oposto (gabaritos: 8 > 4). Porque não se verificou um padrão de melhora dos desempenhos; os estímulos visuais não cumpriram a função esperada de esclarecer o entendimento dos textos.

Em suma, podemos dizer que as imagens visuais dos autores interferiram no desempenho dos leitores. Isso equivale a dizer que elas prejudicaram o desempenho deles na tarefa crítica de inferência.

7.2 Expressividade, intermodalidade e leitura

O fenômeno ora relatado ganha uma nova compreensão quando considerado à luz do exposto nas seções 1 e 2. Supomos que estamos aqui diante de um caso de incongruência dentro de processos perceptuais envolvidos na leitura, e que esse fato cognitivo tem o potencial de desorientar os leitores. Há duas razões que sustentam essa suposição. Uma é a admissão de que a escrita é capaz de “mostrar o que está dentro” na medida em que ela aciona uma imageria mental referente ao seu autor. A outra é que essa imageria, quando não é correspondente ao que é esperado, pode afetar o nível de atenção dos leitores.

Quanto à primeira assunção, o fato de a comunicação oral, como a conversação, ser um evento multimodal determina como vamos perceber a expressividade do texto. Ao falarmos com um conhecido, p.ex., temos uma experiência global na qual diferentes componentes formam um todo inteligível que é a experiência de comunicação oral; i.e., os atos de fala não são fenômenos impessoais. As palavras de nosso conhecido soam de modo particular, com entonação, ritmo e velocidade próprios (seu estilo ou idioleto). Elas sempre vêm associadas a um rosto particular, a fonte da emissão vocal. Além disso, os atos linguísticos que ele executa são incorporados em gesticulação facial e corporal, e por aí vai. Essa descrição se encaixa em um fenômeno estudado pela psicologia conhecido por “correspondência intermodal”.

Correspondência intermodal “se refere a um efeito de compatibilidade entre atributos ou dimensões de um estímulo (i.e., um objeto ou evento) em diferentes modalidades sensoriais

¹⁶ “[...] the trepidation in my voice might enable you to know how my anxiety feels if you are sufficiently empathetic.

¹⁷ Dados não publicados pertencentes ao Projeto de Pesquisa “Teoria da mente e compreensão textual. Como condições de produção podem influenciar a leitura de intenção comunicativa”, coordenado pelo Prof. Dr. Wagner Ferreira Lima (UEL).

¹⁸ “[...] facial expression, often coupled with intonation and other non-verbal behavior, shows states of mind, in particular emotions, in the showing α way [...]”

¹⁹ As fotos usadas na pesquisa foram preparadas para serem congruentes com a valência dos textos. Por motivo de privacidade dos modelos (colaboradores da pesquisa), elas não foram disponibilizadas como anexos neste artigo.

(sejam elas redundantes ou não)”²⁰ (SPENCE, 2011, p. 973). Porque os diferentes componentes estão integrados, o receptor pode considerar correspondência intermodal para melhorar a compreensão do sinal (KELLY; ÖZYÜREK; MARIS, 2010). Ele pode também se valer dela para precisar a origem do som a partir de pistas visuais (BONAH *et al.*, 2007; SPENCE, 2007; ALAIS; BURR, 2004). Pode se basear em pistas de voz para reconhecimento de identidade pessoal (CAMPANELLA; BELIN, 2007; VATAKIS, SPENCE, 2007; LACHS; PISONI, 2004; KAMACHI *et al.* 2003); e em estímulos visuais para aprendizagem de voz (SHEFFERT; OLSON, 2004).

É assim que, durante a leitura, as palavras desencadeiam o sentido de presença do enunciador, mesmo este não estando fisicamente lá (como no exemplo anterior da percepção da chaleira). Textos escritos têm o potencial de evocar aspectos intersubjetivos presentes na oralidade, como afetos, atitudes, enfim, crenças sociais; e de *mostrar* com isso os estados psicológicos do autor. Mas, e quando os estímulos externos não correspondem aos aspectos da imagieria mental evocados pelo texto, violando assim o princípio da integração multimodal? Nesse caso, temos os fenômenos de interferência que encontramos na pesquisa ora relatada. Chegamos assim à nossa segunda razão.

Imagieria mental intermodal podem pré-ativar a percepção de estímulos correspondentes a uma dada modalidade sensorial. Se é verdade que os alunos que leram o texto da Figura 1 criaram uma representação amigável do enunciador, essa representação pode ter vindo por terra quando confrontada com o retrato da pessoa real desse enunciador. Assim, se a realidade frustra a expectativa, é esperado um esforço cognitivo maior da parte dos alunos para resolver esse conflito.

Isso explica, a princípio, por que a imagem do(a) autor(a), pareada com o texto, não funcionou como pistas a facilitarem a compreensão leitora; uma vez que a correspondência intermodal era esperada cumprir essa função cognitiva. Uma leitura elaborada, como a que foi exigida dos participantes, demanda processos inferenciais em vários níveis. Como dissemos antes, inferência é uma forma de função executiva que se baseia na inibição de estímulos distratores e na retenção unicamente daquilo que é funcional. Essa tarefa é amena quando a integração multimodal é suave. Mas, quando a integração é incongruente, o evento distrator, i.e., o aspecto incongruente (no caso a imagem fotográfica dos autores), desvia a atenção; e isso prejudica a compreensão textual.

8 Considerações finais

Gesticulação facial e corporal é uma dica relevante sobre como os interlocutores devem se engajar na interação verbal. Questionamos se isso poderia acontecer também durante a leitura e se, além de engajar os leitores, isso pudesse facilitar a compreensão leitora. Ao contrário do que esperávamos, o pareamento do texto com a fotografia do(a) seu(sua) autor(a) produziu um efeito de interferência cognitiva. Houve uma dispersão nas respostas na condição bimodal (pareamento texto e foto), mais do que na condição unimodal (só texto). Explicamos esses efeitos pela incongruência entre como o texto *mostra* o seu autor e como a imagem visual o faz. Mesmo a linguagem escrita contém uma dimensão expressiva similar à da oralidade, expressividade essa criada pela imagieria mental multimodal do texto. Assim, o(a) autor(a) autoexpressa o estado mental em que se encontra durante a enunciação; e os leitores, por sua vez, são capazes de compreender a sinalização dele(a). Porque essa incongruência gera incoerência intermodal, ela dispersa a atenção dos leitores dando origem aos efeitos expostos acima.

²⁰ “[...] to refer to a compatibility effect between attributes or dimensions of a stimulus (i.e., an object or event) in different sensory modalities (be they redundant or not)”.

Mesmo não oferecendo resultados conclusivos, esta pesquisa abre a possibilidade de se estudar a leitura em outras bases teóricas. Observar o nível de expressividade do texto em termos dos processos perceptuais descritos acima é um dos caminhos. Isso permite entender como os leitores se relacionam com o texto, quais processos perceptuais e cognitivos estão envolvidos, e se efetivamente a intermodalidade é nula como dica extraverbal. Seja lá como for, podemos dizer que este trabalho fez uma importante contribuição com as pesquisas sobre interpretação textual. A implicação para o ensino de leitura entre os adolescentes é a seguinte: É preferível usar imagens pareadas com textos unicamente para redundar o significado; e não para informar características dos autores. Isso pode confundir mais que ajudar os leitores juvenis.

Referências

- ALAIS, D.; BURR, D. The ventriloquist effect results from near-optimal bimodal integration. **Current Biology**, v. 14, n. 3, p. 257-262, 2004. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.407.9929&rep=rep1&type=pdf>> Acesso em: 04 jul. 2020.
- BONATH, B. et al. Neural basis of the ventriloquist illusion. **Current Biology**, v. 17, p. 1697-1703, 2007. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0960982207019069>> Acesso em: 04 jul. 2020.
- BORODITISKY, L. How language shapes the thought. The language we speak affects our perceptions of the world. **Scientific American**, p.63-65, 2011. Disponível em: <<https://www.gwashingtongh.org/ourpages/auto/2013/10/23/68598699/sci-am-2011.pdf>> Acesso em: 09 jul. 2020.
- BOUKO, C. Emotions through texts and images: a multimodal analysis of reactions to the Brexit vote on Flickr. **Pragmatics**, v. 30, n. 2, p. 222-246, 2020. Disponível em: <<https://www.ingentaconnect.com/content/jbp/prag/2020/00000030/00000002/art00004>> Acesso em: 04 jul. 2020.
- CALVERT, G. A.; THESEN, T. Multisensory integration: methodological approaches and emerging principles in the human brain. **Journal of Physiology**, v. 98, n. 1-3, p. 191-205, 2004. Disponível em: <<https://europepmc.org/article/med/15477032>> Acesso em: 10 jul. 2020.
- CAMPANELLA, S.; BELIN, P. Integrating face and voice in person perception. **Trends in Cognitive Sciences**, v. 11, n. 12, p. 535-543, 2007. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S1364661307002495>> Acesso em: 04 jul. 2020.
- CHEN, Y.; SPENCE, C. Assessing the role of the “unity assumption” on multisensory integration: A review. **Frontiers in Psychology**, v. 8, 2017. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00445/full>> Acesso em: 11 jul. 2020.
- DIAMOND, A. Executive functions. **Annual Review of Psychology**, v. 64, p. 135-138, 2013. Disponível em: <<https://www.annualreviews.org/doi/full/10.1146/annurev-psych-113011-143750>> Acesso em: 04 jul. 2020.
- GARROD, S.; PICKERING, M. J. Joint action, interactive alignment, and dialog. **Topics in Cognitive Sciences**, v. 1, p. 292-304, 2009. Disponível em: <<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1756-8765.2009.01020.x>> Acesso em: 07 jul. 2020.
- GAZZANIGA, M. S.; HEATHERTON, T. F. Emoções, estresse e *coping*. In: _____ **Ciência Psicológica**. Mente, cérebro e comportamento. Trad. Maria Adriana Verissimo Veronese. Porto Alegre: Artmed, 2005. p. 312-341.

- GREEN, M. S. Précis of Self-expression (Oxford, 2007). **Acta Anal**, v. 25, p. 65-69, 2010. Disponível em: <<https://link.springer.com/article/10.1007/s12136-009-0082-0>> Acesso em: 04 jul. 2020.
- GREEN, Mitchell S. Speech acts, the handicap principle and the expression of psychological states. **Mind & Language**, v. 24, n. 2, p. 139-163, 2009. Disponível em: <<https://www.semanticscholar.org/paper/Speech-Acts%2C-the-Handicap-Principle-and-the-of-Green/e96efc4c51943ee89775791e2c0e3f8b1247c7dd>> Acesso em 12 jul. 2020.
- GREEN, M. S. Empathy, Expression, and What Artworks Have to Teach. In: HAGBERG, G. (ed.). **Art and Ethical Criticism**. Blackwell, 2008. p. 99-122. Disponível em: <https://www.researchgate.net/publication/228046519_Empathy_Expression_and_What_Art_works_Have_to_Teach> Acesso em: 04 jul. 2020.
- KAMACHI, M. et al. “Putting the face to the voice”: matching identity across modality. **Current Biology**, v. 13, p. 1709-1714, 2003. Disponível em: <[https://www.cell.com/current-biology/pdf/S0960-9822\(03\)00663-8.pdf](https://www.cell.com/current-biology/pdf/S0960-9822(03)00663-8.pdf)> Acesso em: 04 jul. 2020.
- KAYSER, C.; LOGOTHETIS, N. K. Do early sensory integrate cross-modal information? **Brain Structure and Function**, v. 212, p. 121-132, 2007. Disponível em: <<https://link.springer.com/article/10.1007/s00429-007-0154-0>> Acesso em: 10 jul. 2020.
- KELLY, S.; ÖZYÜREK, A.; MARIS, E. Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. **Psychological Science**, v. 21, n. 2, p. 260-267, 2010. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.701.7387&rep=rep1&type=pdf>> Acesso em: 04 jul. 2020.
- KUTAS, M.; FEDERMEIER, K. D. Electrophysiology reveals semantic memory use in language comprehension. **Trends in Science Cognitive**, v. 4, n. 2, p. 463-470, 2000. Disponível em: <https://www.academia.edu/491732/Electrophysiology_reveals_semantic_memory_use_in_language_comprehension> Acesso em: 09 jul. 2020.
- LACHS, L.; PISONI, D. B. Crossmodal source identification in speech perception. **Ecological Psychology**, v. 16, n. 3, p. 159-187, 2004. Disponível em: <<https://psycnet.apa.org/record/2004-21380-001>> Acesso em: 02 jul. 2020.
- LINDQUIST, K. A.; GENDRON, M. What’s in a word? Language constructs emotion perception. **Emotion Review**, v. 5, n. 1, p. 66-71, 2013. Disponível em: <<https://psycnet.apa.org/record/2013-00258-011>> Acesso em: 09 jul. 2020.
- MacLEOD, C. M. Half a century of research on the Stroop effect: An integrative Review. **Psychological Bulletin**, v. 109, n. 2, p. 163-203, 1991. Disponível em: <https://pure.mpg.de/rest/items/item_2355497/component/file_2355496/content> Acesso em: 08 jul. 2020.
- MACALUSO, E. Multisensory processing in sensory-specific cortical areas. **The Neuroscientist**, v. 12, n. 4, p. 327-338, 2006. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.113.4778&rep=rep1&type=pdf>> Acesso em: 10 jul. 2020.
- MEHU, M. The integration of emotional and symbolic components in multimodal communication. **Frontiers in Psychology**, v. 6, 2015. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00961/full>> Acesso em: 07 jul. 2020.
- NANAY, B. Multimodal mental imagery. **Cortex**, v. 105, p. 125-134, 2018. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010945217302277>> Acesso em: 04 jul. 2020.

- NANAY, B. Sensory Substitution and Multimodal Mental Imagery. **Perception**, v. 46, n. 9, p. 1014-1026, 2017. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010945217302277>> Acesso em: 04 jul. 2020.
- NANAY, B. Perception and imagination: amodal perception as mental imagery. **Philosophical Studies: An International Journal for Philosophy in Analytical Tradition**, v. 150, n. 2, p. 239-254, 2009. Disponível em: <<https://www.jstor.org/stable/40856553?seq=1>> Acesso em: 09 jul. 2020.
- NOÉ, P. A. A. B. **Teoria da Mente e Compreensão Leitora**: um Estudo com Alunos Participantes de um Programa de Desenvolvimento Sociocognitivo. Dissertação de Mestrado. Juiz de Fora: Universidade Federal de Juiz de Fora, 2011. Disponível em: <<https://www.ufjf.br/ppgpsicologia/files/2010/01/Priscilla-Aquino-Batista-No%c3%a9.pdf>> Acesso em: 10 jul. 2020.
- PILLAI, J. S. et al. Presence in visual mental imagery. In: **International Visual Informatics Conference (IVIC)**, Malaisia, 2013-11 - Advances in Visual Informatics – 2013. Disponível em: <<https://hal.archives-ouvertes.fr/hal-01206642/document>> Acesso em: 10 jul. 2020.
- SCOTT, M. et al. Inner speech captures perception of external speech. **The Journal of Acoustical Society of America**, v. 133, n. 4, p. 286-292, 2013. Disponível em: <<https://asa.scitation.org/doi/10.1121/1.4794932>> Acesso em: 10 jul. 2020.
- SHEFFERT, S. M.; OLSON, E. Audiovisual speech facilitates voice learning. **Perception & Psychophysics**, v. 66, n. 2, p. 352-362, 2004. Disponível em: <<https://link.springer.com/content/pdf/10.3758/BF03194884.pdf>> Acesso em: 04 jul. 2020.
- ROSENBLUM, L. D. Speech perception as multimodal phenomenon. **Current Directions in Psychological Sciences**, v. 17, n.6, p. 405-409, 2008. Disponível em: <<https://psycnet.apa.org/record/2008-18235-010>> Acesso em: 10 jul. 2020.
- SCHAEFER, R. S. Mental representations in musical processing and their role in action-perception loops. **Empirical Musicology Review**, v. 9, n. 3-4, p. 161-176, 2014. Disponível em: <<https://emusicology.org/article/view/4291>> Acesso em: 10 jul. 2020.
- SPENCE, C. Crossmodal correspondences: a tutorial review. **Attention, Perception and Psychophysics**, v. 73, p. 971-995, 2011. Disponível em: <<https://link.springer.com/article/10.3758/s13414-010-0073-7>> Acesso em: 04 jul. 2020.
- SPENCE, C. Audiovisual multisensory integration. **The Acoustical Society of Kapan**, v. 28, n. 2, p. 61-70, 2007. Disponível em: <https://www.jstage.jst.go.jp/article/ast/28/2/28_2_61/_article> Acesso: 04 jul. 2020.
- TOMASELLO, M. et al. Understanding and sharing intentions: The origins of cultural cognition. **Behavioral and Brain Sciences**, v. 20, p. 000-000, 2005. Disponível em: <https://www.eva.mpg.de/documents/Cambridge/Tomasello_Understanding_BehBrainSci_2005_1555292.pdf> Acesso em: 10 jul. 2020.
- VATAKIS, A.; SPENCE, C. Cross-modal binding: Evaluating the “unity assumption” using audiovisual speech stimuli. **Perception & Psychophysics**, v. 69, n. 5, p. 744-756, 2007. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/17929697/>> Acesso em: 04 jul. 2020.
- WILLIAMS, R. M.; HAGOORT, P. Neural evidence for the interplay between language, gesture, and action: A review. **Brain and Language**, v. 101, n. 3, p. 278-289, 2007. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0093934X07000478>> Acesso em: 07 jul. 2020.