

Metaconceitos, *Insight* e Estrutura Invisível: por que a educação linguística é mais do que decorar listas de nomes — e o que isso revela sobre os limites da IA

Marcus Maia – UFRJ/CNPq



Perspectiva: você vê um cálice preto num fundo branco ou dois rostos brancos de perfil num fundo preto? No desenho Karajá (no alto à esquerda), você vê padrões brancos num fundo preto ou padrões pretos num fundo branco?

Resumo: O artigo argumenta que a educação linguística deve ultrapassar a simples memorização de termos gramaticais e promover o desenvolvimento de metaconceitos cognitivos que reorganizam a percepção e a compreensão da linguagem. Com base em evidências de rastreamento ocular (*eye tracking*), procura-se demonstrar que diferenças no custo de processamento entre orações relativas de sujeito e de objeto, bem como entre relações de causalidade direta e inferência, refletem mudanças de perspectiva e a necessidade de construir e reorganizar estruturas sintáticas e conceituais invisíveis. Defende-se que a noção de **perspectiva** constitui um metaconceito central, articulando-se às noções de atividade epilinguística (Franchi, 1987), de *threshold concepts* (Meyer e Land, 2005) e ao princípio *Structures, Not Strings* (Everaert *et al.*, 2015), segundo o qual a linguagem humana opera sobre relações estruturais hierárquicas, e não apenas sobre sequências lineares de palavras. A aprendizagem profunda ocorre quando os alunos atravessam um limiar conceitual que transforma sua forma de perceber a linguagem, tornando conscientes mecanismos antes apenas intuitivos; nesse processo, o *eye tracking* pode funcionar como recurso pedagógico ao tornar visíveis os próprios processos cognitivos de análise estrutural, leitura e interpretação. Por fim, o artigo contrasta essa reorganização metacognitiva com o funcionamento dos grandes modelos de linguagem (LLMs), argumentando que, embora altamente eficientes na modelagem estatística de padrões linguísticos, eles não necessariamente compartilham a reorganização estrutural e epistêmica característica da cognição humana. Conclui que a missão fundamental da educação linguística é desenvolver flexibilidade cognitiva, consciência estrutural e capacidade de atravessar limiares conceituais.

Sem saber, todos fazemos linguística. Quando, por exemplo, um falante de português ouve a frase “O rapaz que viu a moça já saiu”, ele deve compreender, sem hesitação consciente, que quem saiu foi o rapaz, não a moça. Embora esta possa parecer uma análise simples e óbvia, observe-se que, para isso, o falante precisa resolver um problema computacional complexo: identificar relações hierárquicas entre palavras, reconhecer dependências, organizando constituintes sintáticos (sujeitos e objetos) e atribuindo papéis semânticos (agentes e pacientes) em uma estrutura mental coerente. Não se trata apenas de ouvir uma sequência linear de sons; trata-se de construir internamente uma arquitetura relacional invisível. Até porque, embora, na frase, haja a sequência linear “... a moça já saiu”, tal análise em que “a moça” seria sujeito e agente do verbo “sair” geralmente nem chega a ser considerada, mesmo por uma criança, que pode fazer rápida e intuitivamente a análise não linear apropriada. Por que, então, uma frase como “O rapaz que a moça viu já saiu”, com itens lexicais também verificados na primeira frase seria mais difícil de processar do que a anterior?

Essa diferença entre as duas frases fica bem clara quando observamos a leitura através da técnica de rastreamento ocular — *eye tracking* —, que captura os padrões de fixação do olhar progressivos e regressivos no escaneamento visual de frases e imagens¹, como ilustrado nas figuras e nos vídeos 1 e 2, a seguir:

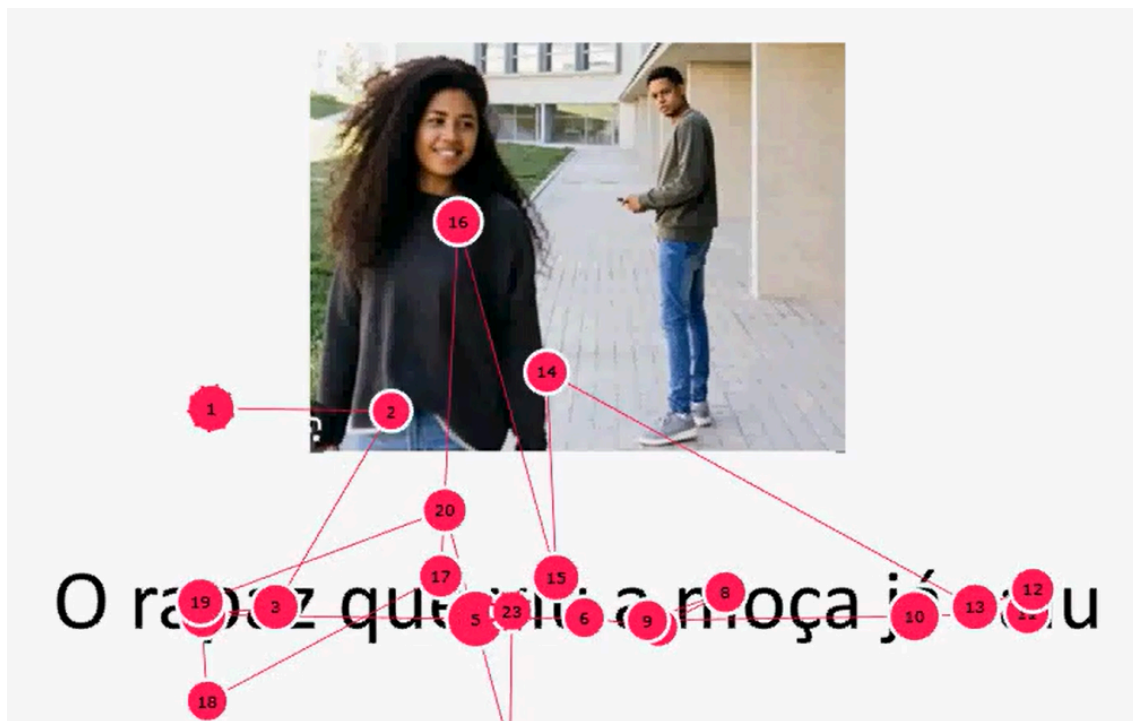


Figura 1.

Vídeo 1 (<https://nosdalinguistica.wordpress.com/wp-content/uploads/2026/07/vc3addeo1-orsujgazeplot.mp4>)

¹ Para uma apreciação mais detalhada da técnica de rastreamento ocular e suas aplicações, ver Maia (2022) e Fonseca & Maia (2022).

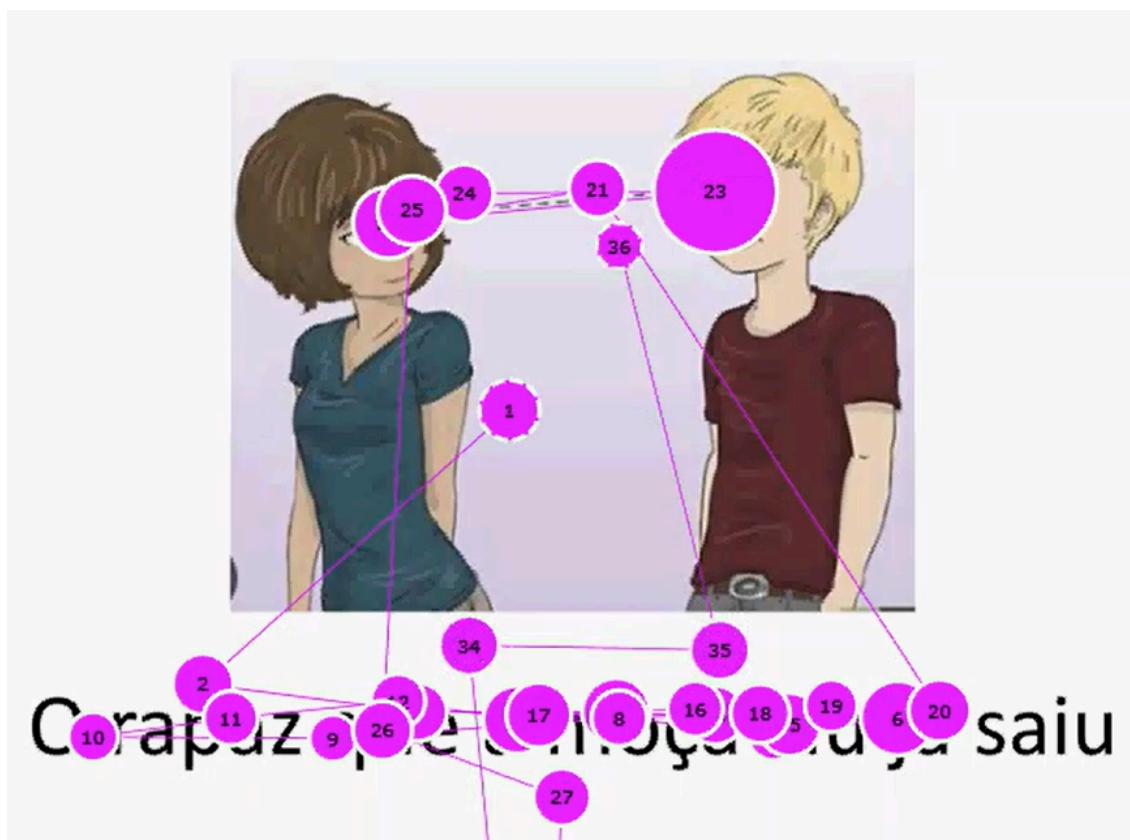


Figura 2.

Vídeo 2 (<https://nosdalinguistica.wordpress.com/wp-content/uploads/2026/07/vc3addeo2-orobjgazeplot.mp4>)

Observe que a frase e a imagem na Figura 1 são processadas com um total de 23 fixações oculares, enquanto na Figura 2 foram necessárias 36 fixações, cerca de mais de 50% de fixações (incluindo mais refixações) do que na Figura 1. Muitos estudos, em diversas línguas², têm demonstrado que orações relativas de objeto são mais custosas de processar do que suas contrapartes de sujeito. Avaliamos que essas diferenças podem estar ocorrendo em função de um **metaconceito cognitivo**³ fundamental: a **Perspectiva**. Observe que, na Figura 1,

² Por exemplo, o trabalho seminal de Traxler, Morris e Seely (2002) demonstra, através da metodologia de rastreamento ocular, que as orações relativas de objeto apresentam tempos de fixação e índices de regressão significativamente maiores do que as orações relativas de sujeito. Em português brasileiro, Corrêa, Augusto, Longchamps e Forster (2012) também indicam maior custo das relativas de objeto, discutindo esse custo adicionalmente em termos de fatores pragmáticos e contextuais. Também Oliveira (2013) e Oliveira, Maia e França (2018) exploram experimentalmente a assimetria sujeito/objeto não só em orações relativas, mas também em interrogativas-QU.

³ Nos termos de Van Rijt e Nygard (2026), um metaconceito cognitivo é um princípio de ordem superior que organiza, integra e orienta a compreensão e a aplicação de múltiplos conceitos específicos, permitindo ao aprendiz reconhecer padrões estruturais, estabelecer relações entre domínios distintos e transferir conhecimento para novas situações. Em vez de descrever um fenômeno particular, um metaconceito fornece um enquadramento cognitivo (*framework*) para interpretar e reorganizar outros conceitos. Entre outros, os autores elencam as funções sintáticas, as estruturas de constituintes e dependências e as estruturas argumentais, como metaconceitos de alto nível.

a perspectiva estabelecida a partir do sujeito “o rapaz” é mantida — o rapaz viu a moça e o rapaz saiu. Em contraste, na Figura 2, há uma mudança da perspectiva inicial — o rapaz é introduzido, mas relativizado como objeto, para em seguida ser retomado como sujeito. Como se vê na comparação dos padrões de fixação e de refixação nas duas figuras, essa mudança de perspectiva exige flexibilidade cognitiva para reorganizar os cenários e tem, portanto, um alto custo de processamento.

Outro exemplo de diferença de custo de processamento, também muito revelador, pode ser apreciado na comparação das frases 3 e 4, a seguir. Na frase 3: — “O professor pediu demissão porque o carro dele não está no estacionamento” — a conjunção *porque* parece introduzir uma causa. Mas basta um instante de reflexão para perceber que não foi o fato de que o carro não está no estacionamento que teria causado a demissão. O que a frase expressa não é causalidade, mas uma **inferência**⁴ explicativa: observando que o carro não se encontra no estacionamento, como de hábito, inferimos, com base em pressupostos, que o professor pode ter se demitido. Compare essa frase inferencial com a frase 4 — “O professor pediu demissão porque o salário dele não dava para as despesas” — que, de fato, expressa causa direta. Nossa mente inicialmente aplica um padrão interpretativo automático — *porque* = causa —, mas pode, em seguida, reorganizar a interpretação à luz de uma análise estrutural e conceitual mais refinada. Nesse breve deslocamento entre interpretação automática e reconstrução reflexiva, vemos em miniatura um fenômeno cognitivo profundo: a passagem do uso intuitivo da linguagem para a consciência de seus mecanismos. Não seria por outra razão que um bom leitor acaba por exibir maior trabalho de processamento na leitura inferencial do que na leitura de causalidade direta, como os dados qualitativos prototípicos de rastreamento ocular exemplificados abaixo sugerem.

Note que, na Figura 4, que indica causa direta, há apenas 12 fixações, todas progressivas (da esquerda para a direita), ao passo que, na Figura 3, que indica inferência, temos 19 fixações, progressivas e regressivas (da direita para a esquerda), indicando maior esforço de processamento:



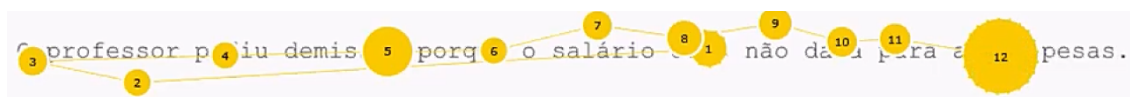
Frase 3.

“O professor pediu demissão porque o carro dele não está no estacionamento”.

Vídeo 3

(<https://nosdalinguistica.wordpress.com/wp-content/uploads/2026/07/vc3addeo3-fraseinferc3aancia.mp4>)

⁴ Para uma visão mais detalhada de estudos psicolinguísticos em português brasileiro sobre o processamento de relações inferenciais, comparando níveis de escolaridade, ver Santos, Garcia e Maia (2019).



Frases 4.

“O professor pediu demissão porque o salário dele não dava para as despesas.”

Vídeo 4 (<https://nosdalinguistica.wordpress.com/wp-content/uploads/2026/07/vc3addeo4-frasecausa.mp4>)

Note-se ainda que há uma importante mudança de perspectiva na comparação entre esses dois exemplos que ilustram causa direta e inferência – no primeiro, a causa ocorre antes do fato, enquanto, na inferência, a perspectiva analítica é concebida como posterior ao fato. Ou seja, em termos dos exemplos em discussão, na causa direta, o fato anterior de o professor não receber um pagamento adequado teria motivado a sua demissão. Já na construção inferencial, o fato de o carro do professor não se encontrar no estacionamento, como de hábito, permite projetar adiante a hipótese de que ele teria se demitido.

A metacognição de linguagem geradora de *insights* estruturais que se exercita aqui é aquilo que linguistas como Carlos Franchi já chamaram de atividade epilinguística⁵: comparando-se diferentes dados linguísticos, um conhecimento implícito, ativo e sofisticado que opera antes da consciência explícita é ativado. Falamos, compreendemos, detectamos estranhezas, arquitetamos reanálises e reperspectivações, percebemos ambiguidades e produzimos causas e inferências sem precisar saber nomear exatamente os processos envolvidos. É como andar de bicicleta: realizamos uma operação altamente complexa sem ter acesso direto aos princípios que a tornam possível. A educação linguística tradicional, porém, muitas vezes, limita-se a acrescentar rótulos técnicos sobre esse funcionamento invisível. Aprende-se o nome “oração subordinada”, “núcleo nominal”, “anáfora”, “pressuposição”, mas frequentemente sem que isso reorganize de fato a maneira como o estudante percebe a linguagem. Nesse caso, a terminologia torna-se apenas espécie de “ornamento conceitual”, não, de fato, transformação cognitiva, como se deveria almejar na educação linguística.

É aqui que emerge uma distinção que nos parece decisiva: um metaconceito só se torna educacionalmente poderoso quando reorganiza percepção, linguagem e identidade cognitiva — não quando é apenas memorizado como terminologia. Essa ideia encontra ressonância profunda no trabalho de Jan H. F. Meyer e Ray Land (2005) sobre a noção de *threshold concept*, ou conceito-limiar. Segundo eles, certos conceitos funcionam como portais cognitivos. Uma vez atravessados, transformam irreversivelmente a forma como um campo do conhecimento é percebido. Não se trata apenas de aprender algo novo, mas de passar a ver o mundo por uma nova lente. Antes de compreender o conceito de entropia, por exemplo, certos fenômenos físicos parecem desconexos; depois, uma nova inteligibilidade emerge. Antes de compreender ironia literária, o texto parece dizer uma coisa; depois, passa a dizer várias ao mesmo tempo. Antes de compreender a estrutura sintática hierárquica, frases são

⁵ Por exemplo, em “Criatividade e Gramática” (1987), Franchi argumenta que a escola tradicional privilegia excessivamente a metalinguagem gramatical em detrimento da reflexão linguística espontânea dos alunos (atividade epilinguística).

percebidas como sequências de palavras; depois, revelam-se como sistemas organizados de relações abstratas.

Essa travessia raramente é simples. Meyer e Land (2005) descrevem um estado intermediário que chamam de liminaridade: um espaço cognitivo de transição, oscilação, dúvida e aparente confusão. É o momento em que antigas intuições já não bastam, mas novas formas de compreensão ainda não se estabilizaram. O estudante fica “entre mundos”. Pode imitar a nova linguagem sem compreendê-la plenamente; pode reproduzir fórmulas; pode “fingir que entendeu”. Mas algo ainda não reorganizou internamente sua visão. Esse estado é desconfortável — e precisamente por isso fértil. A verdadeira aprendizagem profunda costuma acontecer ali, nesse espaço entre o familiar e o ainda incompreensível.

Na linguagem, um conceito-limiar fundamental é a percepção de que frases não são cadeias lineares de palavras, mas estruturas hierárquicas invisíveis. Esse ponto foi sintetizado magistralmente no argumento *Structures, Not Strings*, de Martin Everaert, Chomsky e colaboradores, em artigo de 2015: a linguagem humana não computa apenas sequências lineares; computa **relações estruturais**. Em outras palavras, o cérebro linguístico não lê apenas “palavra após palavra”; ele constrói árvores sintáticas, dependências, encaixamentos, domínios de escopo, arquiteturas pressuposicionais, relações entre posições visíveis e invisíveis. O que parece superfície linear é, cognitivamente, profundidade arquitetônica.

Quando um aluno realmente compreende a noção de estrutura, algo muda qualitativamente. Ele passa a olhar para uma frase e enxergar não apenas palavras, mas relações. Passa a perceber que ambiguidades não estão só “nas palavras”, mas em diferentes arquiteturas possíveis; que inferências emergem de encaixamentos conceituais; que a gramática não é uma lista de proibições escolares, de “certos e errados”, mas uma janela para a computação mental humana. Esse momento traria um *insight* metacognitivo estrutural: a mente toma consciência de como ela própria organiza a linguagem. É um momento de “go meta”, como, interessantemente, o descreve Mahdavi (2014). Esse deslocamento metacognitivo é infraestrutural na implementação por funções executivas, que envolvem recursos de memória e atenção, flexibilidade cognitiva, tais como ocorre na síntese pré-frontal de imagens (cf. Vyshedskiy *et al.*, 2019), entre outras atividades neurocognitivas estruturantes no lobo frontal (cf. Nelson *et al.*, 2017; Pilati & Maia, 2026).

Esse tipo de *insight* pode ser intensificado por ferramentas como o rastreamento ocular, que torna visível o invisível. Quando leitores observam seus próprios padrões de fixação, regressão ocular e hesitação diante de fenômenos da linguagem, como, por exemplo, as construções envolvendo a assimetria sujeito/objeto e relações de causa direta e inferência, discutidas aqui, percebem concretamente que compreender uma frase não é um ato instantâneo, mas um processo dinâmico de hipótese, revisão e reconstrução. A leitura deixa de parecer mágica e passa a ser observável como atividade mental inferida a partir de padrões de fixação e movimentação ocular. Surge então uma forma rara de aprendizagem: não se trata apenas de aprender um conteúdo, mas de ver a própria cognição em funcionamento. E aqui chegamos a um contraste fascinante com a inteligência artificial contemporânea.

Modelos de Inteligência Artificial — *Large Language Models* ou *bots*, como *ChatGPT*, entre vários outros — impressionam porque dominam regularidades estatísticas vastíssimas em *corpora*. Reconhecem — em altíssima velocidade — padrões, frequências de ocorrência, completam sequências, contrastam e integram conteúdos, produzindo textos

fluentíssimos, capturando associações sutis entre contextos linguísticos. Tudo em algoritmos lineares do tipo *NGram* que, item a item, vão incrementando na velocidade da luz as cadeias lineares construídas.

Mas fluência não equivale necessariamente à representação explícita de princípios estruturais profundos. Muito do sucesso desses sistemas deriva da extraordinária capacidade de modelar rapidamente distribuições lineares e contextuais de palavras. Isso pode ser computacionalmente impressionante — mas não parece ser o mesmo tipo de computação estrutural que, como procuramos discutir acima, caracteriza a linguagem humana.

Essa diferença importa educacionalmente. Um sistema pode manipular formas linguísticas sem atravessar qualquer limiar conceitual; pode gerar respostas convincentes e rápidas sem sofrer reorganização perceptiva; pode operar semanticamente sem identidade epistêmica. Já o ser humano aprende profundamente quando um conceito transforma sua forma de ver. A aprendizagem humana mais poderosa não consiste em acumular “respostas corretas”, mas em reconfigurar internamente os princípios que organizam percepção e interpretação.

Por isso, em uma época fascinada por respostas rápidas, simplificação extrema e automatização cognitiva, talvez a tarefa mais importante da educação seja, de fato, cultivar precisamente aquilo que nenhuma automatização substitui facilmente: a capacidade de atravessar limiares conceituais, suportar a liminaridade da dúvida, reperspectivar, reorganizar a percepção e tornar explícita a estrutura invisível que governa pensamento e linguagem.

Saiba mais

Em Pilati & Maia (2026), apresenta-se a metodologia da Gramaticoteca, em que se ilustram atividades de educação linguística ativas em que a metacognição é central.

Em

<https://lefufrij.wordpress.com/2022/02/01/focalizando-e-topicalizando-na-lingua-karaja-brincando-epilinguisticamente/>, Maia exemplifica atividades epilinguísticas levadas a efeito em escola da etnia Karajá.

Referências bibliográficas

CORRÊA, Letícia M. Sicuro; AUGUSTO, Marina R. A.; LONGCHAMPS, Jacqueline; FORSTER, Renê A. M. S. Referência anafórica com relativas restritivas de objeto: custo relativizado na interface gramática-pragmática. **Revista Linguística**, Rio de Janeiro, v. 8, n. 2, pp. 173–195, dez. 2012. ISSN: 1808-835X.

EVERAERT, Martin; HUYBREGTS, Marinus; CHOMSKY, Noam; BERWICK, Robert; BOLHUIS, Johan. Structures, not strings: linguistics as part of the cognitive sciences. **Trends in Cognitive Sciences**, v. 19, n. 12, pp. 729–743, 2015.

FONSECA, Aline; MAIA, Marcus. Na trilha do processamento da linguagem: o uso de rastreadores oculares na análise de dados linguísticos. *In*: OLIVEIRA, Cândido Samuel Fonseca de; SÁ, Thais de. (Org.). **Métodos Experimentais em Psicolinguística**. 1 ed. São Paulo: Pá de Palavra, 2022. v. 1, pp. 55–74.

FRANCHI, Carlos. Criatividade e gramática. **Trabalhos em Linguística Aplicada**, Campinas, v. 9, pp. 5–45, 1987.

MAHDAVI, M. An overview: Metacognition in education. **International Journal of Multidisciplinary Current Research**, v. 2, n. 3, pp. 529–535, mai.-jun. 2014.

MAIA, M. A. R. EYE TRACKING SENTENCES IN LANGUAGE EDUCATION. **Diacrítica**, v. 36, pp. 6–36, 2022.

MEYER, J. H. F.; Land, R. Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. **Higher Education**, v. 49, n. 3, pp. 373–388, 2005.

NELSON, Matthew J. *et al.* Neurophysiological dynamics of phrase-structure building during sentence processing. **Proceedings of the National Academy of Sciences**, v. 114, n. 18, pp. E3669–E3678, 2017.

OLIVEIRA, Fernando Lúcio de. **O processamento da assimetria sujeito-objeto em construções do tipo QU no português brasileiro**: interrogativas e relativas. Dissertação (Mestrado em Linguística), Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.

OLIVEIRA, F. L.; MAIA, M.; FRANÇA, A. I. Processamento de interrogativas-QU em Português Brasileiro: evidências de eyetracking e EEG. **Letras de Hoje**, v. 53, n. 1, pp. 24–38, 2018.

PILATI, E.; MAIA, M. Cognição, gramática e simplexidade: uma argumentação sobre a eficácia da proposta gerativista da gramaticoteca para a educação básica com base nas neurociências e na psicolinguística. **Gragoatá**, v. 31, n. 70, e70606, 2026.

SANTOS, Sabrina Lopes; GARCIA, Daniela; MAIA, Marcus. Relações lógico-inferenciais no uso dos conectivos *porque* e *já que*: Um estudo comparativo entre Ensino Fundamental e Ensino Superior. **Ilha Do Desterro**, v. 72, pp. 123–138, 2019.

TRAXLER, M. J.; MORRIS, R. K.; SEELY, R. E. Processing Subject and Object Relative Clauses: Evidence from Eye Movements. **Journal of Memory and Language**, v. 47, n. 1, pp. 69–90, 2002.

VAN RIJT, Jimmy; NYGÅRD, Mari. Relationships between linguistic theory and L1 grammar education revisited: a Delphi study on the key metaconcepts in the syntax-semantics interface. **Open Linguistics**, v. 12, n. 1, pp. 1–32, 2026.

VYSHEDSKIY, Andrey. Language evolution to revolution: the leap from rich-vocabulary non-recursive communication system to recursive language 70,000 years ago was associated with acquisition of a novel component of imagination, called Prefrontal Synthesis, enabled by a mutation that slowed down the prefrontal cortex maturation simultaneously in two or more children – the Romulus and Remus hypothesis. **Research Ideas and Outcomes**, v. 5, n. 4, e38546, 2019.