

LEITURABILIDADE E EXPERIÊNCIA ESCOLAR: A COMPLEXIDADE SINTÁTICA EM EVIDÊNCIA¹

**READABILITY AND READING EXPERIENCE:
SYNTACTIC COMPLEXITY IN EVIDENCE**

KÁTIA NAZARETH MOURA DE ABREU
Universidade do Estado do Rio de Janeiro
katia.abreu@uerj.br

DANIELA CID DE GARCIA
Universidade Federal do Rio de Janeiro
daniela@letras.ufrj.br

SABRINA LOPES DOS SANTOS
Universidade Federal do Rio de Janeiro
sabrinalopes@letras.ufrj.br

Resultados do PISA de 2022 (OCDE 2023) revelam que cerca de 50% dos estudantes brasileiros examinados não atingem níveis mais altos de proficiência em leitura, em que se espera uma compreensão detalhada de textos longos ou complexos (BRASIL 2023). Fatores específicos sobre o processamento da linguagem podem estar entre os preditores dessa dificuldade. Neste trabalho, focamos na relação entre tamanho do texto e complexidade sintática como fatores que impactam o processamento da leitura. Realizamos um experimento com rastreador ocular em que manipulamos a extensão (com e sem elementos intercalados) e a estrutura (orações relativas de sujeito e de objeto) de períodos, verificando os efeitos em grupos de participantes com diferentes graus de escolaridade: Ensino Fundamental, Médio e Superior. Nossa hipótese foi a de que a introdução de um elemento intercalado na estrutura argumental poderia impactar a leiturabilidade do período, por dificultar o fechamento da unidade sintática. Os resultados indicaram efeito principal dos fatores estrutura e escolaridade. Concluímos ser fundamental que as práticas educacionais voltadas à competência leitora salientem o papel da hierarquia estrutural da língua.

Palavras-chave: Leiturabilidade, complexidade sintática, rastreamento ocular.

¹ Este estudo foi autorizado pelo Comitê de Ética em Pesquisa com Seres Humanos da Universidade do Estado do Rio de Janeiro, em parecer número 4.814.835, por meio de processo de submissão de toda a documentação necessária.

Results from the 2022 PISA (OECD 2023) reveal that around 50% of Brazilian students examined do not reach advanced levels of reading proficiency, where a detailed understanding of long or complex texts is expected (BRASIL 2023). Specific factors related to language processing may contribute to this difficulty. In this study, we focus on the relationship between text length and syntactic complexity as factors that affect reading processing. We conducted an eye-tracking experiment in which we manipulated the length (with and without interspersed elements) and structure (subject and object relative clauses) of sentences, observing the effects on groups of participants with different levels of education: elementary, high school, and higher education. Our hypothesis was that the introduction of an interspersed element within the argument structure could affect sentence readability by making it more difficult to close the syntactic unit. The results indicated a significant effect of both sentence structure and education level. We conclude that educational practices aimed at improving reading competence should emphasize the importance of understanding the structural hierarchy of language.

Keywords: Readability, syntactic complexity, eye tracking.

Recibido: 16 agosto 2024 Aceptado: 15 octubre 2024

1. INTRODUÇÃO

A leitura desempenha um papel crucial no desenvolvimento cognitivo e na formação de cidadãos críticos e informados. No entanto, índices recentes têm revelado um cenário preocupante em relação ao desempenho leitor de jovens, tanto no Brasil quanto no cenário global. Dados provenientes de avaliações internacionais e nacionais, como o Programa Internacional de Avaliação de Alunos (PISA) e o Sistema de Avaliação da Educação Básica (SAEB), mostram que muitos estudantes apresentam dificuldades significativas em compreender e interpretar textos. Os fatores responsáveis por esse desempenho abaixo da média podem ser de diversas naturezas. Este artigo tem como objetivo explorar fatores psicolinguísticos que podem impactar a habilidade em leitura.

Durante muito tempo, o foco da educação esteve centrado na alfabetização, com debates acalorados sobre métodos de ensino, como as chamadas "guerras da leitura" e a ênfase na consciência fonêmica. Esses debates giraram em torno da ideia de que a capacidade de decifrar o código alfabético era a chave para a leitura eficaz. A perspectiva do *Simple View of Reading* (Gough e Tunmer 1986) sustenta que, uma vez decifrado o código, o processamento da leitura seria tão natural quanto o processamento da língua falada. No entanto, a realidade mostra que muitos jovens enfrentam dificuldades significativas ao lidar com textos mais extensos e complexos, indicando que a decifração do código, por si só, não é suficiente para desenvolver a fluência leitora. Essas dificuldades revelam a necessidade de se olhar além da simples decifração e considerar outros fatores que contribuem para uma leitura dinâmica e eficiente. Em particular, este artigo examina a importância de um fator relacionado ao processamento sintático e sua relação com o tamanho do texto, explorando como a compreensão da estrutura das frases e a capacidade de lidar com construções sintáticas complexas podem influenciar o desempenho leitor. Ao analisar esses elementos, busca-se compreender melhor os desafios enfrentados pelos jovens leitores, de modo a contribuir com práticas que possam melhorar a fluência e a compreensão.

Exploramos um efeito do processamento sintático – a assimetria sujeito-objeto – incluindo os fatores extensão e escolaridade. Pressupomos que, ao alongar o constituinte, é possível que o

efeito sintático se dilua com maior demanda da memória de trabalho, sobretudo em leitores menos experientes, que devem imprimir uma leitura mais linear. Espera-se, com isso, estabelecer uma relação entre os estudos psicolinguísticos de processamento de sentenças e as métricas de leiturabilidade, ferramentas que vêm sendo desenvolvidas há algumas décadas para avaliar a facilidade com que um texto pode ser compreendido.

2. LEITURABILIDADE

Abreu e Lima (2022) sugerem uma conversa entre a literatura sobre leiturabilidade e o processamento sintático. O termo leiturabilidade se refere ao “quão facilmente os materiais escritos podem ser lidos e entendidos. Isso depende de vários fatores, incluindo o tamanho médio dos períodos, o número de palavras novas contidas e a complexidade gramatical” (Richards *et al.* 1992). Uma maneira de aprimorar as habilidades de leitura é selecionar textos que apresentem um nível de desafio apropriado, evitando aqueles que sejam muito complexos ou incompreensíveis. Alguns autores argumentam que, quando leitores são expostos a textos que estão dentro de um nível de dificuldade adequado, suas habilidades de leitura tendem a melhorar (Allington 2005; Stanovich 1985; Wolfe *et al.* 1998). Sendo assim, medidas de leiturabilidade começaram a se desenvolver, desde os anos 40, nessa busca de adequar a dificuldade de textos ao nível do leitor.

Essa ideia de se partir de um objeto mais manuseável (Gallego e Chomsky 2020), entender o sistema antes de partir para desafios mais complexos se alinha com propostas recentes de alguns pesquisadores das áreas de Teoria Gramatical e de Psicolinguística que têm promovido uma maior interação entre esses campos da ciência da linguagem e as ciências da educação no Brasil (Maia 2018, 2019a; 2019b; Pilati 2017; Pilati *et al.* 2019; Kenedy 2013; 2021). Esses autores têm destacado o período como uma unidade de bastante relevância na leitura. Uma leitura que identifica e hierarquiza os constituintes sintáticos tende a resultar em melhor recuperação das informações presentes no texto, sobretudo os de maior extensão. Quando falamos em identificar, não estamos nos referindo aqui à classificação de constituintes, mas ao conhecimento intuitivo sobre onde começa e onde termina um período ou sobre quais são os argumentos em uma oração –sobre o que é essencial e o que é satélite. Esse conhecimento deve ter um impacto crucial em como direcionar o olhar para quais informações buscar, por exemplo, em períodos muito longos. Aqui estamos falando em estabelecer conscientemente –metacognitivamente– uma estratégia de leitura dentro do período, aplicando uma leitura que poderia ser considerada *topdown*, para buscar os termos argumentais, descartando os elementos intervenientes, e em seguida procedendo com uma estratégia de combinação desses constituintes (Garcia *et al.* 2020).

Ao estruturar e hierarquizar informação dentro do período, é possível administrar recursos atencionais e de memória para uma leitura mais acurada e efetiva. Hyöna *et al.* (2002) investigaram a relação entre memória e leitura no nível textual. De acordo com os resultados apresentados pelas autoras, leitores que estruturam os diferentes assuntos de um texto em seções apresentam maior capacidade de lembrar o conteúdo. Essa maior capacidade é encontrada em relação a leitores que adotam uma estratégia de leitura linear. Maia (2019a) argumenta que a estruturação sintática desempenha um papel fundamental para a compreensão de sentenças por via da leitura, e apresenta evidências de que estudantes que produzem uma leitura linear tendem a se desengajar da leitura, o que tem impacto negativo na compreensão.

Essa redução do objeto não deve ser confundida com as propostas de simplificação da linguagem que vêm sendo sugeridas recentemente sob um argumento falacioso de inclusão. Se o objetivo da educação formal é alcançar os mais altos graus de compreensão textual, precisamos dedicar nossa atenção em melhorar a qualidade de ensino dos nossos cidadãos e não moldar a escrita de forma que os textos contem com sentenças simples ao invés de unidades complexas. A discussão em curso sobre linguagem simples Projeto de Lei 6256/2019 começou com uma proposta de linguagem cidadã em que:

Linguagem Cidadã é uma forma de comunicação usada para *transmitir informações de maneira simples, objetiva e inclusiva*. Uma comunicação é simples se o público-alvo *consegue entender de forma fácil e rápida, sem precisar reler o texto ou pedir para outra pessoa repetir a mesma mensagem*.

(Projeto de Lei 6256/2019: 2. Grifos nossos)

A ideia de tornar as informações apresentadas pelo serviço público algo simples, objetivo e inclusivo devem ser vistas como um avanço no sentido de evitar uma escrita prolixa e cheia de jargões desconhecidos da população em geral. Infelizmente, o que temos visto são ações prescritivistas sem um embasamento linguístico que sustente a argumentação ficando apenas o cunho moralista de uma percepção de inclusão quando na verdade o acesso à língua fica mais restrito.

Seguimos a ideia de Maia e Rodrigues (2024), que destacam a importância de se conceber a relação entre linguagem e leitura a partir do modelo de arquitetura da linguagem explicitado em Chomsky (2005). Esse modelo estabelece que há três fatores na arquitetura da linguagem. O primeiro fator seria o que é determinado pela genética, os princípios universais e inatos da gramática humana. O segundo seria a experiência, as influências e variações provenientes do ambiente linguístico, que, pela interação com o primeiro fator, ativaria os princípios inatos.

Ninguém ensina a criança fazer estrutura; trata-se de um universal da cognição linguística humana, que constrói binariamente estruturas potencialmente complexas, em relações hierárquicas que, ao serem computadas, permitem a representação mental, a compreensão e a produção da linguagem.

(Maia e Rodrigues, 2024:26)

O terceiro fator diz respeito às métricas de economia e simplicidade, que são fundamentais em várias áreas, influenciando a forma como sistemas e organismos funcionam de maneira otimizada. É um fator externo à linguagem, mas que impõe restrições ao desempenho linguístico (Kenedy 2021). Maia e Rodrigues (2024) destacam a proposta de Trotzke *et al.* (2013) de conceber o processamento da linguagem como terceiro fator, já que modelos de processamento, justamente por restrições impostas pelo nível implementacional, baseiam-se em princípios de economia e simplicidade. Sendo assim, embora a linguagem seja inerentemente recursiva, complexa e capaz de gerar sentenças infinitas, seu desempenho é restringido no processamento, como, por exemplo, por limitações de memória e atenção.

A escrita permitiu que as sentenças se tornassem mais longas, exigindo que nossa memória e atenção lidem com essas sequências complexas. Para enfrentar essa sobrecarga de memória, nossa espécie desenvolveu a capacidade de agrupar informações, um processo conhecido como *chunking* (Miller 1956, Kato 2000, Kleiman 2013). O conceito de *chunking* na psicologia refere-se ao processo de agrupar informações em unidades maiores e mais significativas, chamadas de

chunks ou blocos, para facilitar a memorização e o processamento cognitivo. Esse conceito é especialmente relevante na memória de curto prazo, onde a capacidade de armazenar informações é limitada. À medida que as orações se expandem e se tornam mais complexas, elas não crescem de forma aleatória e linear, mas respeitam uma estrutura básica. No caso do português, reconhecer a estrutura básica sujeito-verbo-objeto (SVO) e buscar essas unidades significativas deve funcionar como uma estratégia útil para percorrer períodos muito longos sem desengajar (Kleiman 2013; Maia 2019a).

3. ASSIMETRIA SUJEITO-OBJETO

Segundo a literatura linguística que trata do processamento de períodos, (Gibson 1998; Warren e Gibson 2002; Traxler *et al.* 2002; Staub 2010; Oliveira 2013; Forster e Corrêa 2017; Oliveira, Maia e França 2018; Guedes 2021; *inter alia*), os períodos que contêm orações relativas extraídas de objeto (OROs) são mais difíceis de processar se comparados a períodos que contêm orações relativas extraídas de sujeito (ORSs). Gibson (1998); Warren e Gibson (2002) atribui a dificuldade da extração de objeto a operações de memória de trabalho, pois o SN da oração principal precisa ser mantido na memória até que o verbo da oração relativa seja encontrado; assim, quanto mais distante o SN estiver da Oração Relativa, mais custosa será a recuperação e, conseqüentemente, maior dificuldade de compreensão leitora. O custo da integração de dois elementos depende da distância entre os dois. Sendo assim, a dependência de longa distância conduz a mais dificuldades para que ocorra essa integração. Se temos uma cadeia estruturalmente menor (extraída de sujeito) menor a dificuldade de integração. Por outro lado, se temos uma cadeia estruturalmente maior, maior a dificuldade de integração.

Traxler e colegas, em artigo de 2002, relatam a pesquisa que conduziram sobre o processamento de períodos com orações relativas de sujeito e de objeto por meio da técnica de rastreamento ocular. Eles realizaram três experimentos. O primeiro, indicou que períodos contendo orações relativas de objeto eram mais difíceis de processar do que períodos contendo orações relativas de sujeito no decurso da oração relativa e do verbo principal. Esses resultados acompanham o padrão que é observado quando essa comparação é feita, ou seja, orações relativas de sujeito X orações relativas de objeto.

O segundo manipulou a plausibilidade do sujeito sentencial e do substantivo dentro da oração relativa como o agente da ação representada pelo verbo na oração relativa. Os participantes leram períodos como: 1 (a) O policial *que prendeu o ladrão* era conhecido por carregar uma faca e 1 (b) O ladrão *que o policial prendeu* era conhecido por carregar uma faca; nesse caso, é plausível que um policial prenda um ladrão, mas não o contrário. Os leitores experimentaram maior dificuldade durante o processamento de períodos contendo orações relativas de objeto do que em períodos com orações relativas de sujeito.

O terceiro, manipulou a animacidade do sujeito sentencial e do substantivo dentro da oração relativa. Esse experimento demonstrou que a dificuldade associada às orações relativas de objeto foi bastante reduzida quando o sujeito sentencial era inanimado. Os autores mostram que essa diferença de dificuldade tem sido atribuída de várias maneiras às limitações da memória de trabalho, a fatores sintáticos e a mudança de perspectiva.

Propostas baseadas em memória atribuem a diferença na dificuldade entre as relativas de sujeito e de objeto ao aumento das demandas da memória de trabalho impostas por construções

relativas ao objeto. Uma abordagem baseada em sintaxe atribui a diferença entre as construções relativas de sujeito e relativas de objeto a diferenças nos papéis que o sujeito sentencial exerce em períodos contendo orações relativas de sujeito e relativas de objeto. A dificuldade de atribuir simultaneamente –em grande parte– papéis sintáticos e temáticos incompatíveis a um único constituinte contribui para o aumento da dificuldade de construções relativas de objeto.

Quanto à abordagem de mudança de perspectiva, essa prevê que os leitores devem mudar sua perspectiva quando encontram o sujeito da oração relativa e novamente quando eles voltam a processar a oração principal. Em período com oração relativa de sujeito, há apenas um constituinte servindo como sujeito ao longo do período e a mudança de perspectiva não é necessária. Em período com oração relativa de objeto, o sujeito da oração principal é diferente do sujeito da oração relativa e a mudança de perspectiva é necessária.

Staub (2010) apresenta duas formas gerais de abordagem para a dificuldade de processamento das orações relativas extraídas de sujeito: abordagem baseada em memória –sobre a qual tratamos acima– e abordagem baseada em expectativa. A abordagem baseada em expectativa prevê que os compreendedores antecipam prováveis continuções de períodos. Conforme essa perspectiva, as orações relativas extraídas de objeto (OROs) são difíceis principalmente por serem incomuns em comparação com as orações relativas de sujeito (ORSs). Vejamos o exemplo (1):

1. a. O repórter que atacou o senador admitiu o erro. (ORS)
- b. O repórter que o senador atacou admitiu o erro. (ORO)

Depois de encontrado o pronome relativo ‘que’ em 1 (a), uma oração relativa de sujeito é uma continuação muito mais provável do que uma ORO e, portanto, uma ORS é preferencialmente esperada pelo leitor ou ouvinte. Uma ORO é, portanto, surpreendente e indutora de dificuldade. O autor pondera que “Há, de fato, evidências substanciais que tanto a recuperação de memória [...] quanto a previsão estrutural [...] estão envolvidas no processamento sintático” (Staub 2010: 72).

Oliveira (2013) em sua pesquisa intitulada “O processamento da assimetria sujeito-objeto em construções do tipo QU no Português Brasileiro: interrogativas e relativas” sustentou que a extração de sujeito descreve uma cadeia linearmente e estruturalmente menor quando comparada à extração de objeto. O autor postulou a existência de processamento facilitado para relativas de sujeito em relação às de objeto, tomando como base o Princípio da Cadeia Mínima–De Vincenzi, 1996 e a Hipótese do Processamento de Interrogativas–Hofmeister *et al*, 2007. Ele defendeu que o fator predominante no processamento de orações relativas é o custo cognitivo oferecido à memória de trabalho.

Em artigo sobre assimetria no custo de processamento de relativas de sujeito e de objeto, Forster e Corrêa (2017) investigaram, no Português Brasileiro, a possibilidade de um efeito de informação contextual no processamento de sentenças contendo orações relativas de sujeito e de objeto, considerando os possíveis impactos dessa integração para a assimetria de custo entre ORSs e OROs. O objetivo era verificar até que ponto um contexto conduzido de forma inadequada poderia substituir a assimetria entre orações relativas de sujeito e objeto. A técnica de rastreamento ocular foi utilizada em uma tarefa de compreensão leitora que teve como foco a comparação entre relativas de sujeito e relativas de objeto em dois contextos. Os resultados sugeriram que orações relativas de objeto (ROs) têm maior custo quando comparadas a relativas de sujeito (RSs), independentemente do contexto discursivo no qual estão inseridas.

Em artigo de 2018, Oliveira, Maia, França investigaram a Assimetria Sujeito-Objeto em interrogativas-QU em Português Brasileiro e utilizaram as técnicas de rastreamento ocular e eletroencefalografia como ferramentas de diagnóstico do processamento linguístico em tempo real com o objetivo de buscar evidências sobre a atuação do mecanismo humano de processamento de frases durante o processamento das funções de sujeito e de objeto nesses tipos de sentenças. Assumem o postulado de Oliveira (2017) sobre a Assimetria Sujeito-Objeto (ASO) em que “Dados dois constituintes, em duas frases distintas, um com função de sujeito e outro com função de objeto, a ASO se instancia se uma das frases for agramatical ou mais difícil de processar na mente.”

Os autores não estão tratando da questão de gramaticalidade, mas sim com a questão da dificuldade de processamento, já comentada anteriormente, que envolve custo cognitivo podendo ser maior ou menor na comparação de uma frase com outra. Em suas conclusões, apontam que a literatura psicolinguística sobre a Assimetria Sujeito-Objeto “demonstra uma clara tendência de que frases cujo constituinte relevante ao estudo desempenha função de sujeito são mais fáceis de processar do que suas contrapartes de objeto” (Oliveira *et al.* 2018: 36).

Em trabalho recente, Guedes (2021) abordou o tema da Assimetria Sujeito-Objeto no processamento de orações relativas por bilíngues português-inglês e investigou se a animacidade do referente interferiria no processamento de orações relativas em inglês por falantes não nativos brasileiros que têm o português como idioma materno e o inglês como segunda língua. Ele observou o desaparecimento da assimetria de processamento entre orações relativas de sujeito e objeto no caso dos falantes de nível avançado, mas não observou o mesmo com os sujeitos de nível intermediário.

Guedes (2021) destacou o trabalho de Mak *et al.*, (2002) no qual descreveram um estudo de *corpus* constituído de textos jornalísticos em alemão e holandês que indica a importância da animacidade na distribuição de ORSs e OROs nas duas línguas. Destacou, também, o trabalho de Traxler *et al.* (2002) já descrito anteriormente, em que o terceiro experimento demonstrou que, uma vez que a animacidade do referente das orações relativas é levada em consideração, a assimetria de processamento de ORSs e OROs é grandemente reduzida.

4. MÉTODO

Neste trabalho, utilizamos a técnica de rastreamento ocular para verificar o impacto de elementos intercalados no processamento de períodos com orações relativas de sujeito e de objeto.

O rastreamento ocular, ou eye-tracking, consiste em uma técnica em que se monitoram, através de um equipamento, as fixações oculares e os movimentos sacádicos progressivos e regressivos, na leitura de frases ou na visualização de imagens”

(Maia 2015: 60).

Com essa técnica, é possível analisar o comportamento do aluno-leitor durante o ato de ler. Dessa forma, o aspecto processual ganha relevo, pois uma questão que se impõe em estudos sobre a leitura com foco no contexto escolar é como os alunos-leitores estão lendo: se do início ao fim do material de leitura, se do início do material de leitura até uma parte com posterior desengaja-

mento da atividade, ou se leem partes distintas do material de leitura. Desse modo, é possível capturar “subprocessos de leitura de modo direto, no momento em que estes ocorrem [...]” (Maia 2019b: 103).

O experimento tem como objetivo geral analisar o processo de leitura e de compreensão leitora de alunos-leitores em relação direta com um aspecto da leiturabilidade: a complexidade gramatical da língua. A complexidade gramatical está sendo tomada aqui pelo processamento da assimetria sujeito-objeto em períodos com orações relativas. Sendo assim, temos como um dos fatores experimentais o tipo de oração relativa (Sujeito ou Objeto). Além disso, manipulamos o tamanho dos períodos intercalando um sintagma adjetival, que alonga o sintagma nominal sujeito, aumentando o material que deve ficar na memória até que a unidade sintática progrida, com verbo e objeto (v. Figura 1).

SUJEITO	VERBO	COMPLEMENTO
O empresário	devolveu	um pedido

SUJEITO	VERBO	COMPLEMENTO
O empresário que maltratou o funcionário	devolveu	um pedido

no último minuto.

SUJEITO	VERBO
O empresário grosso e estressado que maltratou o funcionário por causa de fofoca	devolveu ...

Figura 1. Exemplo do material

As variáveis independentes testadas foram, portanto (i) extensão do período, que se dividiu entre período longo e período curto (com e sem intercalação) e (ii) estrutura oracional do período, que pode ser com relativa de sujeito e com relativa de objeto. Essas variáveis constituíram um *design within subjects* 2x2 (Tabela 1) gerando as condições LS - Período Longo com relativa de Sujeito, LO - Período Longo com relativa de Objeto, CS - Período Curto com relativa de Sujeito, CO - Período Curto com relativa de Objeto. Quanto às variáveis dependentes temos 1) o custo de processamento, materializado no tempo total das fixações (*Total fixation duration*) e no número total de fixações (*Fixation Count*) na oração relativa, e 2) o nível de compreensão, instanciado pelo índice de acerto a uma pergunta de compreensão após a leitura do período.

Extensão/Estrutura	Relativa de Sujeito	Relativa de Objeto
Longo	LS	LO
Curto	CS	CO

Tabela 1. Design do experimento

Consideramos, ainda, o fator escolaridade, considerando três grupos (Ensino Superior, Ensino Médio e Ensino Fundamental), constituindo um *design between subject* 2x2x3. A tarefa

consistia em ler os períodos e responder a uma pergunta de compreensão, conforme o exemplo 2 abaixo. A resposta se dava com a fixação do olhar em uma das opções.

2. O empresário grosso e estressado que maltratou o funcionário por causa de fofoca devolveu um pedido no último minuto.

Quem devolveu um pedido?

O funcionário

O empresário

Nossa hipótese leva em consideração três pressupostos: (1) de que sentenças com orações relativas de objeto são mais difíceis de processar, conforme atestado na literatura sobre o tema (Forster e Corrêa 2017; Oliveira *et al.* 2018; Guedes 2021); (2) de que a intercalação de elementos na estrutura sintática dificulta o processamento e a compreensão de sentenças (Kleiman 2013) e (3) de que maior escolaridade resulta em maior habilidade leitora. Sendo assim, esperamos que alunos com menor escolaridade e, portanto, menos experiência leitora, tenham mais dificuldade com as sentenças com mais intercalações.

A previsão do estudo de rastreamento é, portanto, a de que estudantes do ensino superior devem ler as sentenças integralmente e reproduzir as expectativas de assimetria sujeito-objeto da literatura Psicolinguística, independente do tamanho das sentenças. Ou seja: o processamento de sentenças com relativas de objeto deve ser mais custoso, o que deve se manifestar, nas medidas dependentes deste experimento, como maiores latências de fixação ocular nas orações relativas de objeto. Com relação às perguntas de interpretação, os estudantes do ensino superior devem apresentar altos índices de acerto em todas as condições. Considerando o pressuposto de que leitores com menos experiência são mais impactados pela extensão do material escrito, este estudo prevê que os índices de acerto à questão interpretativa devem reduzir com a escolaridade, em função do tamanho. Com relação aos dados de processamento nos níveis mais baixos de escolaridade, é possível que o efeito da assimetria fique menos proeminente nos períodos com intercalação, já que uma leitura menos estruturante tende a diluir as fronteiras sintáticas, sobretudo em sequências com maior extensão, devido à sobrecarga na memória de trabalho.

4.1. Participantes

A pesquisa foi desenvolvida em turmas de 9º ano do ensino fundamental e de 3º ano do ensino médio da rede pública estadual do Rio de Janeiro e com alunos do ensino superior, a saber: alunos do Curso de Letras da Faculdade de Formação de Professores, Universidade do Estado do Rio de Janeiro do quarto ou do quinto período, totalizando 72 participantes. Havia entre os grupos um intervalo temporal de dois anos, portanto.

Os alunos do Ensino Fundamental, do Ensino Médio e do Ensino Superior foram escolhidos de forma aleatória e demonstraram interesse em participar. Foram considerados, nesse processo, os seguintes critérios de inclusão: ausência de queixas relacionadas à visão; ausência de distúrbios

comportamentais ou cognitivos, de queixas relacionadas a dificuldades do aprendizado da leitura e da escrita ou a dificuldades de aproveitamento escolar.

Os alunos indicados apresentaram o Termo de Assentimento Livre e Esclarecido (TALE) e o Termo de Consentimento Livre e Esclarecido (TCLE) assinado por eles ou pelos pais e/ou responsáveis. Participaram também de uma triagem de leitura, que foi realizada com a finalidade de garantir não só a observância dos critérios inclusivos como também de verificar se possuíam a habilidade de leitura padrão para o referido ano escolar.

4.2. Material

O material do experimento consistiu em 16 (dezesseis) conjuntos de períodos com 4 (quatro) condições cada, digitados em fonte Courier New, considerando a extensão dos períodos e a estrutura gramatical (Tabela 2). Quanto à extensão, nesse experimento, um período Longo possui cinco segmentos e um período Curto possui quatro segmentos (cf. exemplos). Quanto à estrutura gramatical, o Longo é formado por um sintagma nominal complexo SN+SA (segmento1), uma oração relativa de sujeito ou de objeto (segmento 2), um sintagma preposicional (segmento3), uma oração principal (segmento 4) e um sintagma adverbial (segmento 5), assim: SN + SA + ORs ou ORo + SPrep + OP + SAdv. O Curto é formado por um sintagma nominal simples SN, uma oração relativa de sujeito ou de objeto, uma oração principal e um sintagma adverbial, assim: SN + ORs ou ORo + OP + SAdv. Como observamos no conjunto de estímulos, exemplificados a seguir:

Longa de Sujeito (LS)				
Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5
O empresário grosso e estressado	que maltratou o funcionário	por causa de fofoca	devolveu um pedido	no último minuto.
Longa de Objeto (LO)				
Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5
O funcionário responsável e alegre	que o empresário maltratou	por causa de fofoca	revelou o segredo	na hora da saída.
Curta de Sujeito (CS)				
Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5
O empresário	que maltratou o funcionário	devolveu um pedido	no último minuto.	-----
Curta de Objeto (CO)				
Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 5
O funcionário	que o empresário maltratou	revelou o segredo	na hora da saída.	-----

Tabela 2. Exemplo de materiais experimentais.

Além dos dezesseis conjuntos experimentais, foram construídos e incluídos no experimento trinta e dois períodos distratores e houve a distribuição dos materiais em quadrado latino em quatro versões. O material experimental foi implementado por meio do programa Tobii Studio 3.2, que monitora o rastreador ocular Tobii PRO 120 Hz.

4.3. Procedimentos

Cada participante automonitorou a leitura dos períodos apresentados na tela de um notebook com tela de 15.6 polegadas e com resolução 1366 X 768 enquanto suas fixações oculares eram registradas. Cada sessão foi precedida de uma etapa de calibração na qual foi solicitado ao participante que fixasse o olhar e acompanhasse os pontos que se movimentavam na tela. Em seguida, houve uma etapa de treinamento na qual a experimentadora explicou a tarefa e observou os participantes na leitura de dois períodos de prática. Nesse momento, o experimentador salientou que a leitura deveria ser feita com bastante rapidez, mas pretendendo o entendimento, pois haveria uma pergunta de compreensão ao final do período. Com a percepção de que a tarefa estava entendida pelo participante, o experimentador ativava uma das versões do experimento e se retirava da sala. A previsão era de que a tarefa completa durasse 20 minutos.

5. RESULTADOS

5.1. Acurácia nas respostas de compreensão

Os dados relativos a erros e acertos nas respostas às perguntas de compreensão (Tabela 3) foram submetidos, primeiramente, a um teste qui-quadrado para verificar se havia diferença na proporção de acertos em relação a erros entre os grupos. O teste indicou que havia diferença significativa (X-squared 14.225, df=2, p=0.001).

Grupo	CO		CS		LO		LS		Total
	certo	erro	certo	erro	certo	erro	certo	erro	
EF	69	17	57	27	65	18	61	27	341
EM	60	27	70	21	67	22	66	23	356
ES	67	15	80	8	74	10	69	12	335
Total	196	59	207	56	206	50	196	62	1032

Tabela 3. Acertos e erros às respostas de compreensão.

Para verificar os efeitos em relação às condições experimentais, submetemos os resultados a um teste de regressão logística binomial, que apontou diferença significativa em quatro comparações: 1) entre Ensino Superior e Ensino Fundamental nos períodos Curtos com Relativa de Sujeito (*Odds Ratios*=4.74, CI=1.09-11.86, p<0.001), indicando que o Ensino Fundamental errou significativamente mais; 2) entre Ensino Superior e Ensino Médio nos períodos Curtos com

Relativa de Sujeito ($Odds\ Ratios=2.94$, $CI=1.26-7.49$; $p=0.016$), indicando que o Ensino Médio também errou significativamente mais; 3) entre Ensino Superior e Ensino Fundamental nos períodos Longos com Relativas de Sujeito ($Odds\ Ratios=0.39$, $CI=.18-0.83$, $p=0,016$), indicando que o Ensino Fundamental errou significativamente mais; 4) entre o Ensino Superior e o Ensino Médio nos períodos Longos com Relativa de Objeto ($Odds\ Ratios=0.41$, $CI=0.17-0.91$, $p=0.033$), indicando que o Ensino Médio errou significativamente mais.

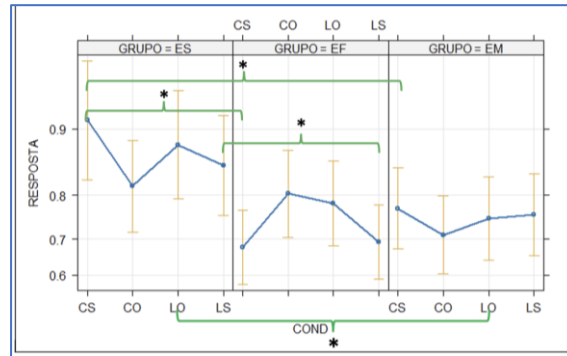


Gráfico 1. Efeitos para os índices de acurácia.

Não houve diferenças significativas nos índices de acurácia dentro de cada grupo. Ou seja, a acurácia das respostas, dentro de cada grupo, não dependeu de se o período era Longo ou Curto ou de se tinha oração relativa de Sujeito ou de Objeto. Em linhas gerais (Gráfico 2), o desempenho dos estudantes de Ensino Superior foi significativamente melhor que o de estudantes do Ensino Médio ($X\text{-squared}=9.83$, $df=1$, $p=0.002$) e que o de estudantes do Ensino Fundamental ($X\text{-squared}=11.39$, $df=1$, $p=0,008$), que não diferiram entre si.

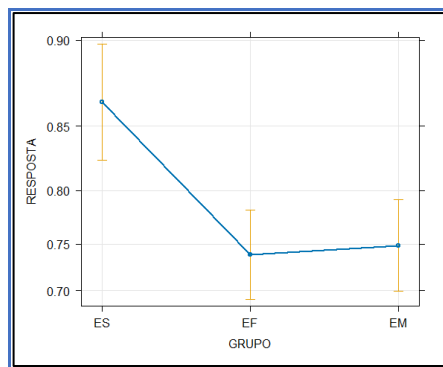


Gráfico 2. Acurácia das perguntas de compreensão.

5.2. Rastreamento Ocular: *Total Fixation Duration*

As médias relativas à duração total de fixação (Tabela 4 - *Total Fixation Duration*) nas orações relativas foram analisados através de uma ANOVA por sujeitos e apontam que houve

efeito principal significativo de grau de escolaridade ($F(2,190) = 3,99$ $p < 0,02$) e de estrutura ($F(1,95) = 10,1$ $p < 0,002$).

Nas comparações pareadas, feitas por meio de um teste de Tukey, tanto o Ensino Fundamental quanto o Ensino Superior reproduziram o efeito clássico de assimetria sujeito-objeto nos períodos curtos. Ou seja, na condição curta, as relativas de objeto foram mais difíceis de processar (ESCOvsESCS: $p=0,003$ $t(95) = 2,20$ $p < 0,05$ / EFCOvsEFCS: $t(95) = 2,44$ $p < 0,0166$). Não houve diferença significativa entre relativas de sujeito e objeto nas demais comparações pareadas, tanto nas longas do ensino Fundamental e Superior quanto nas curtas e longas do Ensino Médio.

Escolaridade	ES	ES	ES	ES	EM	EM	EM	EM	EF	EF	EF	EF
Extensão	C	C	L	L	C	C	L	L	C	C	L	L
Estrutura	O	S	O	S	O	S	O	S	O	S	O	S
Média	3,27	2,74	2,29	2,27	2,80	2,66	2,07	1,84	3,21	2,61	2,56	2,16

Tabela 4. Médias de Duração Total de Fixação na área da OR em segundos

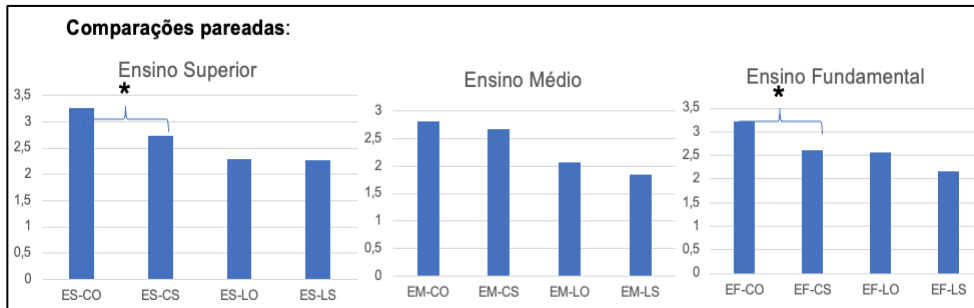


Gráfico 1. Comparativo de médias de Duração Total de Fixação entre os grupos.

5.3. Rastreamento Ocular: *Fixation Count*

Os dados relativos à contagem de fixações (*Fixation Count*) nas orações relativas foram analisados através de uma ANOVA por sujeitos e apontam que houve efeito principal significativo de estrutura ($F(1,95) = 5,81$ $p < 0,018$).

Nas comparações pareadas, tanto o Ensino Fundamental quanto o Ensino Superior reproduziram o efeito clássico de assimetria sujeito-objeto nos períodos curtos. Ou seja, na condição curta, as relativas de objeto foram mais difíceis de processar (ESCOvsESCS: $t(95) = 2,17$ $p < 0,033$ / EFCOvsEFCS: $t(95) = 3,04$ $p < 0,003$). Não houve diferença significativa entre relativas de sujeito e objeto nas demais comparações pareadas, tanto nas longas do ensino Fundamental e Superior quanto nas curtas e longas do Ensino Médio.

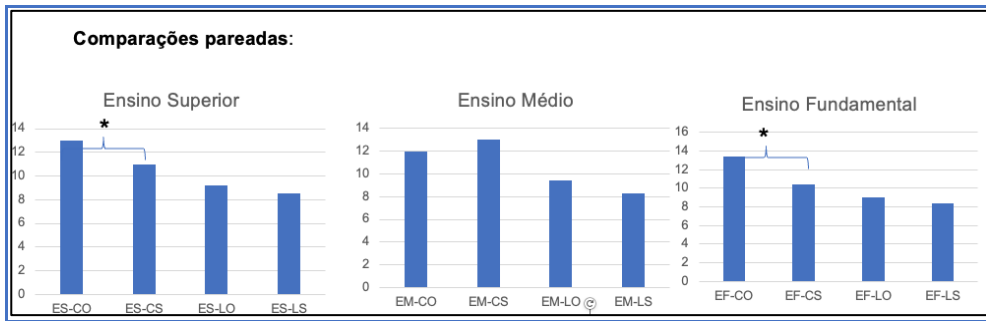


Gráfico 2. Comparativo de Contagem de fixações na área da OR entre os grupos.

	ES	ES	ES	ES	EM	EM	EM	EM	EF	EF	EF	EF
Extensão	C	C	L	L	C	C	L	L	C	C	L	L
Estrutura	O	S	O	S	O	S	O	S	O	S	O	S
Média	13,01	11,01	9,19	8,56	12,00	13,03	9,42	8,28	13,42	10,47	9,00	8,43

Tabela 5. Médias de *Fixation Count* (CF) na área da OR.

6. DISCUSSÃO

A duração total de fixações (*Total Fixation Duration*) e a quantidade de fixações (*Fixation Count*) indicam quanto tempo o participante gastou na leitura e quantas vezes fixou o olhar no segmento alvo -- que, neste caso, foi a oração relativa -- e funcionam como uma medida do custo de processamento. Sendo assim, os leitores do Ensino Superior e do Ensino Fundamental tiveram menos dificuldade para processar as orações relativas de sujeito do que as de objeto, na condição Curta. Na condição longa não houve diferença de processamento em relação ao tipo de oração relativa, indicando, possivelmente, que a intercalação de elementos entre o SN simples de sujeito e o verbo deixaram menos proeminente o efeito estrutural. Independente do efeito estrutural, os estudantes do Ensino Superior tiveram melhor desempenho na compreensão, indicado pelos maiores índices de acurácia nas respostas.

Em linhas gerais, os resultados corroboram, em parte, a nossa hipótese de que o efeito de assimetria sujeito-objeto não seria tão detectável nos períodos mais longos em função da escolaridade. No entanto, o efeito não foi encontrado na condição longa em nenhum dos grupos. Na condição curta, o efeito estava presente tanto no Ensino Superior quanto no Ensino Fundamental. Esse resultado está de acordo com a ideia reiterada em Maia e Rodrigues (2024) de que a capacidade de construir estruturas complexas, em relações hierárquicas, e que permite a compreensão e a produção da linguagem, está presente já no Ensino Fundamental. Como a relativa estava no primeiro segmento, mesmo leitores menos estruturantes tiveram sensibilidade ao efeito estrutural. Nos períodos longos, por outro lado, como houve a intercalação de um sintagma adjetival entre o SN simples e a oração relativa, as limitações do desempenho comprometeram o efeito estrutural, ou seja, não houve diferença no processamento de relativas

de sujeito e de objeto. O resultado mais surpreendente foi que esse comprometimento ocorreu também no Ensino Superior.

Uma possibilidade é também de que os leitores do Ensino Superior também não estejam imprimindo uma estratégia tão estruturante. Ainda que tenham tido o melhor desempenho, ainda houve um índice elevado de erro nas respostas à questão de compreensão, em torno de 20%. Nesse sentido, seria interessante replicar esse experimento considerando o perfil leitor para separar os grupos, no lugar de escolaridade. Os resultados do Ensino Médio podem indicar uma lacuna na passagem de uma etapa escolar para outra, interferindo no desempenho da leitura. Isso também pode ter relação com o alto índice de erros identificado no Ensino Superior.

7. CONSIDERAÇÕES FINAIS

Diversos trabalhos argumentam que estruturar o que se lê libera recursos atencionais e de memória para que seja possível operar uma quantidade maior de informação (Miller 1956; Kleiman 2013; Hyöna, Lorch e Kaakinen 2002; Khalifa e Weir 2009). Essa prática perpassa os diferentes níveis estruturais inerentes à prática de leitura, a saber: decodificação alfabética, fatiamento sintático e integração textual. Hyöna *et al.* (2002) mostram que leitores que estruturam os diferentes assuntos de um texto em seções apresentam maior capacidade de lembrar o conteúdo. Essa maior capacidade é encontrada em relação a leitores que adotam uma estratégia de leitura linear. Neste artigo, exploramos a relação entre conhecimento sintático e leitura no nível do período.

Nossos resultados indicaram que mesmo leitores menos experientes são sensíveis a fatores estruturais na leitura, como indicado pelos dados de rastreamento ocular. Esse trabalho, portanto, contribui para reforçar a necessidade de salientar o papel da hierarquia estrutural da língua nas práticas educacionais voltadas para o desenvolvimento da habilidade em leitura/ competência leitora.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abreu, Kátia e Victor Lima. 2022. A leiturabilidade no ensino fundamental e no superior, em Marcus Maia (org), *Psicolinguística: diversidades, interfaces e aplicações*, São Paulo, Contexto: 109-127.
- Allington, Richard L. 2005. NCLB reading first, and whither the future em *Reading Today*, 23(2): 18.
- Câmara dos Deputados. Projeto de lei 6256/2019. Institui a Política Nacional de Linguagem Simples nos órgãos e entidades da administração pública direta e indireta de todos os entes federativos.
- Chomsky, Noam. 2005. Three factors in language design, em *Linguistic inquiry*, 36(1): 1-22.
- Forster, Renê e Letícia Maria Sicuro Corrêa. 2017. On the asymmetry between subject and object relative clauses in discourse context, em *Revista de Estudos da Linguagem*, [S.l.], 25(3): 1225-1254.
- Gallego, Ángel J. e Noam Chomsky. 2020. A faculdade da linguagem: um objeto biológico, uma janela para a mente e uma ponte entre disciplinas, em *Revista Linguística*, 16: 52-84. [em linha] Disponível em: <https://doi.org/10.31513/linguistica.2020.v16nEsp.a39404>
- Garcia, Daniela Cid de; Marília Uchôa Cavalcanti Lott de Moraes Costa, Cristiane Oliveira da Silva e Rogério Santana Lourenço. 2020. O que está escrito aqui? Redescobrimo a gramática nos caminhos

- para a leitura what's written here? Redescoving grammar on the paths to reading, em *Revista Prolíngua*. 1983
- Gibson, Edward. 1998. Linguistic complexity: locality of syntactic dependencies, em *Cognition* 68(1): 1-76.
- Gough, Philip B. e William E. Tunmer. 1986. Decoding, reading and reading disability em *Remedial and Special Education*, 7(1): 6-10.
- Guedes, Ramon Brasileiro. 2021. *Animacidade e o processamento de orações relativas de sujeito e objeto por bilíngues português-ínglês*. Dissertação (Mestrado), Universidade Federal da Paraíba, João Pessoa. Disponível em <https://repositorio.ufpb.br/jspui/handle/123456789/24069>
- Hyöna, Jukka; Robert F. Lorch Jr e Johanna K. Kaakinen. 2002. Individual differences in reading to summarize expository text: Evidence from eye fixation patterns, em. *Journal of Educational Psychology*, 94: 44-55.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. 2023. *Notas sobre o Brasil no Pisa 2022*. Brasília, DF. [Cit.: INEP]
- Kato, Mary Aizawa. 2000. *No mundo da escrita: uma perspectiva psicolinguística*, 7ª ed., São Paulo, Editora Ática.
- Kenedy, Eduardo. 2013. *Curso básico de linguística gerativa*. São Paulo: Contexto.
- Kenedy, Eduardo. 2021. Heterogeneidade e homogeneidade nas línguas: uma síntese histórica para o surgimento da teoria das múltiplas gramáticas e uma análise sobre a articulação entre orações em sentenças do português brasileiro, em *Cadernos de linguística*, 2 (1): E327-e327.
- Khalifa, Hanan e Cyril J. Weir. 2009. *Studies in Language Testing: Examining Reading: Research and Practice in Assessing Second Language Reading*, 29.
- Kleiman, Angela. 2013. *Oficina de leitura: teoria e prática*. São Paulo, Pontes.
- Maia, Marcus. 2015. Sintaxe Experimental, em Gabriel de Ávila Othero e Eduardo Kenedy (orgs.) *Sintaxe, Sintaxes: uma introdução*. São Paulo, Contexto: 1: 51-72.
- Maia, Marcus. 2018. Computação estrutural e de conjunto na leitura de períodos: um estudo de rastreamento ocular, em Marcus Maia, *Psicolinguística e educação*, Campinas, SP, Mercado de Letras.
- Maia, Marcus (org.). 2019a. *Psicolinguística e metacognição na escola*. Campinas, SP, Mercado de Letras.
- Maia, Marcus. 2019b. Pensando (psico)linguisticamente, experimentalmente, educacionalmente, em Eloisa Pilati, Rozana Naves e Heloisa Salles (orgs.). *Novos olhares para a gramática na sala de aula: questões para estudantes, professores e pesquisadores*. São Paulo, Pontes: 248.
- Maia, Marcus. 2019c. Dimensões do Processamento Sintático, em Maílce Borges Mota e Cristina Name (org.), *Interface linguagem e cognição: contribuições da Psicolinguística*, Tubarão, SC: Copiart, 1: 89-114.
- Maia, Marcus e Gisele Rodrigues. 2024. Linguagem simples e Complexa: Metacognição epilinguística para o Letramento e Cidadania plenos. Dossiê Língua e Tecnologia. São Paulo, Museu da Língua Portuguesa: 25-41.
- Mak, Willem M.; Wietske Vonk e Herbert Schriefers. 2002. The influence of animacy on relative clause processing, em *Journal of memory and language*, 47(1), 50-68.
- Miller, George A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information, em *Psychological Review*, 62(2): 81-97.
- Organização para a Cooperação e Desenvolvimento Econômico. 2023. *PISA 2022 Results. The State of Learning and Equity in Education*, vol I Publishing, Paris. [em linha]Disponível em: <https://doi.org/10.1787/53f23881> (Cit.: OECD)
- Oliveira, Fernando Lúcio de.2013. *O Processamento da Assimetria Sujeito-Objeto em Construções do Tipo QU no Português Brasileiro: Interrogativas e Relativas*. Dissertação (Mestrado), Universidade Federal do Rio de Janeiro, Rio de Janeiro [em linha] Disponível em: <https://ppglinguistica.letras.ufjr.br/teses-e-dissertacoes>
- Oliveira, Fernando Lúcio de. 2017. O Efeito da Lacuna Preenchida na posição de sujeito no português brasileiro, em. *Soletas*, (33): 89-104.

- Oliveira, Fernando Lúcio de; Marcus Maia e Aniela Improtta França. 2018. Processamento de interrogativas-QU em Português Brasileiro: evidências de eyetracking e EEG, em *Letras de Hoje*, 53(1): 24-38.
- Pilati, Eloisa. 2017. *Linguística, gramática e aprendizagem ativa*. Campinas, Pontes Editores.
- Pilati, Eloisa; Rozana Naves e Heloisa Salles (orgs.). 2019. *Novos olhares para a gramática na sala de aula: questões para estudantes, professores e pesquisadores*. São Paulo, Pontes.
- Richards, Jack C; John Talbot Platt e Heidi Platt. 1992. *Longman dictionary of language teaching and applied linguistics*. London, Longman.
- Stanovich, Keith E. 1985. Explaining the variance in reading ability in terms of psychological processes: What have we learned, em *Annals of Dyslexia*, 35(1): 67.
- Staub, Adrian. 2010. Eye movements and processing difficulty in object relative clauses, em *Cognition*, 116: 71-86.
- Traxler, Matthew J., Robin K. Morris e Rachel E. Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements, em *Journal of Memory and Language*, 47(1): 69-90.
- Trotzke, Andreas; Markus Bader e Lyn Frazier. 2013. Third factors and the performance interface in language design, em *Biolinguistics* 7: 001-034.
- Warren, Tessa e Edward Gibson. 2002. The influence of referential processing on sentence complexity, em *Cognition*, 85(1): 79-112.
- Wolfe, Michael, Maureen Elizabeth Schreiner, Bob Rehder, Darrell Laham, Peter Foltz, Walter Kintsch e Thomas Landauer. 1998. Learning from text: Matching readers and texts by latent semantic analysis, em *Discourse Processes*, 25(2-3): 309-336.