

PROCESSAMENTO DE INFERÊNCIAS DE CAUSA NA LEITURA DE SENTENÇAS**PROCESSING OF CAUSE INFERENCES IN READING SENTENCES**Sabrina Santos¹

RESUMO: Este trabalho investigou o processamento de sentenças com relações de inferência de causa interoracionais, comparando dois grupos com diferentes níveis de experiência em leitura, a saber, Ensino Fundamental (EF) e Ensino Superior (ES). Por meio de técnica de rastreamento ocular, contrastaram-se dados experimentais da leitura de sentenças com relações de causa direta (e.g. A aluna recebeu os parabéns porque/já que ela obteve sucesso no exame do ENEM), e de causa indireta (e.g. A aluna escreve textos bons porque/já que ela obteve sucesso no exame do ENEM) avaliando também a influência dos conectivos (*já que* e *porque*) utilizados para articular as orações. As perguntas de pesquisa foram: se e de que modo a experiência em leitura, influencia no processamento de sentenças que estabelecem coerência de causa por mecanismos inferenciais distintos. As hipóteses levantadas foram as de que: i) as relações de causa indireta apresentam maior custo de processamento do que as relações de causa direta, pois requerem geração de inferências a partir do conhecimento de mundo do leitor; ii) as propriedades gramaticais dos conectivos levam a maior, ou menor, facilitação do processamento da sentença; iii) há correlação entre baixa habilidade em leitura e dificuldades de parseamento. Os resultados indicaram que as relações causais indiretas são mais custosas de serem processadas para ambos os grupos de participantes. Também identificamos que estudantes de EF se favorecem mais da presença dos conectivos do que o grupo ES.

PALAVRAS-CHAVE: Processamento. Relações Causais. Conectivos. Inferência na Leitura.

ABSTRACT: This work investigated the processing of sentences with intersentential inference, comparing two groups of different levels of school education, namely, Middle School and Undergraduate participants. Through an eye-tracking technique, experimental data from sentence reading were contrasted with direct cause relationships (e.g., *A aluna recebeu os parabéns porque/já que ela obteve sucesso no exame do ENEM*), and of indirect cause (e.g., *A aluna escreve textos bons porque/já que ela obteve sucesso no exame do ENEM*) Also evaluating the influence of connectors (since and why) used to articulate as prayers. The research questions were: whether and how the experience in reading influences the processing of statements with causal coherence by different inferential mechanisms. The hypotheses were: i) indirect cause relations have a higher processing cost than direct causal relations since the generation of inferences are made from the reader's knowledge of the world; ii) grammatical properties of the connectors lead to higher or lesser facilitation of sentence processing; iii) there is a correlation between low reading skills and parsing difficulties. The results indicated that indirect causal relationships are more costly to be processed for both groups of participants. We also found that E students favor the presence of connectors more than the ES group.

KEYWORDS: Processing. Causal Relationships. Connectives. Inference in Reading.

1 Introdução

Este trabalho investigou o processamento de sentenças com relações de inferência de causa interoracionais, comparando dois grupos de diferentes níveis de formação escolar, a saber, Ensino Fundamental (EF) e Ensino Superior (ES). Por meio de técnica de rastreamento ocular, contrastaram-se dados experimentais da leitura de sentenças com relações de causa

¹ Doutora pelo Programa de Pós-Graduação em Linguística na UFRJ (Bolsista CNPq). É membro pesquisadora do Laboratório de Psicolinguística Experimental (LAPEX/UFRJ) e do Grupo de Educação Multimídia, programa de extensão ligado ao Núcleo Interdisciplinar para o Desenvolvimento Social (NIDES/UFRJ), onde coordena projeto de extensão voltado para a formação de leitores em parceria com escolas da rede pública. E-mail: sabrinalopes@letras.ufrj.br.

direta, como (1), e de causa indireta, como (2), avaliando também a influência dos conectivos (*já que* e *porque*) utilizados para articular as orações.

- (1) A aluna recebeu os parabéns porque/já que ela obteve sucesso no exame do ENEM.
- (2) A aluna escreve textos bons porque/já que ela obteve sucesso no exame do ENEM.

As perguntas de pesquisa foram se e de que modo a experiência em leitura influencia no processamento de sentenças que estabelecem coerência de causa por mecanismos inferenciais distintos, em (1), inferência sentencial e, em (2), inferência discursiva (ou extrassentencial). As hipóteses levantadas foram as de que: i) as relações de causa indireta apresentam maior custo de processamento do que as relações de causa direta, inobstante o nível de habilidade em leitura, pois requerem geração de inferências a partir do conhecimento de mundo do leitor, i.e., exigem informações extrassentenciais, representando, assim, um nível a mais de processamento; ii) as propriedades gramaticais dos conectivos levam a maior, ou menor, facilitação do processamento da sentença; iii) baixa habilidade em leitura pode estar relacionado a dificuldades de parseamento das unidades significativas, refletindo em maior custo de processamento para ambos os tipos de relações causais.

As orações investigadas por este trabalho são chamadas de orações subordinadas causais e orações coordenadas explicativas pela tradição gramatical (CUNHA; CINTRA, 2002). Entretanto, neste artigo, utilizou-se, de modo indiscriminado, o tratamento de orações causais, pois o debate versa sobre mecanismos necessários para se estabelecer coerência de causa necessária à compreensão e não pretende propor análises para a representação gramatical dessas estruturas.

Este trabalho é parte da pesquisa de doutorado da autora (SANTOS, 2020), inserindo-se em um empreendimento que buscou construir um diálogo entre a Psicolinguística e a Educação Básica, fornecendo dados experimentais que possam vir a contribuir para o desenvolvimento de materiais e de práticas pedagógicas de ensino-aprendizagem da leitura. Este projeto também contribui com pesquisas desenvolvidas pela Rede Nacional de Ciência para Educação – Rede CpE – (<http://cienciaparaeducacao.org/>), formada por pesquisadores e laboratórios de diferentes áreas do conhecimento em torno das Ciências Cognitivas. As pesquisas no âmbito das ciências cognitivas são informativas para se compreender como os seres humanos aprendem, mas esses resultados têm sido pouco considerados no desenvolvimento de currículos e práticas e materiais pedagógicos. Fatores econômicos, sociais e políticos representam parte significativa do problema educacional do Brasil, mas, para além disso, é preciso refletir de que maneira resultados das áreas das ciências cognitivas podem contribuir com abordagens de processos de ensino-aprendizagem.

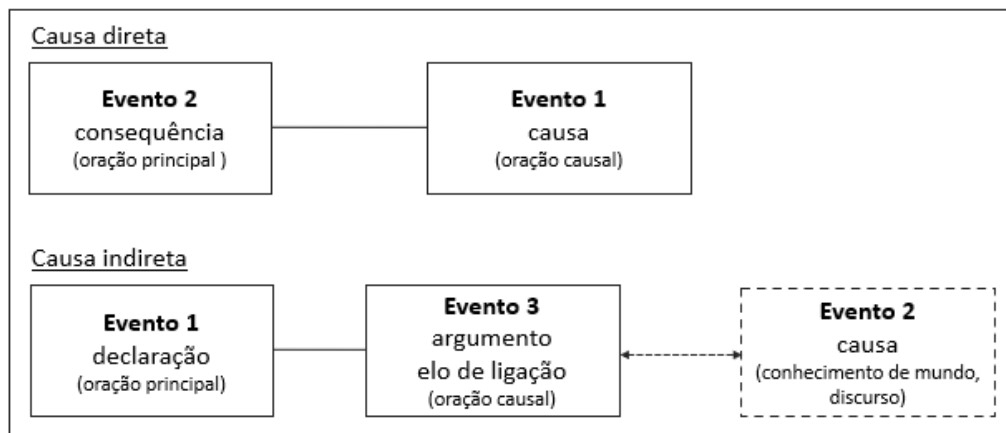
O artigo apresenta a seguinte organização: na seção 2, discute-se o papel da inferência na compreensão da leitura a partir das relações de causalidade; na seção 3, discutem-se alguns trabalhos sobre o processamento de relações causais e a influência dos conectivos causais; na seção 4, descreve-se o experimento realizado a fim de responder às questões de pesquisa levantadas, e, na seção 5, conclui-se e fazem-se as considerações finais do trabalho.

2 Inferência de causa e leitura

As sentenças exemplificadas em (1) e (2) exibem relações de causa que se instanciam por meio das chamadas *inferências necessárias*, consideradas indispensáveis para a compreensão por estabelecerem a coerência entre significados (SINGER, BROOKE, 2012). Porém, há uma distinção entre a inferência derivada em (1) para a derivada em (2). Enquanto em (1), a relação de causa é estabelecida a partir dos significados expressos explicitamente pelos constituintes

das orações ([receber os parabéns] [obter sucesso no exame]), em (2), estabelecer a coerência de causa depende da geração de inferências a partir de informações implícitas, relacionadas ao conhecimento de mundo do leitor e/ou à situação pragmática de enunciação. Em (2), há a necessidade de acessar o conhecimento de mundo para que a coerência de causa se efetue, caracterizando também uma *inferência de ligação – bridging inference* (cf., SINGER, BROOKE, 2012). Segundo Singer e Brooke (2012), no caso de relações de causalidade, a coerência requer uma inferência de ligação em que o primeiro evento tenha causado o segundo. Observando os exemplos (1) e (2), vemos que essa relação está explícita entre os conteúdos das orações de (1), porém não se pode dizer o mesmo entre as orações da sentença em (2). Essa distinção ocorre porque, na sentença em (1), apresenta-se a relação de causalidade *default* do tipo consequência→causa, e, em (2), uma relação de causa não *default* do tipo declaração→argumento. Em (1), vemos uma ordenação temporal em que o evento expresso pela oração à esquerda (*A aluna recebeu os parabéns*) sucede o evento da oração à direita (*já que/ porque ela obteve sucesso no exame do Enem*). Em (2), essa ordenação não se configura, ou seja, não é possível atribuir à oração à direita valor de causa. A oração (causal) é interpretada como uma pista, ou argumento, para inferir-se a declaração expressa pela oração à esquerda (principal). Isso significa que os conteúdos expressos pela oração identificada como causal, em (2), não significam causa real, mas representam uma ponte ou um elo de ligação com a informação implícita, i.e., a causa. A coerência entre as orações se dá, portanto de forma indireta. No caso do exemplo em (2), podemos assumir, por exemplo, essa informação implícita, que tem valor de causa, como “a estudante fez uma boa redação no Enem”, mas isso não está explícito na sentença. Em (1), não precisamos desse nível extrassentencial, resultando em menos material a ser processado relativamente a sentenças do tipo visto em (2). O Quadro 1, abaixo, ilustra a diferença entre os níveis informacionais e de ordenação de eventos dessas relações em uma dada sentença:

Quadro 1 – Representação dos níveis e ordenação das informações em relações causais interoracionais.



Fonte: autora

Na literatura, há evidências robustas indicando que inferências necessárias são realizadas durante a leitura (cf., VAN DEN BROEK, 1994; GARROD; SANFORD, 1994). Porém, as inferências que requerem um nível a mais de informação, isto é, que exigem acesso a informações discursivas, têm apresentado maior custo de processamento comparativamente àquelas que se estabelecem no nível da sentença (e.g., MODESTO; MAIA, 2017; MEDEIROS, 2019; TRAXLER; BYBEE; PICKERING, 1997; TRAXLER et al., 1997, CANESTRELLI; MAK; SANDERS, 2012).

A inferência é uma propriedade fundamental para a compreensão na leitura, pois atende à necessidade de construção da coerência entre os significados dos conteúdos de uma dada sentença ou de um dado texto. Na compreensão da leitura, a coerência precisa ser estabelecida o mais rápido possível para que não haja sobrecarga da memória de trabalho, sistema cognitivo fundamental para que o leitor reconheça e agrupe as unidades significativas do material lido de forma eficiente (KLEIMAN, 2013[1989]). Com uma memória de trabalho bem desenvolvida, o leitor terá habilidade de fazer conexões entre o texto e seu próprio conhecimento de mundo de modo eficiente e de relacionar cada proposição sucessiva ao que foi lido anteriormente sem sobrecarregar esse sistema (TANKERSLEY, 2003). A memória de trabalho tem uma capacidade de armazenamento de informações limitada, assim, se um indivíduo, durante o ato da leitura, despende muito tempo em processos cognitivos como decodificação e reconhecimento de palavras, acabará por sobrecarregar esse mecanismo, resultando em perdas para o processamento. Moraes e Kolinsky (2018) argumentam que, para ser considerado letrado, o indivíduo deve exibir níveis elevados de automaticidade de processamento na leitura (e.g., decodificação e reconhecimento da palavra, relações sintagmáticas), da habilidade de analisar e relacionar informações (e.g., fatos, julgamentos e ideias), de avaliar criticamente o material lido e de expressá-lo por meio da escrita de modo coerente e de um uso criativo da língua. Essa automaticidade libera recursos cognitivos, otimizando, assim, a memória de trabalho (MORAIS, 2017).

Indicadores nacionais e internacionais que avaliam as habilidades em leitura, como IDEB (Índice de Desenvolvimento da Educação Básica), ANA (Avaliação Nacional da Alfabetização), ENEM (Exame Nacional do Ensino Médio), Prova Brasil, INAF (Indicador Nacional de Alfabetismo Funcional), PISA (*Programme for International Student Assessment*), entre outros, revelam a situação precária em variados níveis de escolaridade e nichos socioeconômicos. O PISA (OCDE, 2018), por exemplo, indicou que apenas 2% dos estudantes da Educação Básica no Brasil atingiram os níveis mais elevados em sua escala classificatória de habilidade em leitura.² O texto de base avaliativa e analítica do PISA (OCDE, 2017) aponta especificamente as relações de causa como fator de dificuldade no desenvolvimento da competência leitora: “[...] compreender uma ligação causal específica é mais fácil do que inferir uma relação de causalidade implícita” (OCDE, 2017, p.40)³. Neste sentido, ao investigar o processamento *on-line* das relações de causa interoracionais, este trabalho contribui com dados experimentais que auxiliem a elaboração de práticas pedagógicas de reversão do lamentável quadro que vem sendo apontado por indicadores de avaliação das habilidades em leitura entre os brasileiros.

Estudos psicolinguísticos recentes, reunidos em Maia (2018a), apontaram para uma correlação significativa entre a dificuldade de hierarquização das unidades significativas no processamento sentencial e um efeito de “desaderência da leitura” entre participantes de Ensino Fundamental de uma escola da rede pública no município de Niterói, localizado no estado do Rio de Janeiro. Entre esses trabalhos, Maia (2018b) propõe, a partir da tipologia de leitores proposta por Hyöna, Lorch e Kaakinen (2002), que esses estudantes de EF apresentaram comportamento de um leitor linear incompleto, por abandonarem a leitura antes mesmo de atingirem o final da sentença. A essa desaderência, Maia (2018b) atribuiu sobrecarga da memória de trabalho:

² O PISA, implementado pela OCDE (Organização para a Cooperação e Desenvolvimento Econômico), avalia os conhecimentos e habilidades essenciais de estudantes com idade de 15 anos em leitura, matemática, ciência e capacidade inovadora.

³ Tradução livre, “[...]and comprehending a specified causal link is easier than inferring an implicit causal relationship.” (OCDE, 2017, p. 40).

[...]o grupo fundamental faz uma leitura linear, não estrutural, do período, dedicando menos atenção às orações, progressivamente, ao longo da leitura. Pode-se inferir que, não se conseguindo fazer uma análise estrutural, as orações vão se sucedendo, em um conjunto aparentemente linear, minimamente estruturado, sem se organizar hierarquicamente, sobrecarregando a memória de trabalho e tornando cansativa a leitura, com conseqüente desengajamento progressivo da atenção.(MAIA, 2018b, p.125)

Hyöna, Lorch e Kaakinen (2002) classificaram leitores em dois tipos: linear e seletivo estruturante. Leitores lineares se caracterizam por realizar uma leitura basicamente progressiva, sem exibir índices significativos de regressões no texto, que são considerados padrões de leituras destinados à integração de informações. Leitores considerados seletivos, por outro lado, são aqueles que exibem movimentos regressivos (*lookbacks*) em áreas anteriores do texto que está sendo lido. Essas retomadas podem ser feitas de modo seletivo para regiões específicas do texto ou da sentença. Diante do comportamento identificado, Maia (2018b) instanciou o leitor linear incompleto que se desengaja e abandona a leitura. O presente estudo também permitiu verificar se os participantes de ambos os grupos apresentavam padrões de leitura de acordo com essas tipologias, esperando encontrar diferenças entre os grupos a partir do processamento das relações causais e do papel que os conectivos investigados possam exercer sobre essas relações.

3 Relações causais e a influência de conectivos

O trabalho seminal acerca do processamento de relações causais interoracionais em língua inglesa de Traxler, Bybee e Pickering (1997) já evidenciava o maior custo de processamento para a causa indireta em relação à causa direta⁴. Traxler, Bybee e Pickering (1997), ao compararem sentenças como (3) e (4), por meio de técnica de rastreamento ocular, evidenciaram que os participantes imediatamente consideravam as informações já computadas do resto da sentença ao se depararem com o conectivo.

(3) Causa indireta

Susan didn't take care of money and credit cards because/ she left her purse /at the bus stop.

‘Susan não teve cuidado com dinheiro e cartões de crédito porque ela deixou sua carteira no ponto do ônibus.’

(4) Causa direta

Susan lost her money and her credit cards because /she left/ her purse/ at the bus stop.

‘Susan perdeu seu dinheiro e seus cartões de crédito porque deixou sua carteira no ponto do ônibus.’

Os autores identificaram aumento de custo de processamento significativo em região pós-conectivos, como a destacada nos exemplos (3) e (4), quando a relação de causa se dava de forma indireta. Isso significa que antes de atingirem o final da frase, os participantes já haviam computado a informação contida no conectivo. A causalidade *default* é do tipo consequência→causa, assim, o conectivo causal *because* aciona inicialmente essa relação, na medida em que a leitura avança essa relação não se instancia, resultando em maior custo de processamento por necessidade de reanálise. Os autores atribuíram a diferença de processamento à necessidade de as relações indiretas exigirem inferência discursiva para estabelecer a relação de “consequência causal”. Canestrelli, Mak e Sanders (2012), por sua vez,

⁴ Os autores utilizam *diagnostic statements* e de *causal statements* para se referir à relação que está sendo chamada de causa indireta e causa direta, respectivamente no presente trabalho.

também encontraram essas diferenças na língua holandesa. No entanto, os autores também tinham o objetivo de avaliar se o papel dos conectivos *omdat* e *want* (“porque” e “já que”) influenciam de forma distinta o processamento de relações causais direta e indireta interacionais, pois assumem que há especificação de distribuição dos conectivos causais em língua holandesa.⁵ A hipótese dos autores foi a de que *omdat* é prototípico de causa direta e *want*, prototípico de causa indireta. Ao submeterem participantes a experimento de rastreamento ocular, em que compararam sentenças com combinação prototípica entre conectivos e relações causais, como em (5a) e (6a), a sentenças com combinações invertidas, como em (5b) e (6b), Canestrelli, Mak e Sanders (2012) verificaram que a presença do conectivo *want* resultou em maior custo de processamento independentemente do tipo de relação.⁶

(5) Relação indireta

a. *Hanneke had haast, want ze was vier trappen afgerend om de post te halen.*

b. *Hanneke had haast, omdat ze was vier trappen afgerend om de post te halen.*

‘Hanneke estava com pressa PORQUE ela desceu correndo quatro lances de escada para pegar o correio.’

(6) Relação direta

a. *Hanneke was buiten adem, omdat ze vier trappen was afgerend om de post te halen.*

b. *Hanneke was buiten adem, want ze vier trappen was afgerend om de post te halen.*

‘Hanneke estava sem fôlego PORQUE ela desceu correndo quatro lances de escada para pegar o correio

A presença de *omdat* em contexto declaração → argumento resultou em efeito facilitador em estágio inicial de processamento, mas gerou maior custo no decorrer da leitura da sentença. Os autores atribuem o custo elevado da combinação *omdat* com a relação indireta a um “efeito *garden-path* discursivo”, uma vez que esse conectivo introduz uma relação do tipo consequência → causa (CANESTRELLI; MAK; SANDERS, 2013:23), mas seus conteúdos semânticos não expressam essa relação. A presença desse conectivo em contexto declaração → argumento conduz, portanto, o leitor a uma reanálise.

Em língua portuguesa, não há estudos de processamento que suportem essa especificação, identificada em língua holandesa, entre os conectivos causais, embora, na teoria gramatical, já se indique que não há essa distribuição categórica entre os conectivos e as relações causais (LOBO, 2003). Sob perspectiva teórica da gramática gerativa (CHOMSKY, 1995), Lobo (2003) argumenta que o conectivo *já que* é especificado no léxico com traço [+pressuposicional], assim, o conteúdo oracional introduzido por este conectivo será esperado ou pressuposto, no entanto, também afirma que o conectivo *porque* é subespecificado para esse traço. Essa proposta se alinha com a análise mais recente de Rodrigues (2018). A autora sugere que orações causais introduzidas por *porque* possuem estruturas informacionais distintas.⁷

⁵ Outras línguas também apresentam distribuição específica para as relações de causa direta e indireta, como as línguas francesa, a alemã e o mandarim (cf. STUKKER; SANDERS, 2011; LI ET AL, 2017).

⁶ Trabalhos de processamento do discurso de linha cognitiva, como Canestrelli, Mak e Sanders (2012), frequentemente se referem a relações de causa direta e de causa indireta como objetiva e subjetiva, respectivamente, por considerarem a perspectiva da proposição.

⁷ Define-se pressuposição como a informação considerada pelo falante como pressuposta por seu interlocutor e foco a informação a ser dada pelo falante. Comumente, atribui-se uma distinção entre informação velha (pressuposta) e informação nova (focalizada). No âmbito da Teoria Gerativa, a representação estrutural das sentenças prevê a hierarquização dos domínios gramaticais em três níveis: a camada lexical (vP), a camada flexional (IP) e a camada complementizadora (CP), definida como a periferia esquerda da sentença. Rizzi (1997)

Segundo Rodrigues (2018), essas orações podem exibir uma construção causal associada a uma única estrutura foco-suposição, em que a oração subordinada veicula a suposição, e a oração causal, o foco; ou a construção causal pode ser associada a duas estruturas foco-suposição distintas.

- (7) A Maria foi embora porque o Pedro chegou.
- (8) A Maria saiu, porque as luzes estão apagadas.
- (9) A Maria foi embora, porque o Pedro chegou.

Rodrigues (2018) sugere que orações causais podem exercer tanto função de adjunto de predicado, como (7), quanto de adjunto de frase, como (8) e (9), porém, como pode ser observado nesses exemplos, o conteúdo das sentenças (7) e (9) é o mesmo, porém a função da oração causal será distinta. A autora sugere que no caso de assumir a função de adjunto de frase, haverá uma curva prosódica distinta que pode ser identificada por uma vírgula, como em (9).⁸ Há, portanto, consenso na literatura de processamento e de teoria linguística de que relações causais interoracionais e conectivos causais possuem propriedades distintas. Entretanto, há evidências que sugerem uma correlação entre tendência de uso do conectivo *já que* em contexto de relação de causa indireta e experiência em leitura, comparativamente ao uso do conectivo *porque* (SANTOS, GARCIA, MAIA, 2019) no mesmo contexto de relação causal. Santos, Garcia e Maia (2019) realizaram teste de julgamento de escolha forçada comparativamente entre estudantes de EF e ES. No experimento, os participantes deveriam escolher entre os conectivos *porque* e *já que* para completar o sentido de sentenças com relações de causa direta ou de causa indireta. Os resultados indicaram que os participantes do grupo de EF selecionavam significativamente mais o conectivo *porque* em ambos os contextos de causalidade; os participantes de ES, por outro lado, exibiram preferência significativa em relação ao conectivo *já que* no contexto de causa indireta comparativamente à seleção do mesmo conectivo em contexto de causa direta. O experimento descrito a seguir avaliou como essas diferenças se relacionam com experiência em leitura ao comparar participantes de EF a de ES no processamento *on-line*.

4 O experimento

A técnica de rastreamento ocular (*eye-tracking*) permite aferir tempos e números de fixações oculares em áreas de interesse definidas de acordo com as questões que se desejam investigar, possibilitando estabelecer correlações entre esses padrões de fixações e diferentes processos cognitivos. O rastreador ocular também possibilita mensurar medidas de movimentos regressivos e progressivos, ou seja, para regiões posteriores ou anteriores a partir de uma dada área de interesse. Nesse sentido, uma dada área de interesse pode configurar uma palavra, uma oração ou um sintagma formado por mais de um constituinte, como um sintagma verbal em que se avalia o verbo seu complemento.

O presente experimento teve o objetivo de investigar diferenças de processamento entre sentenças com relações causais interoracionais e a influência que os conectivos *já que* e *porque* exercem sobre cada uma dessas relações, avaliando se o nível de habilidade em leitura dos participantes se refletiria no modo como essas sentenças são processadas. As previsões foram as de que:

expande essa camada para dar conta das relações de interface entre discurso e sintaxe, representando, assim, (a chamada estrutura informacional da sentença (e.g. tópico, foco, força ilocucionária, entre outras).

⁸ A explicação de Rodrigues (2018) para a distinção entre adjunção de predicado e de frase se baseia na proposta de Haegeman (2006), que propõe a distinção de orações adverbiais a partir de operações sintáticas cujas relações oracionais são orientadas para os eventos expressos ou para o discurso.

- i) ambos os grupos deveriam apresentar maior custo de processamento relativos à causa indireta, o que deveria ser traduzido por médias de fixações e refixações mais elevadas;
- ii) o conectivo *já que*, por ser codificado com traço [+pressuposicional], deveria facilitar de modo mais significativo a causa indireta e *porque*, por ser subespecificado, exigiria mais reanálises sobretudo para a causa indireta;
- iii) os participantes de ES deveriam apresentar médias mais elevadas de processamento das relações causais indiretas no estágio inicial do processamento, enquanto o EF deveria ter médias elevadas durante todo o processamento;
- iv) os participantes de EF deveriam apresentar índices de erros significativamente mais elevados nas respostas à pergunta interpretativa final do que os participantes de ES.

4.1 Método

Participantes

O experimento foi aplicado a 30 estudantes de Ensino Fundamental (EF) e 30 de Ensino Superior (ES). Os participantes do EF tinham de 14 a 17 anos e cursavam o 9º ano na E.M. Dilermando Cruz, localizada no bairro de Bonsucesso, na cidade do Rio de Janeiro. O grupo de ES, com idade média de 26,5 anos, foi composto por estudantes de graduação do curso de Letras da Universidade Federal do Rio de Janeiro. Para fins de homogeneidade quanto à competência leitora do grupo de ES, foram selecionados apenas estudantes a partir do 7º período. Este experimento, portanto, se configura em dois modelos: *within subjects* e *between subjects*, uma vez que foram comparados todos os fatores entre todos os participantes, assim como se contrastaram os dados experimentais obtidos entre os grupos de participantes (ES *versus* EF).

Design

O experimento contou com três variáveis independentes em um *design* fatorial 2x3x2 (tipo de causalidade – direta e indireta, conectivo – *já que*, *porque* e sem conectivo, e grupo – Ensino Fundamental e Ensino Superior). O cruzamento das variáveis tipo de causalidade e tipo de conectivo configurou um design 2x3, gerando seis condições de itens experimentais: **IJ – indireta com *já que***; **IP – indireta com *porque***; **IS – indireta sem conectivo**; **DJ – direta com *já que***; **DP – direta com *porque*** e **DS – direta sem conectivo**. A condição sem conectivo deveria garantir efeito principal da presença do fator conectivo, servindo assim como *baseline* do experimento.

As variáveis dependentes, ou seja, as medidas capturadas e analisadas foram as médias de tempo e números de fixações oculares nas áreas de interesse (Quadro 2), configurando as medidas *on-line*. Também se analisaram os índices de resposta às perguntas interpretativas referentes às sentenças experimentais, configurando, assim, as medidas *off-line*.

Condição	Sentenças			
	AI1	AI2	conect	AI3
Direta+ porque	Boa escrita é essencial na vida acadêmica.	A aluna recebeu os parabéns	porque	ela obteve sucesso no exame do ENEM.
Direta+ <i>já que</i>			<i>já que</i>	
Direta s/conect.			∅	
Indireta+ porque		A aluna escreve textos bons	porque	

Indireta+ já que			já que	
Indireta s/conect.			Ø	
Pergunta (causa direta)	Obter sucesso no exame fez a aluna receber os parabéns?		Sim ou Não	
Pergunta (causa indireta)	Obter sucesso no exame fez a aluna escrever textos bons?		Sim ou Não	

Fonte: autora

Materiais

Foram utilizadas 24 sentenças experimentais por condição, distribuídas em 6 listas com 5 participantes por lista, gerado um total de 720 observações por cada área de interesse (cf. Quadro 2). A cada sentença experimental seguiu-se uma pergunta interpretativa de acordo com o tipo de relação causal. As sentenças experimentais foram distribuídas em quadrado latino (*Latin Square*) de modo que todos os participantes vissem todas as condições experimentais, mas não a mesma sentença em cada condição. Os materiais ainda contaram com 48 sentenças distratoras. Note que a área de interesse 3 (AI3), que continha a oração causal, apresentou exatamente a mesma oração para todas as condições. Isso significa que diferenças identificadas nessa região só poderiam ser atribuídas a fatores relativos às demais regiões ou ao grupo de participantes. Como pode ser observado no Quadro 2, as condições experimentais foram constituídas de dois períodos: um, apresentando um contexto prévio, seguido da sentença com as orações que estabeleciam relação de causa. Foram controlados os tamanhos das orações de modo que a oração principal tivesse de 27 a 31 caracteres e a oração subordinada de 36 a 40 caracteres, ambas com estrutura de sujeito+verbo+objeto. O contexto prévio, que também era o mesmo entre as condições, continha orações simples com estrutura de sujeito+verbo de ligação+predicado, totalizando de 40 a 44 caracteres. Nas perguntas interpretativas finais, a oração causal foi mantida como evento causador considerando a semântica do verbo fazer. Nesse sentido, para acertar a pergunta relativa à causa direta, os participantes deveriam responder SIM, mas para acertar à pergunta referente à causa indireta deveriam responder NÃO, já que não pode ser a orações não pode ser interpretada como evento causador daquele que se segue após o verbo (i.e., *Obter sucesso no exame fez a estudante escrever textos bons?*).

Procedimentos

O experimento foi aplicado por meio de equipamento *Eyelink 1000* de registro monocular com frequência de 1000Hz. Os estímulos foram apresentados em uma tela de 32 polegadas com resolução de 1920x1080 utilizando o programa *Experiment Builder 2.2.1*. Os resultados foram analisados no programa *Eyelink Data Viewer 3.2*. Ambos os grupos de participantes realizaram o experimento em sala silenciosa com a presença apenas do experimentador de modo a minimizar a influência de elementos externos ao teste.

Os participantes foram orientados a automonitorar a leitura de períodos que apareceriam, um por vez, na tela de um computador acoplado ao equipamento de rastreamento ocular. Ao final da leitura de cada sentença, os participantes deveriam pressionar a barra de espaço no teclado posicionado à sua frente. Após ler cada sentença, os participantes deveriam pressionar a barra de espaço no teclado para que chamasse a próxima tela em que aparecia a pergunta interpretativa final. Para responder às perguntas, os participantes deveriam pressionar um botão destacado em verde se achassem que a resposta era SIM ou em vermelho, se achassem que a resposta era NÃO. Antes de começar o teste, os participantes foram submetidos a um processo de calibragem a fim de capturar o contraste da pupila e de seu reflexo (*glint*), meio pelo qual o equipamento rastreia os movimentos oculares. A média de tempo de execução

do experimento foi de 40 minutos para o EF e de 30 minutos para o ES, incluindo o período de calibragem.

4.2 Resultados

As seguintes medidas foram computadas: tempos totais de fixação, tempos de primeiras fixações ou primeiro passe, tempo de duração de regressões no primeiro passe e de refixações de segundo e de terceiro passe. Os tempos totais de fixação compreendem a duração total de todas as fixações, mensuradas em milésimos de segundo, sobre a região a ser analisada (RAYNER; LIVERSEDGE, 2004). Os tempos de fixações de primeiro passe se referem ao tempo total da primeira leitura em uma dada região. Essa medida possui caráter mais reflexo, relacionando-se com frequência e reconhecimento de palavras, acesso lexical e estruturação sintática (RAYNER, 1998; RAYNER et al. 2004). A duração de regressões de primeiro passe se refere ao tempo total de regressões, mensuradas em milésimos de segundos, feitas a partir da primeira entrada do olhar em uma determinada região até os olhos se moverem para a próxima região (RAYNER; LIVERSEDGE, 2004). As refixações de segundo e de terceiro passe (*lookbacks*) são medidas de caráter reflexivo, sendo mais informativas quanto a processos integrativos entre unidades significativas das sentenças e dos textos (HYÖNÄ; LORCH; KAAKINEN, 2002). Na Tabela 1, apresentam-se as médias acompanhadas do desvio padrão entre parênteses:

Tabela 1 – Médias de tempos totais de fixação (TTF), de primeiro passe (PP), de regressão de primeiro passe (RPP) e de *lookbacks* (regressões de segundo e terceiro passe) em milésimos de segundos. Área de interesse 3 (AI3); área de interesse 2 (AI2) e área de interesse 1 (AI1)

	Ensino Fundamental (EF)					
	DJ	DP	DS	IJ	IP	IS
<i>TTF (AI3)</i>	1451(465)	1509(501)	1552(447)	1380(419)	1643(433)	1799(572)
<i>PP (AI3)</i>	1028(404)	1215(405)	1338(390)	1232(387)	1282(458)	1529(567)
<i>RPP (AI3)</i>	1814(652)	1866(792)	1857(677)	1567(677)	2040(772)	2238(960)
<i>Lookbacks (AI1+AI2)</i>	638(420)	687(325)	800(350)	744(393)	628(458)	994(221)
	Ensino Superior (ES)					
	DJ	DP	DS	IJ	IP	IS
<i>TTF (AI3)</i>	1341(431)	1359(392)	1449(500)	1475(652)	1559(485)	1640(555)
<i>PP (AI3)</i>	916(379)	1074(334)	1027(370)	847(401)	1127(383)	1149(378)
<i>RPP (AI3)</i>	2036(919)	1979(724)	1994(969)	2253(1213)	2481(788)	2122(1164)
<i>Lookbacks (AI1+AI2)</i>	683(293)	635(264)	628(285)	673(236)	688(206)	694(386)

Fonte: autora

As análises estatísticas foram realizadas em programa de código aberto *ezANOVA* que permite a condução de análise de variância por sujeitos e por grupos e de Teste de Tukey para verificar as possíveis diferenças resultantes de comparação pareada entre as condições experimentais nos grupos (*within subjects*) e entre os grupos (*between subjects*).

4.2.1 Tempos totais de fixação sobre a área da oração causal

Obtiveram-se efeitos principais significativos para os três fatores: **grupo** ($F(1,188) = 9,24$ $p < 0,002^*$), **causa** ($F(1,188) = 24,2$ $p < 0,000002^{***}$) e **conectivo** ($F(2,376) = 19,1$

$p < 0,000001^{***}$), evidenciando que as diferenças encontradas podem ser atribuídas a cada um deles individualmente. Houve, ainda, identificação de efeitos de interação **grupo e conectivo** ($F(2,376) = 3,97$ $p < 0,01^*$) e **entre os três fatores** ($F(2,376) = 7,26$ $p < 0,0008^{**}$). Na direção esperada, a relação de causa indireta e a ausência de conectivos em ambos os tipos de causalidade resultaram em maior custo de processamento para ambos os grupos, considerando-se as médias obtidas.

Como esperado, o grupo de ES apresentou médias significativamente menores para a causa direta independentemente do tipo de conectivo ([ES_D_P 1359ms]vs[ES_I_P 1559ms] PAIRED $t(94)=3,27$ $p < 0,001^*$; [ES_D_J 1341ms]vs[ES_I_J 1475ms] PAIRED $t(94)=2,74$ $p < 0,007^{**}$). A ausência de conectivo na causa direta resultou em média significativamente mais elevada do que sua contraparte com a presença do conectivo *já que* ([ES_D_J 1341ms]vs[ES_D_S 1449ms] PAIRED $t(94)=3,01$ $p < 0,003^{**}$), mas não em oposição ao uso do conectivo *porque* ([ES_D_P 1359ms]vs[ES_D_S 1449ms] PAIRED $t(94)=1,44$ $p < 0,15^{ns}$).

O grupo de EF, como previsto, também exibiu médias de tempos totais de fixação significativamente maiores para a causa indireta comparativamente à causa direta ([EF_D_P 1509ms]vs[EF_I_P 1643ms] PAIRED $t(94)=2,09$ $p < 0,03^*$; [EF_D_S 1552ms]vs[EF_I_S 1799ms] PAIRED $t(94)=3,91$ $p < 0,0002^{***}$). Por outro lado, distintamente do que foi observado nos resultados do grupo de ES, a presença do conectivo *já que* na causa indireta levou a uma redução significativa das médias de tempos totais de fixação sobre a oração causal ao serem comparadas às contrapartes com ausência de conectivo e com o uso do conectivo *porque* ([EF_I_J 1380ms]vs[EF_I_P 1643ms] PAIRED $t(94)=5,36$ $p < 0,0001^{***}$; [EF_I_J 1380ms]vs[EF_I_S 1799ms] PAIRED $t(94)=7,15$ $p < 0,0001^{***}$). Esse resultado era esperado para o grupo de ES, mas não para o grupo de EF. Ainda de modo distinto ao grupo de ES, o grupo de EF apresentou diferenças significativas entre a causa indireta com conectivo *porque* e indireta sem conectivo ([EF_I_P 1643ms]vs[EF_I_S 1799ms] PAIRED $t(94)=2,49$ $p < 0,01^*$). Esse resultado pode indicar que para o grupo de EF a presença do conectivo influencia mais a compreensão da leitura do que para o grupo de ES no sentido de facilitar o processamento. Na comparação dos resultados entre os grupos, foi esperado que os participantes de EF apresentassem, de modo geral, maiores médias de tempos totais de fixação durante a leitura das sentenças. Essa previsão se confirmou, identificando-se diferenças significativas entre as médias da causa direta+já que ([EF_D_J 1451ms]vs[ES_D_J 1341ms] $t(188)=2,84$ $p < 0,004^{**}$), direta+porque ([EF_D_P 1509ms]vs[ES_D_P 1359ms] $t(188)=2,11$ $p < 0,03^*$) e entre indireta sem conectivo ([EF_I_S 1799ms]vs[ES_I_S 1640ms] $t(188)=4,00$ $p < 0,0001^{***}$).

4.2.2 Tempos de fixação de primeiro passe sobre a oração causal

A análise de variância sobre às médias de tempo de fixações de primeiro passe identificou efeito principal altamente significativo sobre os três fatores: **grupo** ($F(1,164) = 75,7$ $p < 0,000001^{***}$), **causa** ($F(1,164) = 19,3$ $p < 0,00002^{***}$) e **conectivo** ($F(2,328) = 33,5$ $p < 0,000001^{***}$). Houve ainda efeito significativo de interação entre os fatores **grupo e causa** ($F(1,164) = 7,57$ $p < 0,006^{**}$) e **grupo e conectivo** ($F(2,328) = 4,86$ $p < 0,008^{**}$). Novamente, o resultado foi no sentido esperado com a causa indireta apresentando as maiores médias. Por ser uma medida reflexa, foi esperado que apenas o grupo de ES apresentasse efeito de facilitação relativo ao conectivo *já que* no processamento da causa indireta ([ES_I_J 846ms]vs[ES_I_P 1127ms] PAIRED $t(82)=4,38$ $p < 0,0001^{***}$, [ES_I_J 846ms]vs[ES_I_S 1149ms] PAIRED $t(82)=4,61$ $p < 0,0001^{***}$). Assinala-se que o uso do conectivo *já que* também resultou em médias significativamente menores em contexto de causa direta comparativamente à presença do conectivo *porque* ([ES_D_J 916ms]vs[ES_D_P 1073ms]

PAIRED $t(82)=2,82$ $p < 0,006^{***}$). Esse resultado embora não tenha sido esperado não se opõe às previsões levantadas.

No grupo de EF, por outro lado, as médias de tempo de fixação de primeiro passe para a causa indireta +PORQUE em oposição à causa indireta+JÁ QUE não apresentaram diferenças significativas entre si ([EF_I_J 1232ms]vs[EF_I_P 1282ms] PAIRED $t(82)=0,79$ $p < 0,42^{ns}$). Na causa direta, porém, o conectivo JÁ QUE resultou em médias de fixação de primeiro passe significativamente menores em relação ao uso do conectivo PORQUE ([EF_D_J 1028ms]vs[EF_D_P 1215ms] PAIRED $t(82)=3,09$ $p < 0,002^{**}$). Ainda, observou-se que para o grupo de EF as médias de fixação de primeiro passe foram significativamente menores na presença de conectivos em ambos os contextos([EF_D_J 1028ms]vs[EF_D_S 1338ms] PAIRED $t(82)=4,54$ $p < 0,0001^{***}$, [EF_I_J 1232ms]vs[EF_I_S 1529ms] PAIRED $t(82)=4,03$ $p < 0,0001^{***}$, [EF_I_P 1282ms]vs[EF_I_S 1529ms] PAIRED $t(82)=3,14$ $p < 0,002^{**}$). Distintamente do grupo de EF, a leitura das sentenças com causa indireta e conectivo *porque* resultou em médias de fixação de primeiro passe semelhantes às médias das sentenças de causa indireta sem conectivo ([ES_I_P 1127ms]vs[ES_I_S 1149ms] PAIRED $t(82)=0,40$ $p < 0,69^{ns}$) para o grupo de ES, ou seja, não houve diferença significativa entre esses pares.

A comparação entre os grupos (*between subjects*), como esperado, resultou em médias significativamente mais elevadas, em quase todos os contextos pareados, para o grupo de EF: [EF_D_P 1215ms]vs[ES_D_P 1074ms] $t(164)=2,45$ $p < 0,01^*$, [EF_D_S 1338ms]vs[ES_D_S 1027ms] $t(164)=5,28$ $p < 0,0001^{***}$, [EF_I_J 1232ms]vs[ES_I_J 847ms] $t(164)=6,30$ $p < 0,0001^{***}$, [EF_I_P 1282ms]vs[ES_I_P 1127ms] $t(164)=2,37$ $p < 0,01^*$, [EF_I_S 1529ms]vs[ES_I_S 1179ms] $t(164)=5,08$ $p < 0,0001^{***}$.

4.2.3 Tempos de regressão de primeiro passe a partir da oração causal

A ANOVA identificou efeito principal significativo para os três fatores: **grupo** ($F(1,202) = 19,1$ $p < 0,00002^{***}$), **causa** ($F(1,202) = 17,3$ $p < 0,00004^{***}$) e **conectivo** ($F(2,404) = 4,52$ $p < 0,01^*$). A análise também apontou efeito de interação significativo entre os fatores **grupo e conectivo** ($F(2,404) = 6,73$ $p < 0,001^{**}$), **causa e conectivo** ($F(2,404) = 4,70$ $p < 0,009^{**}$) e entre os **três fatores** ($F(2,404) = 5,00$ $p < 0,007^*$), porém os fatores **grupo e causa** resultaram em efeito de interação marginal ($F(1,202) = 3,78$ $p < 0,05$). Identificaram-se, portanto, diferenças significativas entre os fatores analisados. O grupo de ES apresentou médias de regressão de primeiro passe mais elevadas para a condição de causa indireta/PORQUE, exibindo distinção significativa no pareamento com sua contraparte sem conectivo ([ES_I_P 2481ms]vs[ES_I_S 2122ms] PAIRED $t(101)=2,71$ $p < 0,007^{***}$). A causa indireta/PORQUE também exibiu médias significativamente mais elevadas em oposição à causa direta/PORQUE ([ES_D_P 1979ms]vs[ES_I_P 2481ms] PAIRED $t(101)=4,97$ $p < 0,0001^{***}$). Por outro lado, o grupo de EF apresentou médias mais elevadas para a causa indireta sem conectivo relativamente às demais condições, exibindo distinção significativa em relação à causa indireta/JÁ QUE ([EF_I_J 1567ms]vs[EF_I_S 2238] PAIRED $t(101)=5,49$ $p < 0,0001^{***}$) e à causa direta sem conectivo ([EF_D_S 1857ms]vs[EF_I_S 2238ms] PAIRED $t(101)=2,86$ $p < 0,005^{**}$). Salienta-se que, no grupo de EF, a causa indireta/JÁ QUE resultou em média significativamente menor em relação às demais condições. Distintamente do grupo de ES, o grupo de EF também apresentou diferenças significativas entre a causa direta/JÁ QUE e indireta/JÁ QUE ([EF_D_J 1814ms]vs[EF_I_J 1567ms] PAIRED $t(101)=3,15$ $p < 0,002^{**}$) e entre a causa indireta/PORQUE e indireta/JÁ QUE ([EF_I_J 1567ms]vs[EF_I_P 2040ms] PAIRED $t(101)=4,11$ $p < 0,0001^{***}$). O pareamento entre os grupos de EF e de ES relativo aos tempos de regressões de primeiro passe exibiu médias mais elevadas para o grupo de ES, com distinção significativa entre a causa direta/JÁ QUE ([EF_D_J 1814ms]vs[ES_D_J 2036ms] $t(202)=1,98$ $p < 0,04^*$), a causa indireta/JÁ QUE ([EF_I_J 1567ms]vs[ES_I_J 2253ms] $t(202)=4,98$ $p <$

0,0001***) e a causa indireta/PORQUE ([EF_I_P 2040ms]vs[ES_I_P 2481ms] $t(202)=4,04$ $p<0,0001***$). Esse resultado vai em sentido inverso aos resultados obtidos no pareamento entre grupos das médias de tempos totais de fixação e de tempos de fixações de primeiro passe.

4.2.4 Lookbacks de segundo e terceiro passe

Os tempos médios de leituras de segundo e terceiro passe analisados se referem às áreas de interesse 1 e 2, correspondentes ao contexto prévio e à oração principal (cf. Quadro2). A ANOVA por sujeitos e por grupos dos tempos médios de *lookbacks* identificou efeito principal altamente significativo para cada um dos três fatores: **grupo** ($F(1,502) = 37,0$ $p<0,000001***$), **causa** ($F(1,502) = 25,6$ $p<0,000001***$) e **conectivo** ($F(2,1004) = 41,9$ $p<0,000001*$). Assim, é possível afirmar que houve diferenças significativas entre as médias de *lookbacks* resultante da manipulação dos fatores investigados. A análise também apontou efeito de interação altamente significativo entre os fatores **grupo e conectivo** ($F(2,1004) = 47,7$ $p<0,000001***$), entre **causa e conectivo** ($F(2,1004) = 10,1$ $p<0,00004**$) e entre os **três fatores** ($F(2,1004) = 10,3$ $p<0,00003**$), porém os fatores **grupo e causa** exibiram efeito de interação marginal ($F(1,502) = 3,65$ $p<0,05$). A seguir, apresentam-se os resultados de teste de Tukey para fins de comparação pareada.

No ES, a releitura das áreas de interesse 1 e 2 das sentenças com causa direta/JÁ QUE resultou em médias de refixação mais elevadas do que a releitura da causa direta/PORQUE ([ES_D_J 683ms]vs[ES_D_P 635ms] PAIRED $t(251)=2,00$ $p<0,04*$) e do que a causa direta sem conectivo ([ES_D_J 683ms]vs[ES_D_S 628ms] PAIRED $t(251)=2,29$ $p<0,02*$). Ainda, analisando os pareamentos no grupo de ES, observamos que a causa indireta/PORQUE exibiu médias significativamente maiores do que a causa direta/PORQUE ([ES_D_P 635ms]vs[ES_I_P 688ms] PAIRED $t(251)=2,69$ $p<0,007*$) o que foi esperado. No mesmo sentido, a causa indireta sem conectivo exibiu médias significativamente mais elevadas do que a causa direta sem conectivo ([ES_D_S 628ms]vs[ES_I_S 694ms] PAIRED $t(251)=2,28$ $p<0,02*$). Por outro lado, a condição exibindo ausência de conectivos em contexto de causa direta gerou médias significativamente mais elevadas no grupo de EF quando comparada às contrapartes com uso de conectivos ([EF_D_J 638ms]vs[EF_D_S 800ms] PAIRED $t(251)=5,12$ $p<0,0001**$, [EF_D_P 687ms]vs[EF_D_S 800ms] PAIRED $t(251)=3,80$ $p<0,0002*$). Ainda de modo adverso ao ES, o grupo de EF apresentou diferenças entre as médias dos pares direta/JÁ QUE e indireta/JÁ QUE ([EF_D_J 638ms]vs[EF_I_J 744ms] PAIRED $t(251)=2,96$ $p<0,003*$), entre direta/PORQUE e direta sem conectivo ([EF_D_P 687ms]vs[EF_D_S 800ms] PAIRED $t(251)=3,80$ $p<0,0002*$). Também de modo distinto ao ES, a causa indireta/JÁ QUE exibiu médias significativamente mais elevadas do que a causa indireta/PORQUE ([EF_I_J 744ms]vs[EF_I_P 628ms] PAIRED $t(251)=3,15$ $p<0,001**$). A causa indireta sem conectivo apresentou médias significativamente maiores do que suas contrapartes com uso de conectivos ([EF_I_J 744ms]vs[EF_I_S 994ms] PAIRED $t(251)=9,06$ $p<0,0001**$; [EF_I_P 628ms]vs[EF_I_S 994ms] PAIRED $t(251)=11,03$ $p<0,0001**$). Essa diferença não surpreende, porém cabe salientar que não foi identificada no grupo de ES. A causa indireta sem conectivo também teve médias significativamente mais elevadas que a causa direta sem conectivo ([EF_D_S 800ms]vs[EF_I_S 994ms] PAIRED $t(251)=7,01$ $p<0,0001***$), nesse caso, no mesmo sentido que o grupo de ES. Analisando os dados comparativos entre os grupos, constataram-se diferenças significativas entre as médias da causa direta sem conectivos ([EF_D_S 800ms]vs[ES_D_S 628ms] $t(502)=6,04$ $p<0,0001**$), entre a causa indireta/JÁ QUE ([EF_I_J 744ms]vs[ES_I_J 673ms] $t(502)=2,43$ $p<0,01*$) e entre a causa indireta sem conectivo ([EF_I_S 994ms]vs[ES_I_S 694ms] $t(502)=10,70$ $p<0,0001*$). Entre esses pareamentos, o EF exibiu as médias significativamente mais elevadas do que o grupo de ES.

Por fim, apresentamos os resultados dos índices percentuais de erros por grupo de participantes à pergunta interpretativa (Tabela 2).

Tabela 2 – Índices percentuais por grupo de erros à pergunta interpretativa final.

Grupo	Condição					
	DJ	DP	DS	IJ	IP(*)	IS
EF	14%	18%	21%	73%	78%	69%
ES	17%	19%	23%	68%	58%	66%

Para acertar a pergunta relativa às sentenças com relação de causa direta, os participantes deveriam responder SIM, e, para acertar as perguntas referentes às sentenças com relação de causa indireta, os participantes deveriam responder NÃO. A previsão foi a de que o grupo de EF apresentasse índice de erros significativamente maiores do que o grupo de ES relativamente à causa indireta. Porém, na comparação entre os grupos, observou-se que apenas a causa indireta/PORQUE resultou em índices de erros significativamente mais elevados para o grupo de EF do que para o grupo de ES ([EF_I_P 78%]vs[ES_I_P 58%] $t(238)=3,24$ $p<0,001^{***}$).

Para analisar as diferenças entre os índices de erros, realizamos uma ANOVA bivariada que identificou efeito altamente significativo para o fator **causa** ($F(1,238) = 318$ $p<0,00001^{***}$) e efeito de interação significativo entre os fatores **grupo e causa** ($F(1,238) = 4,10$ $p<0,04^*$). Na análise pareada, resultante de teste de Tukey, identificou-se, para ambos os grupos, que a causa indireta teve significativamente mais erros do que às contrapartes de causa direta ([ES_D_J 19%]vs[ES_I_J 58%] PAIRED $t(119)=8,59$ $p<0,0001^{***}$; [ES_D_P]vs[ES_I_P] PAIRED $t(119)=6,33$ $p<0,0001^{***}$; [ES_D_S 23%]vs[ES_I_S 66%] PAIRED $t(119)=6,95$ $p<0,0001^{***}$, [EF_D_J 14%]vs[EF_I_J 73%] PAIRED $t(119)=11,32$ $p<0,0001^{***}$; [EF_D_P 18%]vs[EF_I_P 78%] PAIRED $t(119)=11,22$ $p<0,0001^{***}$; [EF_D_S 21%]vs[EF_I_S 69%] PAIRED $t(119)=8,01$ $p<0,0001^{***}$).

Relativamente à variável tempo de resposta à pergunta interpretativa final (Tabela 3), as médias do grupo de EF foram significativamente mais elevadas do que do grupo de ES. A condução de teste de ANOVA por sujeitos identificou efeito principal altamente significativo, relativamente aos fatores **grupo** ($F(1,224) = 51,8$ $p<0,000001^{***}$) e **causa** ($F(1,224) = 48,9$ $p<0,000001^{***}$). Também se identificou efeito de interação significativo entre os fatores **causa e grupo** ($F(1,224) = 18,7$ $p<0,00002^{***}$).

Tabela 3 – Médias de tempo de resposta, em milésimos de segundos, referentes à pergunta interpretativa final.

Grupo	Condição					
	DJ(*)	DP(*)	DS(*)	IJ	IP(*)	IS(*)
EF	4313	4232	4602	4305	4712	4700
ES	3478	3412	3537	4360	4218	4270

Na análise pareada, resultante de teste de Tukey, identificamos que na comparação entre os grupos (*between subjects*), apenas a condição indireta+JÁ QUE não exibiu diferença significativa entre as médias de tempo de resposta. Relativamente na comparação considerando a oposição entre as condições por grupo (*within subjects*), verificamos que ambos os grupos apresentaram diferenças significativas entre a causa direta+PORQUE e a causa indireta+porque ([EF_D_P 4232ms]vs[EF_I_P 4712ms] PAIRED $t(112)=2,81$ $p<0,005^{**}$, [ES_D_P 3412ms]vs[ES_I_P 4218ms] PAIRED $t(112)=5,41$ $p<0,0001^{***}$) em que a contraparte com

causalidade indireta exibiu as maiores médias. O ES também apresentou também diferenças significativas entre a causa direta+JÁ QUE e indireta+JÁ QUE ([ES_D_J 3478ms]vs[ES_I_J 4360ms] PAIRED $t(112)=4,14$ $p < 0,0001^{***}$), cujas médias mais elevadas recaíram sobre a causa indireta. O grupo de EF exibiu diferenças significativas entre a causa indireta+JÁ QUE e indireta+PORQUE ([EF_I_J 4305ms]vs[EF_I_P 4712ms] PAIRED $t(112)=2,47$ $p < 0,01^*$) em que obteve-se médias significativamente mais elevadas para a contraparte com PORQUE. Por outro lado, o grupo de ES não apresentou diferenças significativas entre esses pares. O grupo de EF também apresentou diferenças significativas entre a causa direta+PORQUE e direta sem uso de conectivo e entre a causa indireta+JÁ QUE e indireta sem uso de conectivo ([EF_D_P 4232ms]vs[EF_D_S 4602ms] PAIRED $t(112)=2,56$ $p < 0,01^*$, [EF_I_J 4305ms]vs[EF_I_S 4700ms] PAIRED $t(112)=2,26$ $p < 0,02^*$). Em ambos os contextos de pareamento, a condição sem conectivo, independentemente do tipo de causa, resultou em médias de tempo de resposta significativamente mais elevadas, como previsto. O grupo de ES, por sua vez, não apresentou diferenças significativas entre esses pares. Distintamente do grupo de EF, o grupo de ES exibiu diferenças significativas entre a causa direta e indireta sem uso de conectivo ([ES_D_S 3557ms]vs[ES_I_S 4270ms] PAIRED $t(112)=3,93$ $p < 0,0001^{***}$). Como esperado, a causa indireta exibiu médias de tempo de resposta significativamente maiores do que a causa direta nesse contexto de pareamento.

5 Discussão

Os resultados indicaram que de fato as relações causais indiretas representam maior custo de processamento para ambos os grupos de participantes. A partir dos resultados, também foi possível verificar que estudantes de EF se baseiam mais nos conectivos presentes nas sentenças durante todo o processamento, enquanto os participantes de ES utilizam informações desses itens de forma mais significativa no estágio inicial do processamento. Assim, é possível afirmar que a habilidade em leitura reflete no modo como relações de causa são processadas, porém não foi possível avaliar se essas diferenças observadas se refletem na compreensão. Os índices de acurácia, ao contrário do esperado, não apontaram divergências significativas em relação entre os grupos no sentido esperado. Apenas houve diferença significativa entre os índices de erros à resposta da pergunta interpretativa nas condições de causa indireta+PORQUE. É possível considerar a hipótese de que a própria estrutura da pergunta (e.g., *Obter sucesso no exame fez a aluna escrever textos bons?*) pode ter gerado um artefato experimental, resultante de um efeito *good-enough* (FERREIRA; BAILEY; FERRARO, 2002). O chamado efeito *good-enough*⁹ ocorre por um equívoco de análise estrutural em que o processamento parcial de sentenças gera interpretações semânticas incompletas que a não refletem o conteúdo de seu significado (FERREIRA; BAILEY; FERRARO, 2002). Como estratégia, o *parser* acaba indo por caminhos que conduzem a reanálises passíveis de interpretações implausíveis. O verbo *fazer* é prototipicamente causativo de sujeito agentivo, logo a estrutura sujeito+fazer+complemento pode ter levado ao efeito citado. Há ainda o fato de que para acertar as perguntas indiretas os participantes deveriam responder NÃO, o que é amplamente aceito como mais difícil do que responder SIM. Contudo, salienta-se que a diferença significativa entre os grupos relativamente ao índice de acurácia da causa indireta+PORQUE é um indício relevante que aponta para maior atenção crítica na compreensão da leitura por parte de ES, mesmo que nos demais contextos de causa indireta esse mesmo grupo não tenha apresentado resultado satisfatório como foi esperado. Por outro lado, as análises das médias de tempo de resposta indicaram, no sentido esperado, que a causa indireta gerou maior custo de processamento, uma vez que ambos os grupos demoraram

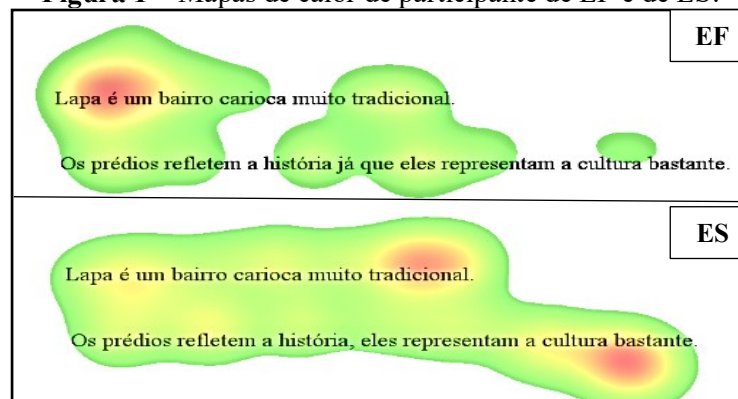
⁹ Em tradução literal ‘suficientemente bom’.

significativamente mais para responder as perguntas relativas a esse tipo de causalidade do que às perguntas relativas à causa direta. No entanto, o fato de os participantes terem que negar, ao invés de afirmar, pode ter enviesado o resultado.

Como dito no início deste texto, ambos os tipos de relação de causa exigem inferências necessárias para que a coerência entre os significados das orações seja estabelecida e o leitor possa compreender a sentença. Contudo, a relação de causa indireta toma a oração causal como um elo de ligação com a causa real que não está explícita, mas deve estar disponível no conhecimento de mundo do leitor. Esse processo tem um nível a mais de informação que precisa ser processada, o que conseqüentemente gera custo a mais em relação à causa direta. Os resultados também indicam que a proposta de Lobo (2003) para especificação de traço[+pressuposicional] para o conectivo *já que* encontra correlato empírico uma vez que a presença desse item facilitou mais o processamento tanto da causa direta quanto da causa indireta para ambos os grupos do que o conectivo *porque*. Como também argumentado por Lobo (2003), o conectivo *porque* é subespecificado no léxico para traço[+pressuposicional] e, embora não resulte em perdas tão significativas quanto a ausência de conectivo, não é um elemento de efeito facilitador do processamento do mesmo modo que o conectivo *já que*.

Uma análise qualitativa dos padrões de monitoramento ocular também possibilita afirmar, no mesmo sentido de Maia (2018a), que alguns participantes, ao se depararem com dificuldades na leitura, abandonam a tarefa. Na Figura 1, observa-se o contraste entre os mapas de calor extraídos de resultados de participante do grupo de EF e do grupo de ES.

Figura 1 – Mapas de calor de participante de EF e de ES.



Relativamente às distinções entre os grupos, é possível assumir que o grupo de ES apresentou leitura mais automatizada do que o grupo de EF. Segundo Morais (2013), um indivíduo para ser considerado letrado necessariamente desempenhar leitura e escrita de forma automática. Para ele, automaticidade em leitura e escrita se refere ao acesso imediato e direto a formas fonológicas e ortográficas de todas as palavras, isoladamente ou inseridas em um texto, que se alinhe ao seu nível de conhecimento da língua. Esse acesso às representações de palavras ou partes delas e ao modo como são estruturadas se dá de forma inconsciente. Morais (2017) afirma que “a vantagem da automaticidade vem do fato de ela libertar os recursos cognitivos conscientes para a compreensão, a reflexão e a memorização durante a leitura, e para a concepção do texto e a estruturação dos seus diferentes níveis na escrita.” Morais e Kolinsky (2018) argumentam que essa automatização é necessária para que as pessoas venham a efetuar uma leitura interpretativa e crítica e uma escrita argumentativa e original, uma vez que possibilita o exercício das capacidades cognitivas de atenção, reflexão, memória, planificação e a aquisição de conhecimentos e de métodos de análise e síntese da informação necessário (p.7). Assim, é possível que a dificuldade de estruturação das sentenças na leitura por parte dos participantes do EF tenha os levado a abandonar a leitura. Esse comportamento também

explicaria as medidas significativamente menores da condição de causa indireta+JÁ QUE, direta+ JÁ QUE e indireta+ PORQUE comparativamente com o grupo de ES exibidas pelas médias de regressão de primeiro passe. Essa medida é considerada de caráter reflexo, mas também pode ser informativa sobre procedimentos de reanálises em estágio inicial do processamento da leitura, uma vez que mensura as refixações antes mesmo de o participante terminar a leitura da área de interesse, no caso do presente experimento, a região referente à oração causal. Observando os dados referentes aos *lookbacks*, vemos que esse padrão de médias de reanálise se inverte, com isso o ES passa a exibir as menores médias. Isso é um indicativo de que, na fase inicial de processamento, os participantes de ES executaram as reanálises necessárias para a computação e compreensão da sentença, enquanto o grupo de EF ainda estaria tentando formar uma estrutura possível de ser interpretada.

6 Considerações Finais

Este trabalho teve como objetivo investigar o processamento da leitura de sentenças que estabelecem coerência de causa de forma direta e indireta por meio processos inferenciais distintos, avaliando a influência dos conectivos *já que* e *porque* ao serem utilizados como articuladores das orações. Também buscamos avaliar de que modo a experiência em leitura influencia o processamento de sentenças que combinam esses dois fatores (relações causais e conectivos) em sua configuração sentencial. Nesse sentido, comparamos dados do processamento da leitura obtidos por meio de rastreamento ocular de participantes de Ensino Fundamental (EF) e Ensino Superior (ES). Especulamos inicialmente que as relações de causa indireta apresentariam maior custo de processamento do que as relações de causa direta para ambos os grupos o que pode ser identificado frente aos resultados apresentados na seção anterior. A segunda hipótese levantada foi a de que a modo como os conectivos influenciam o processamento depende do nível de habilidade em leitura do indivíduo. Nesse sentido, os dados possibilitaram afirmar que o processamento dos participantes de EF se favorece mais da presença dos conectivos do que o dos participantes do grupo de ES, que parecem ser mais guiados por conhecimento de mundo. As médias significativamente mais elevadas para a condição indireta+porque comparativamente às médias da condição indireta sem conectivo, por exemplo, nas fases de reanálise, configuram evidências para essa afirmação. Por outro lado, o EF não exibiu o mesmo padrão em relação a essas condições. Isso significa que a ausência de conectivo gera mais perdas para o processamento do grupo com menos experiência em leitura (i.e., EF). Por fim, também aventamos a possibilidade de baixa habilidade em leitura estar relacionado a dificuldades de parseamento das unidades significativas, refletindo em maior custo de processamento para ambos os tipos de relações causais. Essa hipótese encontrou suporte nos resultados que indicam abandono da leitura, possivelmente diante de alguma dificuldade encontrada por parte do leitor menos experiente. Esse padrão de desengajamento na leitura já havia sido identificado por trabalhos anteriores apresentados em Maia (2018). Assim, algo que deve ser avaliado com cautela na pesquisa de processamento com grupos de participantes com pouca habilidade em leitura é a relação entre alto ou baixo custo de processamento e tempo de processamento, uma vez que o abandono da leitura gera menores médias. Os resultados apresentados do presente trabalho indicam que médias menores de tempo não significam necessariamente que o participante teve facilidade em processar o material a que fora exposto. Concluímos que os participantes de EF apresentam uma leitura ainda pouco automatizada ao contrário do identificado no grupo de ES, embora esse segundo grupo possa ter apresentado baixos índices de criticidade, tendo em vista o alto índice de erros verificado nas respostas às perguntas interpretativas finais.

Gerar inferências é um processo fundamental para a compreensão da leitura, assim, esperamos que este trabalho venha a contribuir com as pesquisas no âmbito da psicolinguística

educacional, servindo também de referência para reflexões acerca de práticas educacionais voltadas para o desenvolvimento de habilidade em leitura. Avaliamos que os resultados apresentados neste trabalho indicam que, para se tornar um indivíduo letrado, é necessário estabelecer um diálogo entre atividades que tornem salientes as unidades significativas da língua (e.g., propriedades gramaticais), mas que estejam atreladas ao desenvolvimento do conhecimento de mundo dos indivíduos, envolvendo principalmente a prática constante de atividades de leitura.

Referências

- AÇÃO EDUCATIVA; INSTITUTO PAULO MONTENEGRO. Indicador de Alfabetismo Funcional – INAF – Estudo especial sobre alfabetismo e mundo do trabalho, 2016.
- CANESTRELLI, A. R.; MAK, W. M.; SANDERS, T. J. M. Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements, **Language and Cognitive Processes**, 28:9, 1394-1413, 2013.
- CHOMSKY, N. The minimalist program. *Current Studies in Linguistics* 28. Cambridge, MA: MIT Press, 1995.
- CUNHA, C.; CINTRA, L. Nova Gramática do Português Contemporâneo. Rio de Janeiro: Nova Fronteira, 3a ed., 2001.
- FERREIRA, F.; BAILEY, K. G. D.; FERRARO, V. Good-enough representations in language comprehension. *Current Directions*. In: **Psychological Science**, 11(1), 11–15, 2002. <https://doi.org/10.1111/1467-8721.00158>
- GARROD, S.C.; SANFORD, A.J. Resolving sentences in a discourse context: How discourse representation affects language understanding, 1994
- HAEGEMAN, L. Conditionals, factives and the left periphery. **Lingua**, 116(10), 1651–1669, 2006. <https://doi.org/10.1016/j.lingua.2005.03.014>
- HYÖNÄ, J.; LORCH, R. F.; KAAKINEN, J. K. Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. **Journal of Educational Psychology**, 94(1), 44–55, 2002. <http://dx.doi.org/10.1037/0022-0663.94.1.44>
- KLEIMAN, A. Oficina de Leitura: Teoria e Prática, Campinas: Pontes, 15, 2013.
- LOBO, M. Aspectos da Sintaxe das Orações Subordinadas Adverbiais do Português. Tese (Doutorado) Universidade de Lisboa, Lisboa, 2003. (n.d.).
- MAIA, M. A. **Psicolinguística e educação**. Campinas: Mercado de Letras, 2018.
- _____. Computação Estrutural de Conjunto na Leitura de Períodos: Um Estudo de Rastreamento Ocular. In: _____. (org.) **Psicolinguística e educação**. Campinas: Mercado de Letras, 103-132, 2018.
- MEDEIROS, L. *Processamento das construções de tópico estilo-Chinês no Português do Brasil*. In: XI Congresso Internacional da ABRALIN 2019. Apresentação de Trabalho/Pôster, Maceió, Alagoas. 2019.
- MODESTO, M.; MAIA, M. *Representation and Processing of the Inflected Infinitive in Brazilian Portuguese: an eye-tracking study*. **Revista de Estudos da Linguagem**, 25(3), 1183-1224, 2017. <https://doi.org/10.17851/2237-2083.25.3>
- MORAIS, J. 2013 Criar Leitores Para Uma Sociedade Democrática. Santa Cruz do Sul, v. 38, Especial, p. 2-28, jul. dez. 2013. <http://online.unisc.br/seer/index.php/signo>
- _____. Preparar e Formar Letras. **Jornadas Internacionais de Leitura, Escrita e Sucesso escolar**. Universidade do Minho/ Instituto de Educação, Braga, Portugal, 2017.
- MORAIS, J.; KOLINSKY, Régine. Mudar a Educação no Brasil. **Pátio**, v. 22, n. 85, p. 7-9, 2018.
- OECD. Education at a Glance 2018: **OECD Indicators**, OECD Publishing, 2018.

- _____. PISA for development assessment and analytical framework. **OECD Indicators**, OECD Publishing, 2017.
- RAYNER, K. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422, 1998. <https://doi.org/10.1037/00332909.124.3.372>
- RIZZI, L. The fine structure of the left periphery. In: **Elements of grammar**. Springer, Dordrecht, 281-337, 1997.
- RODRIGUES, P. Orações causais e estrutura informacional da sentença. **Gragoatá**, 23(46), 544-565, 2018. <https://doi.org/10.22409/gragoata.2018n46a1135>
- SANTOS, S. Relações de causa e o uso dos conectivos *porque* e *já que*: contribuições da Psicolinguística para a Educação. Tese (Doutorado), PPG-Linguística da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2020 (n.d.).
- SANTOS, S.; GARCIA, D.; MAIA, M. 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, n. 3, p. 123-137, 2019.
- SINGER, M.; BROOKE, L. (2012). Inference and reasoning in discourse comprehension. **Cognitive Pragmatics**, (c), 85–120. <https://doi.org/10.1515/9783110214215.85>
- STUKKER, N.; SANDERS, T., Subjectivity and prototype structure in causal connectives: A cross-linguistic perspective. **Journal of Pragmatics**, 44(2), 169–190, 2011. <https://doi.org/10.1016/j.pragma.2011.06.011>
- TANKERSLEY, K. *Threads of Reading: Strategies for Literacy Development*. Alexandria, Va: Association for Supervision and Curriculum Development, 2003.
- TRAXLER, M. J.; BYBEE, M. D.; PICKERING, M. J. Influence of connectives on language comprehension: Eye tracking evidence for incremental interpretation. **The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology**, 50(3), 1997b, 481–497. <http://doi.org/10.1080/027249897391982>
- VAN DEN BROEK, P.; LORCH JR, R.F. Network representations of causal relations in memory for narrative texts: Evidence from primed recognition. **Discourse processes**, v. 16, n. 1-2, p. 75-98, 1993.

Submetido em 22/10/2020

Aceito em 08/02/2021