

**A INFLUÊNCIA DA ALFABETIZAÇÃO NAS COGNIÇÕES DE
RECONHECIMENTO DE FACES E DOS SINAIS DA FALA: UM ESTUDO
PSICOLINGUÍSTICO**

**THE INFLUENCE OF LITERACY ON FACE RECOGNITION AND SPEECH
SIGN COGNITIONS: A PSYCHOLINGUISTIC STUDY**

Isadora Rodrigues de Andrade ¹
Aniela Improta França ²

RESUMO: Testes de imagem cerebral mostram que depois da alfabetização, durante o reconhecimento de grafemas, surge uma ativação na região ventro-occipito-temporal esquerda: a VWFA. Acredita-se que a VWFA impacte o mosaico de regiões vizinhas envolvidas no reconhecimento de objetos, lugares, partes do corpo e faces. Ela também influencia a região que abriga o reconhecimento dos sinais da fala. Este estudo testa crianças nas fases pré e pós alfabetização, procurando correlatos comportamentais para esses achados neurofisiológicos. Encontramos impacto positivo da alfabetização no processamento da fala, mas, perda provisória de acuidade no reconhecimento de faces dos participantes recém-alfabetizados.

PALAVRAS-CHAVE: Aquisição da Leitura. Reciclagem Neuronal. Reconhecimento de Face. Processamento da Fala. Psicolinguística.

ABSTRACT: Brain imaging tests show that after literacy, an activation in the left ventro-occipito-temporal region occurs in the brain during grapheme recognition. This region, known as VWFA (Visual Word Form Area) is believed to impact the mosaic of neighboring regions involved in the recognition of objects, places, body parts and faces. It also influences the region that hosts the recognition of speech signals. This study tests children in the pre- and post-literacy phases, looking for behavioral correlates for these neurophysiological findings. We found positive impact of literacy on speech processing, but provisional loss of acuity in recognizing faces of new readers.

KEYWORDS: Reading Acquisition. Neuronal Recycling. Face Recognition. Speech Recognition. Psycholinguistics.

1 Introdução

Como bem podemos calcular, a habilidade para leitura não se desenvolve espontaneamente na criança como um traço inato. A literacia³ é facultativa, isto é, não obrigatória e é normalmente adquirida na infância, ou até mais tarde, mediante instrução formal (DEHAENE, 2012; MORAIS, 2013). O presente trabalho dedicou-se em aferir experimentalmente os efeitos comportamentais dessa aquisição sob aspectos de outras

¹ Doutoranda em Linguística no Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro (UFRJ). Rio de Janeiro, Brasil. ORCID: <https://orcid.org/0000-0002-5822-3501>. E-mail: andradeisadorar@gmail.com

² Professora Associada do Departamento de Linguística da Universidade Federal do Rio de Janeiro (UFRJ) e diretora Adjunta da Pós-Graduação da Faculdade de Letras (UFRJ). ORCID: <https://orcid.org/0000-0002-3614-2970>. E-mail: aniela@gmail.com

³ Segundo Morais, literacia é “o conjunto das habilidades da leitura e da escrita (identificação das palavras escritas, conhecimento da ortografia das palavras, aplicação aos textos dos processos linguísticos e cognitivos de compreensão).” (MORAIS, 2013, p.30).

duas importantes cognições espontâneas que se estabelecem mais precocemente nas crianças: reconhecimento de faces e reconhecimento dos sinais da fala.

A esse respeito, estudos no campo da neurociência têm revelado que, ainda que o sistema visual humano para detecção de bigramas ou, mais abrangentemente, de grafemas, não tenha tido tempo evolutivo significativo para se inscrever no DNA da espécie, há uma área no cérebro na região occípito-temporal ventral, que passa a responder preferencialmente às palavras escritas depois da alfabetização (cf. Dehaene 2012, para uma revisão). Essa área ficou conhecida como Área da Forma Visual da Palavra (cf. Fig. 1), do inglês *Visual Word Form Area* (VWFA). A VWFA aparentemente se especifica para detecção de grafemas, antes do pareamento com os fonemas (COHEN *et al.*, 2002; MCCANDLISS; COHEN; DEHAENE, 2003; COHEN; DEHAENE, 2004).

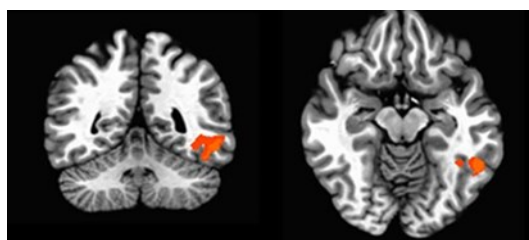


Figura 1: Área da Forma Visual da Palavra em laranja (Visual Word Form Area - VWFA) adaptado de ZHOU *et al.*, 2019, p. 1, vista através de dois cortes coronais mostrando implicação do córtex occípito-temporal ventral lateralizado (vOT esquerdo). Ativação registrada por fMRI, durante a leitura de palavras.

Outras pesquisas mostram ainda que o surgimento da VWFA impacta todo o mosaico de regiões ventrais vizinhas envolvidas no processamento de outras cognições visuais, como reconhecimento de objetos, lugares, parte do corpo e faces e também de uma região ventro-temporal que abriga o reconhecimento dos sinais da fala. Ou seja, aprender a ler interfere na anatomia e na fisiologia de uma importante região cerebral (DEHAENE *et al.*, 2010; DEHAENE, 2005; DEHAENE *et al.*, 2015).

O surgimento da VWFA revela um processo de competição por espaço no córtex visual de neurônios pouco especializados, já presentes, e de neurônios que vão se especializando gradualmente no curso da alfabetização. Esse novo córtex, fruto da plasticidade cerebral, assume um arranjo modular diferente depois da escolarização, como é exemplificado no esquema na Figura 2.

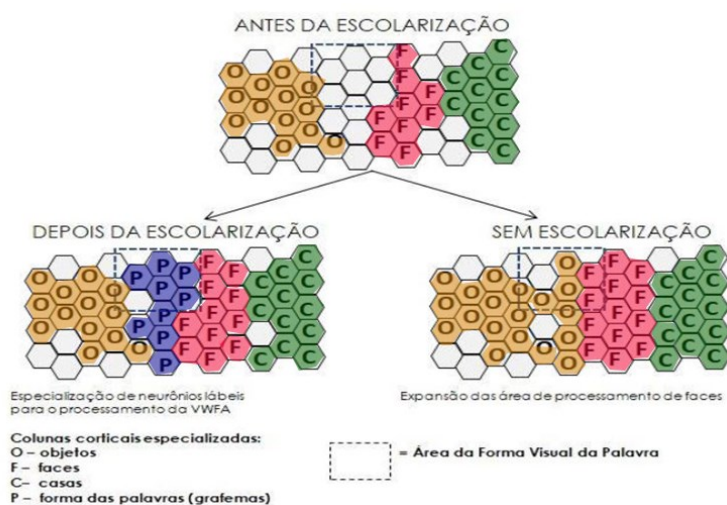


Figura 2: Modelo esquemático do desenvolvimento visual ventral. Os hexágonos representam colunas corticais especializadas para uma determinada categoria (objetos, faces, casas) ou então pequenos grupos de neurônios em um

estado lábil, não comprometido. Cada grupo de hexágonos esquematiza o estado do mosaico visual ventral do hemisfério esquerdo em uma determinada idade. Antes da escolarização (na imagem mais acima, à esquerda), algumas colunas já estão comprometidas com uma determinada categoria, com uma organização lateral-mesial sistemática (esquerda-direita), mas muitas ainda são instáveis. Quando a escolaridade começa, algumas colunas se comprometem com as palavras escritas (representadas pela letra “P”, no canto inferior esquerdo, em roxo). Na ausência de escolaridade, as mesmas colunas são progressivamente invadidas por representações próximas de objetos e faces (no canto inferior direito, ilustrados pelas letras O e F e as cores amarelo e rosa, respectivamente). A caixa tracejada ilustra como um único fMRI voxel pode compreender uma mistura de respostas corticais para objetos e faces e palavras. O modelo pode explicar por que: (1) em crianças alfabetizadas, os voxels VWFA tornam-se seletivos às palavras, mantendo sua resposta prévia aos objetos; (2) a alfabetização bloqueia a expansão das respostas da face no hemisfério esquerdo, restringindo seu crescimento ao hemisfério direito; e (3) em analfabetos, em relação aos alfabetizados, as respostas faciais são maiores no hemisfério esquerdo e, portanto, menos assimétricas em favor do hemisfério direito. P, palavras escritas; F, faces; RMf, ressonância magnética funcional; C, casas; hexágono vazio, lábil; O, objetos; VWFA, área visual da palavra (Adaptado de DEHAENE-LAMBERTZ; MONZALVO; DEHAENE, 2018, p.26).

Pesquisas realizadas tanto com metodologias hemodinâmicas quanto com métodos eletrofisiológicos convergem no achado de que a Área da Forma Visual da Palavra estabelece conexões com as áreas dedicadas ao tratamento da linguagem. Sua localização excepcionalmente estratégica permite que a área das palavras escritas se conecte eficientemente às regiões do Plano Temporal no hemisfério esquerdo, envolvidas no processamento da linguagem falada (DEHAENE *et al.*, 2010; DEHAENE *et al.*, 2015).

Note-se que a aprendizagem da leitura gera mudanças na organização anatômica cerebral (DEHAENE *et al.*, 2010). Mais precisamente, a VWFA cria projeções anatômicas bidirecionais com o giro temporal superior, que comporta notadamente representações neurais de fonemas consonantais e vocálicos (CHANG *et al.*, 2010; MESGARANI *et al.*, 2014), conforme ilustra a figura 3.

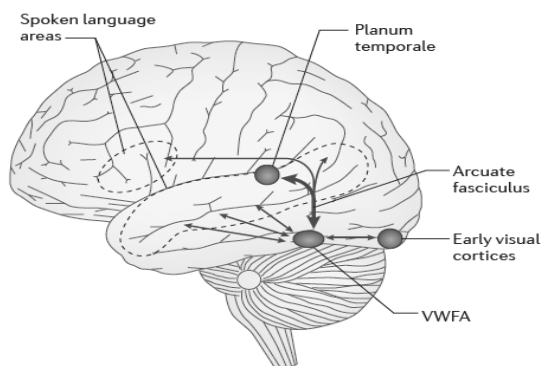


Figura 3: Representação da interface entre áreas do sistema visual e da linguagem falada. As linhas tracejadas ilustram todas as regiões envolvidas no processamento da fala antes da alfabetização. Foi demonstrado que a área da forma visual da palavra (VWFA), os córtices visuais iniciais e o plano temporal (os pequenos círculos) vão sendo refinados no decorrer da aquisição da leitura. As conexões entre essas regiões, representadas pelas setas, também podem ser aperfeiçoadas (DEHAENE *et al.*, 2015, p.4).

Uma quantidade expressiva de pesquisas comportamentais sugerem diferenças significativas na percepção e no reconhecimento e em análises conscientes dos sinais de fala antes e depois a alfabetização (KOLINSKY; PATTAMADILOK; MORAIS, 2012). Indivíduos alfabetizados se mostraram consistentemente mais acurados do que sujeitos não alfabetizados em tarefas como contagem de fonemas (ENHI; WILCE, 1979), detecção de rima (SEIDENBERG; TANENHAUS, 1979), apagamento e manipulação de fonemas ortograficamente transparentes (CASTLES *et al.*, 2003) e até numa tarefa de julgamento explícito das informações de tom, no nível suprasegmental (PATTAMADILOK *et al.*, 2008).

Estudos da neurociência apontam ainda que os efeitos da alfabetização não se limitam à língua, mas também afetam os processos visuais não-linguísticos (KOLINSKY; VERHAEGHE, 2011). No local da VWFA, aprender a ler parece instaurar uma competição com a representação cortical de objetos visuais não linguísticos. Isso significa que, à medida que a VWFA, no hemisfério esquerdo, torna-se cada vez mais sensível às sequências de letras, no momento que os indivíduos adquirem a leitura; ela se torna cada vez menos sensível aos rostos, que se tornam mais lateralizado para o hemisfério direito, em relação aos alfabetizados (DEHAENE *et al.*, 2010). Dessa forma, estudos sugerem que a dominância significativa do hemisfério direito para o tratamento visual das faces pode ter suas causas, em grande medida, na aquisição de símbolos escritos que ocorre no hemisfério esquerdo (DEHAENE *et al.*, 2015).

Apesar de amplamente aferidos neurofisiologicamente, os estudos que avaliam as possíveis consequências comportamentais dessa lateralização ainda são escassos. Uma importante contribuição foi fornecida pelo trabalho de Paulo Ventura e colegas (2013). Utilizando o paradigma de faces compostas, que avalia sob quais condições uma face pode ser processada holisticamente, os autores testaram a percepção facial de indivíduos letrados, ex-iletrados e iletrados. Os resultados revelaram que os iletrados apresentaram uma percepção das faces consistentemente mais holística, considerada mais confiável (EYSENCK; KEANE, 2017), do que os alfabetizados. Os autores sugerem que este comportamento esteja relacionado à forte dominância do hemisfério esquerdo para mecanismos holísticos e o direito para processos mais analíticos, correlacionando assim, seus resultados à reorganização hemisférica observada.

Embora os estudos anteriormente citados tenham se dedicado à investigação dos efeitos comportamentais da aquisição da leitura, a maioria deles foi realizado com adultos e faz parte da literatura internacional. Além disso, nenhum estudo observou as consequências da aprendizagem da leitura de forma integrada a outros aspectos cognitivos. Na presente pesquisa, avaliamos sincronicamente os efeitos comportamentais da alfabetização sob aspectos do reconhecimento de faces e dos sinais da fala. Para dar conta desse complexo estado de coisas, utilizando uma metodologia experimental *off-line* bastante simples, aplicamos uma série de dois testes psicolinguísticos em turmas de pré-escola, alfabetização e ensino fundamental.

A primeira tarefa consistiu no reconhecimento auditivo de sons silábicos, seguido por uma etapa de decisão lexical. Para os bebês, desde o nascimento, as sílabas que formam as palavras são uma realidade perceptual (VAN OOIJEN *et al.*, 1998). A realidade silábica tem sido detectada em bebês tanto para sílabas fortes como fracas em palavras isoladas, apesar, é claro de que as sílabas átonas se comportam de forma diferente das tônicas nas sentenças: as fronteiras átonas podem se misturar com outras mais facilmente e os limites das sílabas das palavras podem não ficar tão nítidos. São exemplos dessas alterações de percepção silábica nas sentenças, então, os casos em que a primeira palavra em posição de coda silábica torna-se *onset* da primeira sílaba do vocábulo seguinte; por exemplo, arroz amarelo/arro[za]marelo em contraste com [O Paulo comprou arroz.] (Para uma revisão desses processos morfofonológicos cf. LEAL, 2015).

Esse tipo de alteração de percepção silábica pode fazer crianças postularem representações subjacentes como [zamarelo], como uma variação da palavra amarelo. Questiona-se aqui se uma vez que a criança é alfabetizada, esse tipo de alteração de percepção auditiva ainda perdura ou se a consciência fonêmica é aguçada e incita uma reinterpretação da percepção. A pergunta aqui foi se o fato de se conhecer como as palavras são grafadas oferece uma melhor percepção auditiva das fronteiras entre as

palavras? Ou seja, haveria sincretismo cognitivo, no sentido de que a percepção visual poderia vir a afetar a percepção auditiva?

A segunda tarefa consistiu no reconhecimento de faces conhecidas e desconhecidas. Os estudos sobre percepção facial convergem na ideia de que a identificação de um rosto envolve, pelo menos, dois processamentos macro: um processamento holístico ou configuracional (TANAKA; FARAH, 1993; DIAMOND; CAREY, 1986; FREIRE; LEE; SYMONS, 2000) e um processamento analítico (COLLISHAW; HOLE, 2000; LEDER; BRUCE, 2000). O processamento holístico consiste na computação integrada das informações de toda a face, como numa *gestalt*; enquanto a via analítica está envolvida no tratamento individual das características faciais. Pesquisas indicam ainda que o processamento de rostos pode ser influenciado por diversos fatores, como o Efeito de Inversão de Faces (MCKONE; KANWISHER; DUCHAINE, 2007), o uso de apetrechos (DIAMOND E CAREY, 1977) etc. Aqui, desejamos verificar se o desenvolvimento da leitura também pode estar associada a algum comprometimento na percepção facial de crianças não alfabetizadas e alfabetizadas.

Em síntese, as hipóteses que norteiam este trabalho são: (i) o conhecimento ortográfico atuará como um facilitador na missão de julgar se as sílabas faladas integram ou não as palavras ouvidas, na primeira tarefa; e na função de organizar as sílabas ouvidas para a formação de palavras, na segunda tarefa. Nesse sentido, espera-se que, quanto maior for o nível de conhecimento ortográfico do sujeito, melhor será o seu desempenho em termos de tempo de resposta e acuidade; (ii) a percepção facial dos participantes não alfabetizados será mais acurada e seu desempenho mobilizará tempos de resposta mais baixos em relação aos sujeitos com níveis de escolarização mais altos. Supõe-se que essa hipótese, se confirmada, possa trazer à luz um possível efeito comportamental relacionável ao processo de lateralização da cognição de processamento facial.

2 Os experimentos

Sabendo que a pesquisa com crianças de diferentes faixas etárias impõe grandes desafios de controle executivo (PINTO; FRANÇA, 2018), especialmente quando a tarefa experimental exige um nível de controle inibitório ou atenção, esse estudo usou uma metodologia ecológica, capaz de minimizar as chances de perdas de participantes por desengajamento ou falta de atenção à atividade.

3 Experimento I: reconhecimento auditivo de sílaba na palavra⁴

Para examinar se conhecimento ortográfico sobre a escrita das palavras é capaz de influenciar a percepção dos sons foi realizado um teste de reconhecimento auditivo de sílaba na palavra com decisão pré-lexical em crianças cursando quatro níveis diferentes escolares: Pré Escola 1, Pré Escola 2, 1º Ano e 2º Ano.

Nossa previsão era a de que as crianças alfabetizadas, cursando o 1º Ano ou o 2º Ano apresentariam um índice mais alto de acurácia na tarefa, enquanto as crianças não alfabetizadas, cursando o Pré 1 ou o Pré 2 teriam um desempenho inferior. Isso aconteceria na medida em que contrastes e fronteiras entre palavras podem ser revisitados pela criança e ganham maior clareza através da escrita, fazendo com que compreensão da fala e a própria fala fiquem mais especializadas (DEHAENE *et al.*, 2015).

⁴ Para visualizar o vídeo estímulo: https://www.youtube.com/watch?v=U65X_elp3SE&feature=youtu.be.

3.1 Metodologia

Desenho Experimental

As variáveis independentes que se desejam investigar no presente experimento foram: (i) transparência ortográfica, em dois níveis: +transparente e -transparente; e (ii) tonicidade, em dois níveis: +tônico e -tônico. Já as variáveis dependentes foram o índice de acurácia e uma medida off-line de tempo de reação para cada estímulo. O desenho fatorial do experimento foi estabelecido a partir do número de variáveis independentes e da quantidade de níveis apresentados por elas. Sabendo que as variáveis independentes apresentam dois níveis, o cruzamento entre elas produz um design fatorial 2x2, totalizando 4 condições (células amarelas), como mostra a Tabela 1.

Nelas, os itens experimentais foram dispostos em quadrado latino, garantindo que cada participante ouvisse pelo menos um *token* de cada condição e que não visse nada repetido para que os vieses de *priming* não se operem sob os resultados de processamento. Como tínhamos 4 condições, trabalhamos com 4 versões de estímulos (de 1 a 4).

| cruzamento sujeitos | [+TRANS +TON] | [+TRANS-TON] | [-TRANS-TON] | [-TRANS+TON] |
|------------------------|---------------|--------------|--------------|--------------|
| 1 | Bala | Palito | Chapéu | Casaco |
| 2 | Sapato | Lupa | Casa | Chato |
| 3 | Barro | Batata | Caça | Passado |
| 4 | Pássaro | Sopa | Caixa | Cansado |

Tabela 1: Condições obtidas a partir do cruzamento das variáveis independentes (amarelo) e os itens experimentais organizados em quadrado latino.

Cada versão foi primeiro aleatorizada através da função RANDOM do MS Excel, que gera uma sequência de números aleatórios em uma célula em cada linha de dados. A aleatorização se efetiva quando se aplica a função de ordenar do maior para o menor esses números aleatórios. As linhas com estímulos passam a obedecer à ordenação crescente aleatória e saem da organização linear das condições do experimento. Depois de chegarmos à ordem, aplicada à versão 1, a lista foi verificada em relação à ocorrência de alguma repetição indesejável, por exemplo, se por acaso três itens da mesma condição apareciam seguidos. Se houvesse alguma ocorrência desse gênero, a ordem aleatória do *Excel* era manipulada manualmente para chegar-se a uma lista balanceada, como o menor número de manipulações. Assim, chegou-se a uma ordenação pseudoaleatória para a primeira versão. Depois, a mesma sequência de números ordenada pseudoaleatoriamente foi copiada e aplicada às demais versões, de modo que não houve viés de ordenação entre as versões. Desse modo, todos os participantes tiveram acesso a todos os estímulos, mas não aos mesmos itens, mantendo-se rigorosamente a mesma sequência para cada versão, seguindo, então, uma Distribuição equilibrada Intra-Grupos (*Within-Subject Distribution*).

A tarefa do teste

Após a fase de treinamento, a experimentadora pausava o vídeo, perguntava se o participante havia entendido o funcionamento da tarefa e reiterava que, primeiramente, ele deveria ouvir com atenção a sentença ditada pela narradora (ex.: Tem ['da] em

['dadu]?). Depois, deveria mostrar se o som ouvido (ex.: ['da]) poderia ser encontrado na palavra falada (ex.: ['dadu]). A tarefa do participante seria apontar para o botão verde, caso a sílaba ouvida pudesse ser encontrada na palavra pronunciada, ou o botão vermelho, se a sílaba não pudesse ser encontrada na palavra ouvida.

Durante a estimulação auditiva, para que a atenção da criança se mantivesse no teste, foi utilizada uma âncora visual: aparecia na tela uma animação com círculos concêntricos coloridos, retirados de um vídeo-compilação de trabalhos do pintor húngaro Victor Vasarely, encontrado no site YouTube (<https://www.youtube.com/watch?v=ny1C-O37PCY&t=3s>).

Imediatamente ao fim da estimulação auditiva, a narradora/avatar reaparecia na tela. Nesse momento, um botão verde e outro vermelho surgiam ao lado dela, à direita. Esta cena era congelada por 5 segundos, tempo disponível para que os participantes fizessem sua escolha. Ao final da escolha, a narradora reaparecia para cumprimentar o participante pela escolha. A sequência do teste para um item pode ser observada na figura 4.

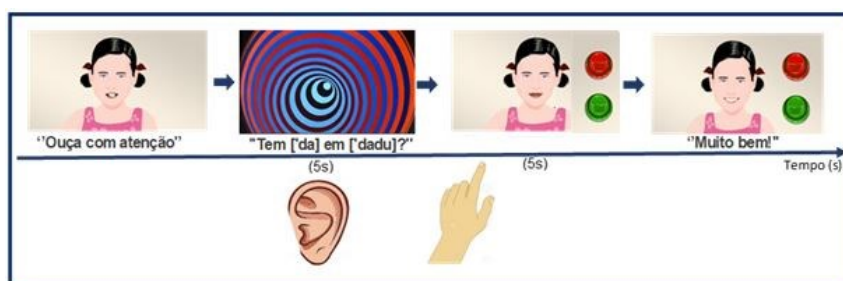


Figura 4: A cronologia do teste.

Embora a tarefa estabelecesse que o participante deveria apontar para o botão escolhido, todos eles pressionaram na tela a opção desejada. Por isso, o tempo de reação do participante foi marcado a partir do momento em que eles tocaram o dedo em um dos botões.

Tanto as respostas obtidas quanto os seus respectivos tempos foram extraídos por meio do programa de criação e edição de vídeos Wondershare Filmora 9, gratuitamente acessível no site (<https://filmora.wondershare.com/pt-br/editor-de-video/>). Os vídeos foram convertidos para o modo *slow motion* e cortados em 60 frames por segundo. Em todos os casos, o ângulo de captura do vídeo permitiu o registro preciso do momento exato em que o participante fazia a sua escolha.

Os resultados obtidos foram analisados a partir de dois tipos de medidas ou variáveis dependentes: acurácia, medida em número de acertos categóricos e tempo de resposta. Os dados categóricos de acurácia (resposta certa Vs errada) serão tratados com o teste Qui-quadrado do software MSEXcel que faz a comparação entre os grupos de frequências observadas e as frequências esperadas, observando-se um limiar de significância de $p > 0,05$. Já os dados não categóricos - tempos de resposta em milissegundos (response times - RTs) - foram tratados com o pacote EZanova, através de uma análise de variância monofatorial intrassujeitos (ANOVA: Design 1 Within Subject Factors), também observando-se um $p > 0,05$.

3.2 Materiais

Para este teste foram utilizados um conjunto de 16 sentenças-estímulo, divididas em 4 versões, resultando em 4 sentenças experimentais por lista. As frases tinham uma

estrutura padrão: “Tem X em Y?”, na qual X era uma combinação de uma grafema consonantal ortograficamente transparente, ou seja, que apresenta uma relação direta com o fonema correspondente, e uma vogal. Foram utilizados apenas os grafemas *P* e *B* e a vogal *A*. Por sua vez, Y representa a palavra a ser pareada à sílaba ouvida.

Além disso, foi usado um mesmo grupo de 4 sentenças distratoras, frases inseridas com a finalidade de “distrair” o participante, de modo que ele não perceba o que está sendo verificado no experimento. Para compor esta categoria, foram utilizadas pseudopalavras como *tifi*, *bogo* e outras (cf. Tabela 1).

3.3 Participantes

O teste foi aplicado a quatro grupos distintos: 8 crianças de Pré-Escola 1 (5 meninos e 3 meninas), com idade média de 4,71 anos (desvio padrão 0,49); 8 crianças de Pré-Escola 2 (5 meninos e 3 meninas), com idade média de 5,25 anos (desvio padrão 0,46); 8 crianças de 1º Ano (7 meninos e 1 menina), com idade média de 6,87 (desvio padrão 0,35) e 6 crianças de 2º ano (4 meninos e 2 meninas), com idade média de 8 anos (desvio padrão 0). Todos os participantes eram alunos do colégio Centro de Educação Infantil Meu Pequeno Príncipe, localizado no Recreio dos Bandeirantes, Zona Oeste da cidade do Rio de Janeiro e realizaram os experimentos voluntariamente e seus responsáveis legais assinaram o Termo de Consentimento Livre e Esclarecido.

Foram realizadas entrevistas com as professoras responsáveis pelas turmas com o propósito de caracterizar os níveis de leitura dos grupos. Segundo os relatos das respectivas professoras, os alunos do Pré-1 são capazes de identificar as letras maiúsculas e minúsculas do alfabeto, nos tipos bastão e cursiva e fazer a correspondência com o fonema correspondente. As crianças do Pré-2 conseguem reconhecer as letras do alfabeto, agrupá-las formando sílabas e ler palavras curtas. Já o 1º ano, é capaz de realizar a leitura de frases e textos curtos, enquanto os alunos do 2º ano apresentam leitura fluente de textos curtos e longos.

Foram realizadas entrevistas com as professoras responsáveis pelas turmas com o propósito de caracterizar os níveis de leitura dos grupos. Segundo os relatos das respectivas professoras, os alunos do Pré-1 são capazes de identificar as letras maiúsculas e minúsculas do alfabeto, nos tipos bastão e cursiva e fazer a correspondência com o fonema correspondente. As crianças do Pré-2 conseguem reconhecer as letras do alfabeto, agrupá-las formando sílabas e ler palavras curtas. Já o 1º ano, é capaz de realizar a leitura de frases e textos curtos, enquanto os alunos do 2º ano apresentam leitura fluente de textos curtos e longos.

3.4 Resultados e Discussão

Nesta tarefa, avaliou-se se os fatores *transparência ortográfica* e *tonicidade*, conhecimentos habitualmente adquiridos e exercitados em sala de aula, eram utilizados por crianças durante o processamento de sons da fala, facilitando seu reconhecimento. Dessa forma, esperava-se que a influência desses conhecimentos no processamento pudesse ser refletida em um maior grau de acerto e menores tempos de reação dos participantes alfabetizados, em comparação aos não alfabetizados.

A princípio, fizemos a predição de que as crianças sendo alfabetizadas teriam mais recursos metalinguísticos e poderiam pensar sobre linguagem com mais afastamento e com maior facilidade para lidar com a tarefa que pressupõe domínio da consciência

relativa do que crianças que não lidaram com as perspectivas absolutas e relativas de forma tão explícita.

No que diz respeito à acurácia, a comparação entre as condições [+TRANS+TON]; [+TRANS-TON]; [-TRANS-TON]; [-TRANS+TON] e as turmas (Pré 1, Pré 2, 1º ano e 2º ano) não resultou em efeito principal. No entanto, foi possível observar uma forte tendência do 2º ano aos maiores índices de acerto no reconhecimento dos estímulos auditivos, enquanto o Pré 1 apresentou o pior desempenho.

Com a finalidade de analisar a influência dos fatores transparência e tonicidade, subdividimos e avaliamos tais fatores separadamente, gerando os itens [+transparente], [-transparente], [+tônico] e [-tônico]. Os resultados estão expressos na figura 5. Constatamos efeito estatístico para tonicidade apenas no nível [-tônico] ($p=0,0544$). Nesse nível, podemos observar que, novamente, os níveis de acerto mais alto e mais baixo foram atribuídos ao 2º ano e ao Pré1, respectivamente. O 1º ano apresentou uma vantagem de apenas 6% em comparação ao Pré 2, obtendo, assim, resultados proporcionais. No nível [+tônico], a maioria das turmas (Pré 2, 1º e 2º anos) apresentou um bom desempenho com exceção do Pré 1, que registrou apenas 63%.

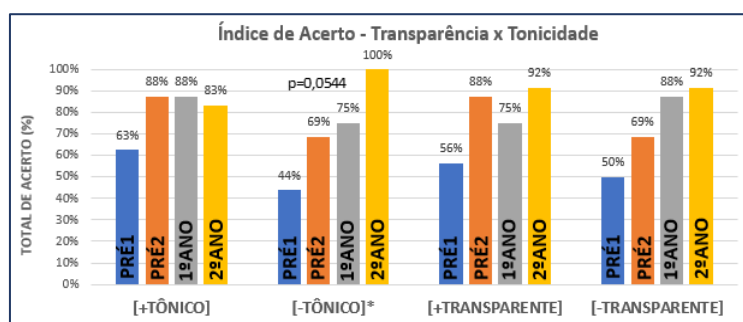


Figura 5: Índice de acerto para os fatores transparência e tonicidade das turmas testadas.

No que diz respeito ao nível [+transparente], pode-se observar que o 2º ano manteve o maior nível de precisão no reconhecimento dos estímulos, sendo seguido pelo Pré2, que obteve a segunda percentagem mais alta, 88%. O 1º ano apresentou 75% de acurácia na tarefa e o Pré1, o menor índice de acerto, somente 56%. Já nos itens [-transparente], é possível perceber uma tendência interessante: as turmas mais escolarizadas, 2º e o 1º anos, apresentam um desempenho similar, com altos índices de acerto (92% e 88%). As turmas não alfabetizadas, Pré1 e Pré2, também tiveram resultados equivalentes, mas com percentuais mais baixos, 50% e 69%.

Acreditando que os resultados relativos à condição [-transparente] possam ser consequência de uma interferência quantitativa dos dados e não refletir, necessariamente, a percepção sobre esse sons, reagrupamos as turmas em dois subconjuntos: alfabetizados (composto pelas turmas que já possuem algum conhecimento ortográfico 1º e 2º anos) e não alfabetizados (Pré-Escola 1 e 2) para os dados de acurácia, apresentados na figura 6.

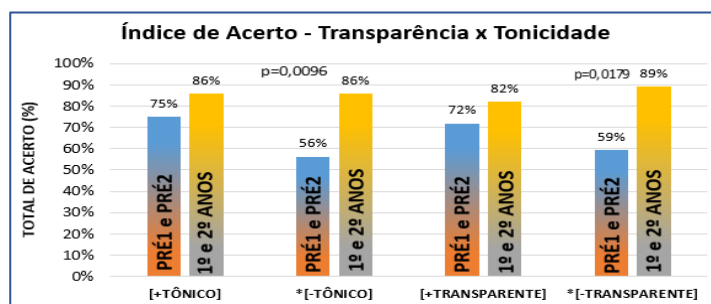


Figura 6: Índice de acerto para os fatores transparência e tonicidade das turmas agrupadas.

Pode-se notar que depois do agrupamento, foi captado efeito estatístico para o fator transparência, no nível [-transparente] $p=0,0179$. O fator tonicidade, no nível [-tônico] desempenhou, mais uma vez, um papel estatisticamente relevante no reconhecimento auditivo das sílabas dos dois grupos $p=0,0096$. Isso significa que a falta de informação sobre a intensidade das sílabas e a relação inconsistente dos fonemas consonantais e das letras correspondentes impactaram de diferentes formas as turmas avaliadas. No entanto, mesmo na ausência dessas importantes fontes de informação sobre a sílaba, os grupos do 1º e 2º anos apresentaram um desempenho significativamente melhor em comparação às turmas de Pré 1 e Pré 2.

Quanto ao tempo de resposta, obtivemos dados em milissegundos medidos a partir de controle dos fotogramas da filmagem dos testes com câmera (Webcam UltraHD 4K BRIO, Logitech). A análise quadro a quadro com o programa Filmora Wondershare 9, forneceram tempos de resposta *off-line* para o reconhecimento desses sons. Assim, as categorias transparência e tonicidade foram divididas em [+transparente], [-transparente], [-tônico] e [+tônico] e seus tempos foram analisados.

No que se refere à transparência, foi verificado efeito principal para esse fator no nível [+transparente] $F(3,36) = 4,07$ $p < 0,013702$. Na investigação par a par também foi encontrado efeito estatístico para Pré1 vs Pré2 $t(12)=3,43$ $p < 0,0050$; Pré1 vs 2º ano $t(12)=2,13$ $p < 0,0547$; Pré2 vs 1º ano $t(12)=2,37$ $p < 0,0351$ e Pré2 vs 2º ano $t(12)=2,11$ $p < 0,0567$ (figura 7).

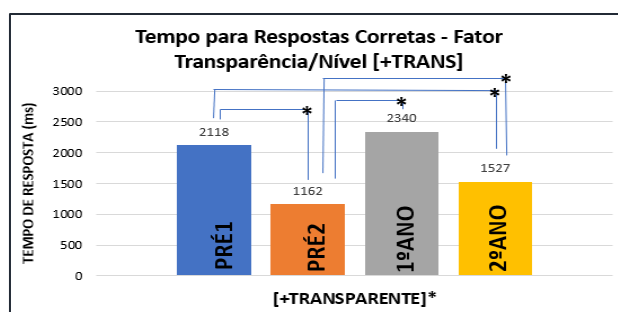


Figura 7: Tempo de reação para as respostas corretas no nível [+transparente].

O maior tempo gasto no reconhecimento dos itens [+transparente] foi registrado na turma de 1º ano, 2340ms, seguido pelo Pré 1, com tempo médio de 2118ms. Esse dado evidencia que essas condições foram mais custosas, em termos temporais, para as referidas turmas do que para as demais. Além disso, não foi verificada diferença estatística entre esses grupos, indicando que o Pré 1 e o 1º ano apresentaram um comportamento muito semelhante nessa categoria. Por outro lado, o tempo mais baixo foi atribuído ao Pré 2, 1162ms, enquanto o 2º ano ocupa uma posição regular, levando em torno de 1527ms para realizar o reconhecimento dos sons silábicos.

É interessante notar que a medida do tempo gasto no reconhecimento das sílabas em relação ao nível [-transparente] também capturou maior dificuldade dos participantes do 1º ano. Os resultados relativos a esse nível podem ser observados na figura 8, que representa o tempo gasto para as respostas corretas nas quatro turmas testadas. Foi constatada relevância estatística na comparação entre as duas turmas de 1º Ano (2121ms) e 2º Ano (1466ms), $t(11)=2,24$ $p < 0,0469$, cujos tempos foram o mais alto e mais baixo, respectivamente.

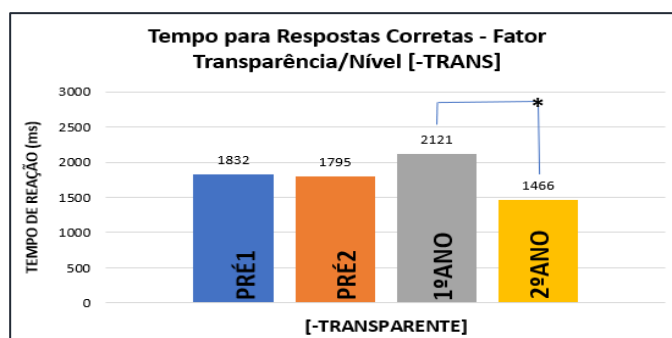


Figura 8: Tempo de reação para as respostas corretas no nível [-transparente].

A respeito do aspecto tonicidade, foi observado que os grupos estudados apresentaram comportamentos bem diferentes nos níveis [-tônico] e [+tônico]. A análise do fator tonicidade, nível [+tônico] não resultou em relevância estatística. Apesar disso, pode-se observar uma tendência do 1º ano mais uma vez apresentar o maior tempo de reação (1906ms), em comparação aos outros grupos (figura 9).

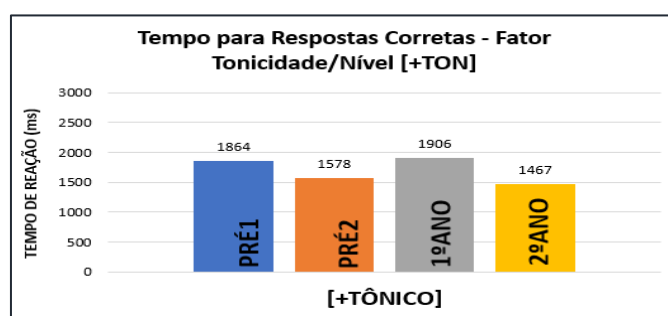


Figura 9: Tempo de reação para respostas corretas no nível [+tônico].

É interessante observar, na figura abaixo (10), o efeito estatístico obtido na análise dos tempos de resposta do fator tonicidade agora no nível [-tônico]: $F(3,33) = 3,61$ $p < 0,0018$. Na comparação par a par, constatamos efeito estatístico na comparação entre 1º ano vs 2º ano $t(11)=3,67$ $p < 0,0037$ e Pré 1 vs 2º ano $t(11)=4,08$ $p < 0,0018$. A ausência de diferença estatística entre os tempos do Pré 1 e do 1º ano, evidenciando que ambos os grupos se comportaram de forma semelhante neste nível, sendo que o Pré 2 apresenta tempos menores, é, mais uma vez, sinal de que existe um atrapalhador na performance dos participantes do 1º ano, efeito esse aqui capturado como tendência. O 2º ano, já recuperado desse momento em que as decisões foram mais lentas, acertou significativamente mais do que Pré 1 e 1º ano. Decorre disso, a conclusão de que não houve uma diminuição progressiva dos RTs, indo do Pré-1 até o 2º ano. O Pré 2 executou a tarefa em menor tempo que o 1º ano.

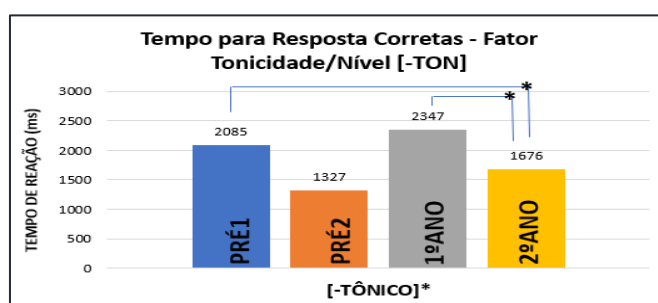


Figura 10: Tempo de reação para respostas corretas no nível [-tônico].

De fato, esse achado não é de todo inesperado. Estritamente dentro das previsões mais espetaculares da teoria de Dehaene (2005) está a especulação de que durante a alfabetização, ou seja, aqui para esse experimento, durante o 1º ano, a exposição paulatina aos grafemas e ao sistema da escrita no cérebro, na região do giro fusiforme (VWFA), avizinhada pelas representações de objetos e faces, provoque ajustes no tecido neural a ponto de afetar momentaneamente o processamento em toda aquela região, impactando a cognição de linguagem e também a cognição de reconhecimento facial. Embora se espere um aprimoramento nas tarefas de linguagem depois da alfabetização, também há a predição de uma piora transiente dessas tarefas durante o 1º ano, quando um novo sistema de decodificação grafema-fonema acontece. Um pequeno indicio disso pode ter sido capturado com esse experimento.

Os resultados do Experimento 1 revelaram que, de maneira geral, a turma alfabetizada, 2º ano, alcançou os maiores níveis de acerto. Por outro lado, o Pré 1, turma não alfabetizada, se manteve com os níveis mais baixos de acurácia. Nas condições positivas, a maioria dos grupos apresentou um bom desempenho, em torno de 80% nas condições [+tônico] e 70% nas condições [+transparente]. Isso indica a possibilidade desse conhecimento estar sendo igualmente aproveitado pelos grupos. Entretanto, nas condições nas quais a tonicidade e a transparência foram neutralizadas, somente a turma de 2º ano manteve seu alto índice de acerto, cerca de 90%, demonstrando que a ausência desse recurso não foi capaz de prejudicar de maneira significativa o processamento dos sons silábicos de crianças alfabetizadas.

Esse achado fortalece uma das abordagens que visa explicar a influência do conhecimento linguístico sob o processamento da fala. As condições em que o fator transparência foi operado poderiam ser muito problemáticas para as turmas alfabetizadas, visto que os participantes deveriam desassociar quaisquer relações que se estabelecem entre o grafema e o fonema para que o processamento auditivo não fosse perturbado. Nossos resultados demonstraram que, mesmo nessas condições mais críticas, foi verificado que o grupo alfabetizado foi substancialmente melhor no reconhecimento dos sons ouvidos. Essa constatação é coerente com a explicação de que o conhecimento ortográfico altera representações fonológicas já existentes, tornando a percepção desses sons mais aprimorada, ao invés da visão de que o desenvolvimento da alfabetização produz automaticamente representações visuais das palavras sempre que ouvidas (KOLINSKY, ATTAMADILOK; MORAIS, 2012).

Já a análise dos tempos de reação para as respostas corretas demonstrou um padrão de rapidez de processamento da turma mais escolarizada (2º Ano) nos seus julgamentos. A lentidão observada para as turmas de 1º ano em relação aos fatores transparência e tonicidade pode ser explicada como a tentativa de organização em relação às acomodações corticais que estão ocorrendo no momento da alfabetização. Isso pode indicar que a influência do novo conhecimento ortográfico sobre o processamento auditivo da fala se manifesta, pelo menos de forma imediata, apenas no nível de precisão na identificação desses sons, demandando no início ainda um alto custo temporal no processamento.

4 Experimento II: pareamento de faces estocadas na memória⁵

Este estudo investigará a percepção facial de crianças pré-alfabetizadas, em processo de alfabetização e crianças já alfabetizadas. Para tal, foi aplicado um teste de

⁵Para visualizar o vídeo estímulo: <https://www.youtube.com/watch?v=7N888YK6h0s&feature=youtu.be>.

reconhecimento de faces conhecidas e desconhecidas a grupos de diferentes níveis de escolaridade.

Para esse experimento, a previsão é a de que as crianças não alfabetizadas apresentarão tempos mais baixos e terão um maior índice de acurácia no teste de reconhecimento facial, ao passo que as crianças alfabetizadas levarão mais tempo na mesma tarefa e terão menor índice de acurácia.

4.1 Metodologia

Desenho Experimental

As variáveis independentes deste experimento são: (i) tipo de face, dividido em dois grupos: faces de pessoas *conhecidas* e *desconhecidas*; e (ii) percepção da face: *mesma pessoa* e *pessoas diferentes*. As variáveis dependentes foram índice de acurácia e uma medida *off-line* de tempo de reação para cada estímulo.

O conjunto de rostos conhecidos compreende imagens de pessoas próximas aos participantes, com as quais eles se relacionam diariamente, como suas professoras e assistentes escolares. Neste grupo, também estão imagens de celebridades. A seleção desses últimos foi feita a partir de nomes sugeridos pelas próprias crianças numa sondagem pré-teste. A categoria de faces desconhecidas é composta por faces de anônimos. Aqui também foram incluídas as imagens distratoras.

O desenho fatorial do experimento foi estabelecido a partir do número de variáveis independentes e da quantidade de níveis apresentados por elas. Considerando que ambas as variáveis independentes apresentam dois níveis cada, o cruzamento entre elas produz um design fatorial 2x2, totalizando 4 condições, como mostra a tabela 2:

| Cruzamento das variáveis | Face Conhecida | Face Desconhecida |
|--------------------------|----------------|-------------------|
| Igual | FCI | FDI |
| Diferente | FCD | FDD |

Tabela 2: Condições obtidas a partir do cruzamento das variáveis independentes.

Os estímulos utilizados no teste foram organizados de forma que todos os participantes tiveram acesso a todos os itens experimentais, mas não os mesmos itens, seguindo, então, uma distribuição intra-participantes – *Within Subject*.

Fase de Instrução

Antes do teste ser iniciado, o participante foi informado pela experimentadora que, nesta tarefa, ele deveria visualizar com atenção duas imagens, uma após a outra, e reconhecê-las como referentes à mesma pessoa ou a pessoas diferentes. Para simular a situação do teste e para ilustrar todas as condições experimentais definidas, foi preparada uma apresentação no programa *PowerPoint*, tal como exemplificado pela sequência a seguir (figura 11).



Figura 11: Sequência da apresentação em PowerPoint, exibida antes experimento.

Os participantes foram orientados a classificar os pares como pertencentes à mesma pessoa quando o indivíduo da primeira imagem (*prime*) e o da segunda imagem (alvo) forem completamente idênticos e na ocasião em que o alvo for apresentado com algum tipo de manipulação externa – modificações superficiais que não são capazes de descaracterizar a identidade visual do sujeito, como a inclusão de apetrechos (óculos) (figura 12).



Figura 12: Os pares usados para representar as condições referentes à mesma pessoa.

Os pares deveriam ser reconhecidos como relativos a pessoas diferentes quando o alvo apresentasse manipulações internas – alterações capazes de transformar características fundamentais da face, como tamanho e formato dos olhos, boca e nariz; e quando se tratasse de pessoas inteiramente distintas, como demonstrado na figura 13.



Figura 13: Os pares utilizados para representar as condições correspondentes a pessoas diferentes.

Contexto Motivacional

O vídeo que traz a tarefa experimental, reconhecer se a face de uma pessoa é igual à face que é apresentada em seguida, tem duração total de 10 m (figura 14).

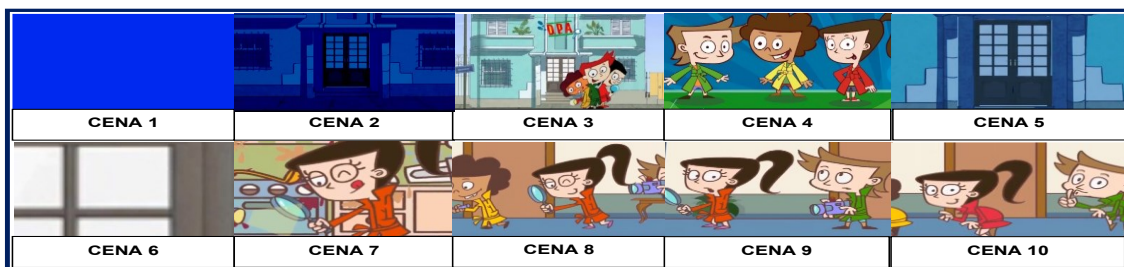


Figura 14: Cenas do contexto motivacional criado no teste e exibido antes do participante iniciar a tarefa.

Durante os primeiros 55 segundos, foi apresentado o contexto *role playing*. O vídeo começa com uma tela azul durante o período de 3 segundos (cena 1). Depois, a imagem do prédio azul é exibida como se fosse durante as primeiras horas da madrugada (cena 1). Enquanto isso, a narradora, pergunta se o participante quer ajudar os detetives a desvendar uma história misteriosa e brincar com os detetives do prédio azul (cena 2). Em seguida, é exibida uma animação curta dos detetives (cena 3). Nela, os personagens se apresentam e relatam que "coisas estranhas" estão acontecendo no prédio azul. Por isso, precisam da ajuda de seus "aprendizes". Um dos personagens ainda adverte que os participantes devem estar sempre atentos às surpresas do prédio azul. Na cena seguinte (cena 4), a imagem de prédio azul é retomada enquanto a narradora conta que, na noite passada, o prédio foi invadido. Por meio de um efeito de aproximação – zoom, o participante é transportado para dentro do prédio (cena 5) onde se encontram os detetives. A narradora chama a atenção para o estado de desordem em que se encontra o edifício (cena 6). Ela assegura que os detetives já estão investigando o caso (cena 7), mas precisam de ajuda. Nesse momento, se dirigindo diretamente ao participante, a narradora pergunta se ele deseja testar suas habilidades de detetive. A imagem é congelada (cena 8) por 3 segundos para que o voluntário possa manifestar a sua vontade de participar da tarefa. Na última cena, os detetives saem de cena enquanto a narradora convoca os participantes a um treinamento (cena 9). É um vídeo participativo com uma rotina que envolveu as crianças e proporcionou ínfimas perdas de resposta.

A tarefa do teste

Após essa parte introdutória, iniciava-se um treinamento. O treinamento era composto por dois *trials* e a dinâmica adotada foi a mesma empregada em todo o teste. Já na fase do teste, o participante era, primeiramente, avisado pela narradora sobre o sumiço de um determinado objeto pertencente a um morador fictício do prédio azul. Em seguida, o voluntário era orientado a olhar com atenção para a imagem do suspeito que aparecia na câmera de segurança do prédio. A imagem permanecia na tela por 5 segundos. Posteriormente, a câmera de segurança era substituída pela lupa dos detetives, onde a imagem do suspeito era apresentada durante um período de 5 segundos. A narradora, então, perguntava ao participante se a pessoa que tinha aparecido na câmera era a mesma pessoa que aparecia na lupa. Se a resposta fosse positiva, o participante deveria apontar para o botão verde. Caso fossem diferentes, o participante deveria apontar para o botão vermelho. Após a identificação do suspeito, o objeto sumido reaparecia em uma planilha de controle dos detetives. A narradora reagia às respostas dadas pelos participantes com expressões de incentivo, como “recuperamos o objeto perdido!”, “pegou mais um!”, “estamos recuperando tudo!” etc.

Após a validação da resposta, o objeto roubado, agora recuperado pelo participante, era levado, por meio de uma espécie de nuvem mágica, a uma grade que continha todos os objetos que deveriam ser resgatados, em preto em branco. A medida

que as coisas fossem sendo recuperadas, o objeto era substituído pela sua versão colorida. Essa tática de troca de cor foi utilizada para que os participantes pudessem perceber que os objetos haviam sido reconquistados e para que pudessem acompanhar a evolução do jogo, sabendo exatamente quantos objetos ainda precisavam ser capturados. A cronologia do teste pode ser vista na figura 15.

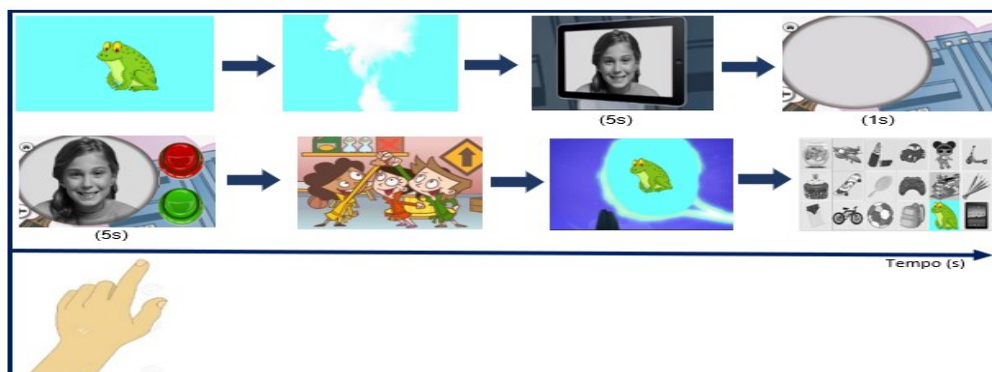


Figura 15: Cronologia do teste: o objeto é apresentado (o sapinho do apartamento 102). Depois, ele desaparece. A face do suspeito (prime) é mostrada primeiro no monitor da câmera de segurança e, em seguida, na lupa (alvo). O participante fazia a escolha através dos botões. Os detetives comemoram e um raio leva o objeto recuperado à planilha de controle.

4.2 Materiais

Para este estudo, foram usados 36 pares de imagens de rostos, distribuídos em duas versões, totalizando 72 imagens. Cada versão conta com 18 pares, dos quais 2 integram a fase de treinamento; 16 constituem as categorias experimentais (4 para as faces conhecidas; 4 para as faces famosas; 4 para as faces desconhecidas) e 4 compõem o grupo de distratores. As imagens apresentadas na câmera de segurança do prédio possuíam um corte retangular, já as exibidas na lupa dos detetives apresentavam um corte circular. Todas as imagens utilizadas no experimento foram neutralizadas com um filtro preto e branco e retiradas de sites de acesso livre de direitos autorais.

4.3 Participantes

Os mesmos participantes do experimento anterior realizaram o presente experimento e os dados tiveram o mesmo tipo de tratamento.

4.4 Resultados e Discussão

Os seres humanos possuem uma capacidade excepcional de reconhecer faces que já pode ser percebida nos primeiros meses de vida. Acredita-se que essa habilidade se desenvolva progressivamente e se especialize com a experiência. No entanto, também é sustentada a visão de que durante as primeiras fases da aprendizagem da leitura, a capacidade de identificação de um rosto e a decodificação de grafemas podem entrar circunstancialmente em rota de competição. Com o objetivo de conhecer esses efeitos, foi aplicado um teste de reconhecimento facial e seus resultados serão expostos e discutidos a seguir.

O presente experimento consistiu em um teste de discriminação de rostos de pessoas conhecidas (professoras e funcionários da escola), famosas (personalidades do

universo infantil) e pessoas desconhecidas, retidas em banco de fotografias de acesso livre. As imagens poderiam ser classificadas como igual, quando ambas as fotografias eram referentes à mesma pessoa; ou diferentes, quando se referiam a outras pessoas.

A figura 16 representa o total de respostas corretas para as três condições estabelecidas nas quatro turmas testadas: Pré-1, Pré-2, 1º e 2º anos. Neste experimento, não se fez necessária a redistribuição das turmas, pois a quantidade de dados obtidos se mostrou suficiente para a realização das análises por turma.

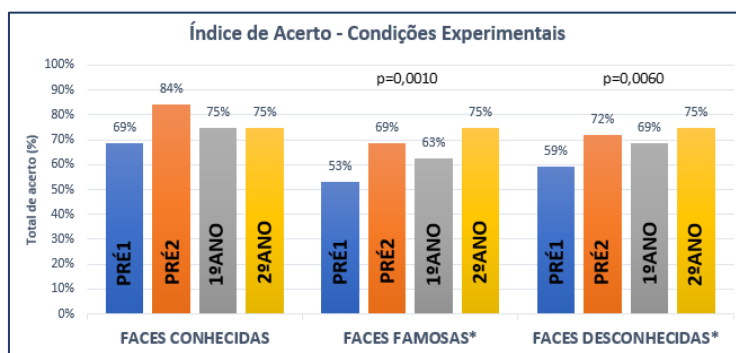


Figura 16: Índice de acerto para as condições experimentais.

A comparação entre as condições e as turmas resultou em significância estatística para as faces famosas ($p=0,0010$) e desconhecidas ($p=0,0060$). Acredita-se que a frequência com que os rostos conhecidos são vistos, sob os mais variados ângulos, considerando as formas mais sutis de alteração (o uso de maquiagem, a remoção, a inserção ou a troca de um acessório novo, como óculos etc.) modulou as respostas dos participantes de modo a torná-los especialistas no reconhecimento dessas imagens tão familiares. As faces famosas e desconhecidas, por sua vez, parecem demandar mais atenção no reconhecimento, mobilizando, possivelmente, estratégias distintas de identificação.

Além disso, esse dado também está se alinhando com as descobertas referentes ao reconhecimento de faces familiares. É bem aceito na literatura que as medidas internas da face são preferencialmente implicadas no reconhecimento de rostos conhecidos por indivíduos adultos (MCKONE *et al.*, 2012). Isso significa que o processamento holístico parece se sobrepôr ao processamento analítico nesse tipo específico de face. Os poucos estudos realizados constataram que este padrão também pode ser observado em crianças de 5 a 11 anos de idade (WILSON *et al.*, 2010). Nosso achado foi além, sugerindo que esse mecanismo pode ser verificado ainda mais cedo, em crianças com uma média 4 anos de idade. O aumento da amostra em testagens futuras pode ser capaz de ratificar, estatisticamente, a tendência observada.

O resultado expresso pela figura 16 demonstra ainda que a condição de faces conhecidas foi a que obteve o maior índice de acerto comparada às demais condições. O Pré 2 foi a turma que mais acertou, totalizando 84%, sendo seguido pelo 1º e 2º anos, que empataram em 75%. Por fim, está o Pré-1, com um percentual de 69% de respostas corretas.

A segunda condição mais acertada foi a das faces desconhecidas. Mais uma vez, os alunos do 2º ano apresentaram o maior índice de acerto em relação às outras turmas, atingindo um total de 75%. O Pré-2 ocupou a segunda colocação no ranking de melhor desempenho, apresentando 72% de precisão no reconhecimento. O 1º ano e Pré-1 apresentaram os índices mais baixos de acerto, 69% e 59%, respectivamente.

As condições referentes às faces de pessoas famosas foram as menos acertadas. Esse resultado joga luz sobre a importância da frequência de exposição e da interação no reconhecimento de faces. Embora a maioria dos participantes tenha reconhecido a imagem das celebridades utilizadas experimento, nomeando-as, inclusive (ex.: Neymar, Anitta, Lucas Netto etc.), esse conhecimento parece não ter sido suficiente para a realização de uma discriminação acurada desses rostos, resultando no menor índice de acerto observado. Também é importante considerar que tais rostos famosos são conhecidos das telas de celulares, TVs, etc – e não de experiência tridimensional. Assim, a interação e, conseqüentemente, a frequência de exposição parecem realmente ajudar a refinar o conhecimento sobre as faces pouco conhecidas, de modo que, sem a presença desses fatores, a discriminação desses rostos é menos precisa.

Como pode ser notado, o 2º ano apresentou, novamente, o melhor desempenho nessa categoria, totalizando 75% de acerto. O Pré 2 também demonstrou um bom nível de precisão no reconhecimento, ficando apenas 3% abaixo do 2º ano. As turmas de Pré 1 e 1º ano tiveram um desempenho inferior as demais turmas, alcançando 53% e 63% de acerto.

A fim de investigar mais profundamente as diferenças na percepção facial desses grupos, avaliamos também os resultados associados ao índice de respostas corretas nas condições em que as faces alvo sofreram três tipos distintos de manipulação: alterações externas, que incluíam o uso de um adereço, como óculos ou mudanças no cabelo; alterações internas no formato do olho, nariz e/ou boca; e a apresentação de uma nova face, pertencente a uma outra pessoa. Os resultados serão apresentados a seguir.

Na figura 17, está representado o índice de acertos dos participantes das quatro turmas nas condições em que foram adicionados adereços às faces. Pode-se observar uma tendência progressiva nas respostas dadas, representada pelo efeito *degradê* no tom das colunas.

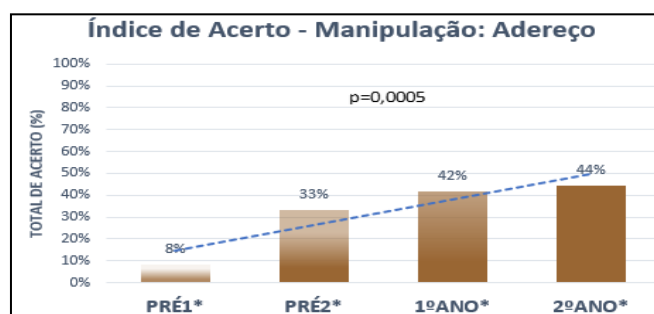


Figura 17: Índice de acerto para as condições manipuladas com o uso de adereço.

O grupo mais jovem, Pré-1, obteve o menor índice de acerto, apenas 8%. A precisão no reconhecimento foi aumentada conforme o passar dos anos, passando pelo Pré 2, com 33%, o 1º ano, com 42% de acerto e tendo sua maior ocorrência na turma de 2º ano, com 44%. Foi observada relevância estatística na comparação entre turmas dessa categoria ($p=0,0005$).

Os resultados referentes à taxa de acerto na apresentação da imagem de uma pessoa diferente da primeira apontam para um desempenho dentro das expectativas, sem grandes diferenças entre as faixas etárias estudadas. Assim, não foi constatado relevância estatística na comparação entre os grupos. Todas as turmas obtiveram uma tendência ao acerto superior a 85%.

Os resultados para índice de acertos em relação à última manipulação, *thatcher*, foram processados e plotados na figura 18. Verificamos significância estatística na

proporção de acerto dessa manipulação: $p=0,0003$ $df=3$. Diferentemente da gradação observada no uso de apetrechos e da regularidade vista no reconhecimento de uma outra face, a modificação *thatcher* foi sentida especialmente pelos participantes do 1º ano.

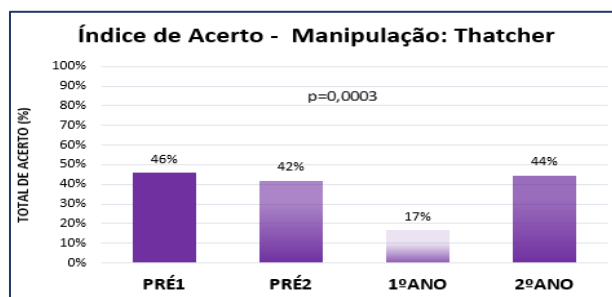


Figura 18: Índice de acerto para as condições de manipulação do tipo thatcher.

Como pode ser visualizado, as turmas de Pré-1 (46%), Pré-2 (42%) e 2º (44%) ano apresentaram resultados relativamente equivalentes, mantendo-se todos na faixa dos 40% de acerto. Porém, o nível de respostas corretas dos alunos de 1º ano, 17%, foi radicalmente mais baixo daqueles dos outros grupos.

Uma medida *off-line* do tempo gasto no reconhecimento dos três tipos de faces – conhecidas, famosas e desconhecidas – também foram computados. Os resultados podem ser observados nas figuras 19, 20 e 21. Na figura 19, estão apresentados os resultados referentes ao tempo médio gasto pelos participantes das diferentes turmas na discriminação de faces conhecidas.

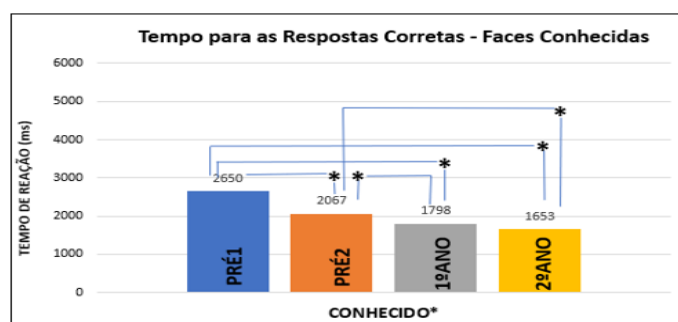


Figura 19: Tempo de reação para as respostas corretas nas condições de faces conhecidas.

Ainda que não tenha sido constatado efeito estatístico no quesito acuidade, na análise dos tempos de reação foi detectada relevância estatística desta fator $F(3,63)=18,6$ $p<0,000001$. O tempo empregado no reconhecimento foi reduzindo gradativamente de acordo com o nível de escolarização das turmas testadas. Assim, o grupo do Pré 1 foi aquele que demorou mais tempo na discriminação de faces conhecidas, cerca de 2650 milissegundos, em comparação com os outros grupos. O 2º ano, por sua vez, levou menos tempo nessa mesma tarefa, 1653ms. O Pré 2 e o 1º ano ocuparam uma posição intermediária, atingindo uma marca de 2067ms e 1798ms, nessa ordem. Essa foi uma constatação par a par que resultou em significância estatística (Pré 1 vs Pré 2 $t(21)=8,00$ $p<0,0001$; Pré 1 vs 1º ano $t(21)=7,65$ $p<0,0001$; Pré 1 vs 2º ano $t(21)=4,55$ $p<0,0002$). De maneira geral, o resultado sugere que a rapidez no reconhecimento de faces conhecidas pode estar associada à experiência: à medida que um indivíduo frequenta diferentes espaços, aumentando o círculo de pessoas com as quais convive, ele tem mais possibilidades de ser exposto a diversas características faciais, automatizando, assim, a habilidade de reconhecimento com o passar dos anos.

No que diz respeito a faces famosas, encontramos também ainda efeito principal desse fator: $F(3,54) = 11,2$ $p < 0,000008$. A média dos tempos está disposta na figura 20. Assim como observado nos tempos das condições anteriores, as respostas às faces famosas também demonstraram uma gradação. Os tempos médios registrados expressam que o Pré 1 levou mais tempo na discriminação dos rostos de celebridades, levando 2616ms no reconhecimento, enquanto o 2º ano foi o grupo mais rápido, gastando 1674ms nesta categoria. O Pré 2 (2005ms) e o 1º ano (1947ms) ficaram, novamente, numa posição média. A comparação par a par evidenciou relevância estatística desse fator entre as turmas de Pré 1 vs Pré 2 $t(18)=2,94$ $p < 0,0087$; Pré 1 vs 1ºano $t(18)=3,62$ $p < 0,0019$; Pré 1 vs 2ºano $t(18)=4,63$ $p < 0,0002$; Pré 2 vs 2ºano $t(18)=2,06$ $p < 0,0546$.

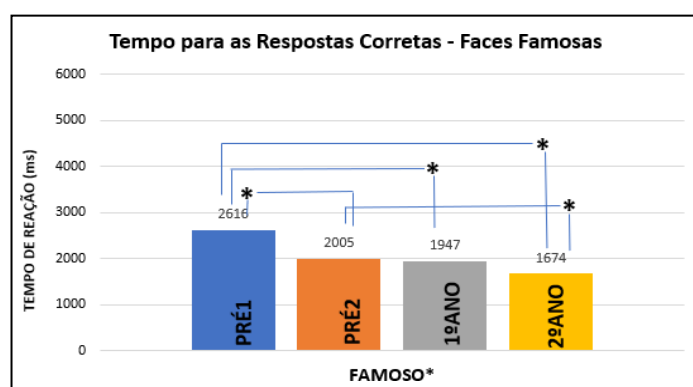


Figura 20: Tempo de reação para as respostas corretas nas condições de faces famosas.

De maneira diferente ao que vem sendo reportado, os tempos relativos ao reconhecimento de faces desconhecidas mostraram uma relativa estabilidade nas turmas de Pré 1, Pré 2 e 2º ano. No entanto, o grupo do 1º ano se destaca apresentando o maior tempo registrado, apresentado na figura 21. A análise dos tempos de resposta referentes a esse fator resultou em relevância estatística $F(3,57) = 5,41$ $p < 0,002395$.



Figura 21: Tempo de reação para as respostas corretas nas condições faces desconhecidas.

Como pode ser notado, os alunos do 1º ano foram os que gastaram maior tempo no reconhecimento de rostos desconhecidos, cerca de 2303ms. Em seguida, vêm os participantes da Pré-Escola 2, totalizando 1911ms. Os grupos de Pré 1 (1843ms) e do 2º ano (1882ms) foram os mais rápidos na discriminação dessas faces. Foram encontradas diferenças significativas somente na comparação do 1º ano com as demais turmas (Pré 1 vs 1ºano $t(19)=4,28$ $p < 0,0004$; Pré 2 vs 1ºano $t(19)=4,68$ $p < 0,0002$; 1ºano vs 2ºano $t(19)=2,66$ $p < 0,0156$). De maneira geral, os dados de tempo são complementares ao que foi observado na análise das respostas certas. As condições referentes às faces conhecidas foram discriminadas com mais rapidez e maior precisão por todas as turmas. Isso indica que processos distintos estão envolvidos no reconhecido, como propõem Bruce e Young

(1986) em seu modelo de reconhecimento de faces. O reconhecimento de faces desconhecidas, por sua vez, foi mais significativamente custosa em termos temporais para os alunos do 1º ano do que para qualquer outro grupo testado. Esse é um dado interessante não previsto na elaboração do experimento.

Na próxima figura (22), podemos observar o tempo empregado pelos participantes para responder de maneira correta às mesmas manipulações já verificadas. Foi encontrado efeito principal da manipulação do tipo adereço $F(3,6) = 4,41$ $p < 0,058014$. A comparação par a par não resultou em relevância estatística, indicando que as turmas gastaram um tempo aproximado no julgamento. É possível observar, contudo, que o maior tempo gasto na discriminação das faces foi do Pré 1 (4952ms), seguido pelo Pré 2 (3274ms), o 2º ano (3109ms) e, por fim, o 1º ano (3109ms).

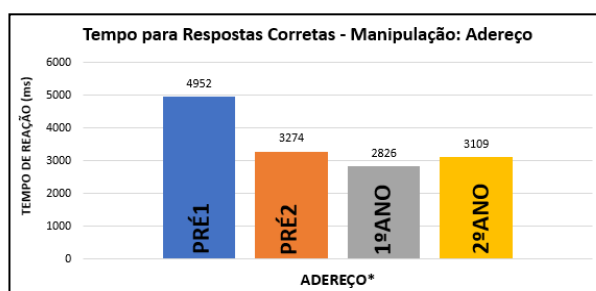


Figura 22: Tempo de reação das turmas testadas às manipulações do tipo adereço.

Os resultados de tempo de resposta e índice de acurácia mostraram que, mesmo levando mais tempo na tarefa, a turma de Pré 1 teve o pior desempenho no reconhecimento de rostos com algum tipo de adereço. Os tempos mais baixos e a maior precisão na discriminação foram atribuídos às turmas mais velhas, de 2º e 1º anos. Essa ocorrência pode estar também associada à experiência. Ao passo que a cognição de processamento visual vai se desenvolvendo, as características mais superficiais, aquelas que são mais facilmente mutáveis, como alterações no cabelo, uso de um boné ou óculos de grau, deixam gradativamente de desempenhar um papel importante no reconhecimento, que vai se tornando especializado na identificação das características fundamentais.

É interessante perceber que apesar de não ter havido relevância estatística entre as faixas etárias estudadas para número de acertos na percepção de face diferente, houve efeito principal da manipulação do tipo outra face: $F(3,48) = 23,6$ $p < 0,000001$ (cf. figura 23). Além disso, muitos resultados par a par foram também estatisticamente relevantes: (Pré 1 vs Pré 2 $t(16)=4,15$ $p < 0,0007$; Pré 1 vs 2ºano $t(16)=5,67$ $p < 0,0001$; Pré 2 vs 1ºano $t(16)=7,89$ $p < 0,0001$; Pré 2 vs 2ºano $t(16)=3,14$ $p < 0,0063$; 1ºano vs 2ºano $t(16)=6,42$ $p < 0,0001$).

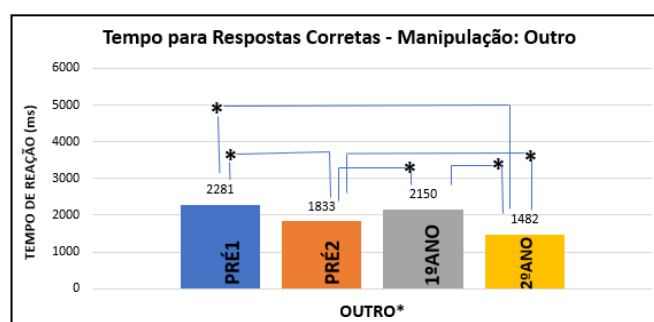


Figura 23: Tempo de reação das turmas testadas às manipulações do tipo outra face.

Entre esses resultados, de novo, obtivemos um índice de interesse focal para essa investigação no que diz respeito ao tempo de decisão dos participantes do 1º ano. Note-se que os tempos resposta não tiveram uma diminuição gradual indo do Pré 1 ao 2º ano. Os participantes do 1º ano apresentaram tempo de decisão bastante elevado em relação ao grupo, equivalente ao tempo dos participantes do Pré 1 e estatisticamente maior do que o tempo dos participantes do Pré 2 e do 2º ano. Esse achado, sugerindo uma maior dificuldade para os alunos da classe de alfabetização (1º ano) em lidar com sinais de reconhecimento facial e discriminação entre faces diferentes, está de acordo com as previsões teóricas de Dehaene (2005), com estudos de imagem (*cf.* LI, 2013; CANTLON, 2011) e também se modula de forma semelhante aos achados do Experimento 1, que detectou maior dificuldade dessa mesma população (tempos de resposta estatisticamente mais longos do que os do Pré 2 e 2º ano) em relação às condições [-TON] e [-TRANS].

Os tempos relativos à discriminação das faces thatcherizadas (figura 24) não se converteram em efeito principal. Todavia foi captada relevância estatística na comparação das turmas de Pré 1 e de 2º ano $t(7)=3,07$ $p < 0,0182$. Ainda assim, podemos observar que as turmas não alfabetizadas (Pré 1 – 2567ms e Pré 2 – 2401ms) foram mais rápidas no reconhecimento – e mais corretas – do que os alunos alfabetizados (1º ano – 2792ms e 2º ano – 3035ms).

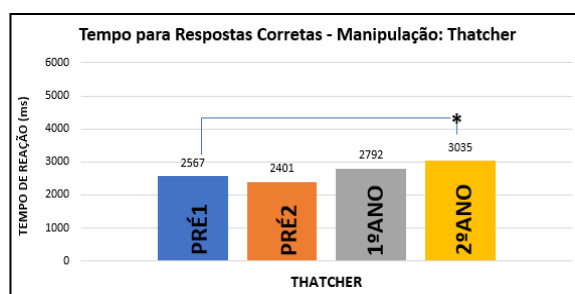


Figura 24: Tempo para as respostas corretas nas condições de manipulação *thatcher*.

Os resultados de tempo de resposta e acuidade das turmas mais escolarizadas (1º e 2º anos) relativos à manipulação do tipo *thatcher*, nos leva a supor que o desenvolvimento da alfabetização esteja impactando negativamente o reconhecimento de faces nessa categoria. Nas condições thatcherizadas, o formato dos olhos, nariz ou boca das imagens alvo foram alterados através do uso de um editor de imagens (Photoshop 2017), forçando o participante a recrutar estratégias mais específicas e acuradas na percepção das faces. Considerando que os indivíduos não alfabetizados são sistematicamente mais holísticos que os alfabetizados, na visualização de faces (VENTURA *et al.*, 2013), e que o processamento holístico tende a ser mais inflexível, porém, mais confiável (EYSENCK; KEANE, 2017), é possível acreditar que a aprendizagem da leitura tenha afetado momentaneamente a percepção das crianças em processo de alfabetização a essas condições, provocando um baixo nível de acurácia e mobilizando altos tempos de resposta.

Finalmente, com a intenção de avaliar a percepção geral dos nossos participantes no julgamento de um rosto tido como pertencente à mesma pessoa ou a pessoas distintas, analisamos a quantidade de respostas verdes (mesma pessoa) e vermelhas (pessoa diferente) dadas por todos os integrantes das quatro turmas para cada um dos itens experimentais, independente da acurácia na discriminação. Os resultados podem ser observados na figura abaixo (figura 25).

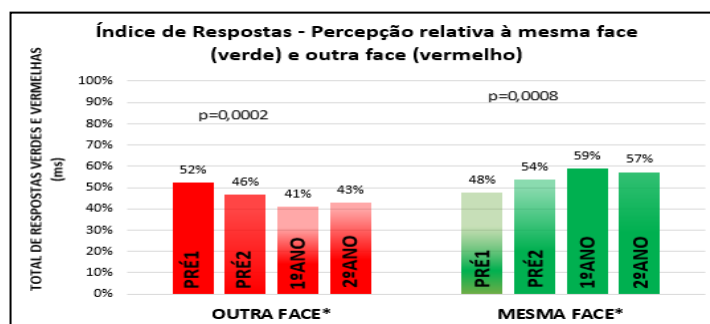


Figura 25: Total de respostas verdes, percepção relativa à mesma face e vermelhas, outra face.

Foi verificada relevância estatística de ambas as possibilidades de resposta: vermelho ($p=0,0002$) e verde ($p=0,0008$). Note-se que os resultados revelam uma diminuição gradativa nas respostas diferentes, sendo os participantes do Pré 1 os mais restritivos, avaliando como diferentes a maioria dos pares exibidos. Esse julgamento vai sendo flexibilizado conforme aumentam os níveis de escolarização dos participantes. O inverso pode ser observado na análise das respostas verdes. As turmas mais escolarizadas apresentam os maiores níveis de resposta igual à discriminação dos estímulos apresentados. Esse comportamento vai sendo paulatinamente mais restritivo à medida que observamos os dados provenientes dos participantes mais novos.

Esse dado pode evidenciar a utilização de mecanismos mais holísticos e restritos pelos participantes não alfabetizados em todas as condições experimentais propostas, julgando como diferentes os alvos cujas manipulações vão desde alterações mais sutis, até as mais perceptíveis. Por outro lado, os sujeitos alfabetizados parecem aplicar um mecanismo mais analítico no processamento de faces, decidindo que duas imagens são pertencentes à mesma pessoa com diferentes níveis de modificação.

5 Considerações Finais

Este trabalho se dedicou a compreender melhor os efeitos do desenvolvimento da leitura sob as cognições de processamento de faces e dos sons da fala. Para este fim, foi aplicada uma série de dois testes psicolinguísticos em turmas de pré-escola, alfabetização e ensino fundamental. A primeira tarefa consistiu no reconhecimento de sons silábicos em palavras e pseudopalavras. A segunda se resumiu na discriminação de faces de pessoas conhecidas, famosas e desconhecidas.

Nossos achados em relação ao Experimento I (reconhecimento auditivo de sílaba na palavra) foram compatíveis com a hipótese de que a informação ortográfica atua como um facilitador na missão de julgar se as sílabas faladas integram ou não as palavras ouvidas. A predição de que quanto maior for o nível de conhecimento ortográfico do sujeito, melhor seria o seu desempenho em termos de tempo de resposta e acuidade também foi confirmada.

Os resultados obtidos revelaram que os participantes recém alfabetizados (2º ano) foram consistentemente mais acurados do que crianças não alfabetizadas (Pré-1 e Pré-2) na tarefa auditiva de reconhecimento de sons silábicos, tendo, o 2º ano, mobilizado tempos de reação mais baixos. Mesmo nas condições em que os fatores tonicidade e a transparência foram anulados, foi verificado que o grupo que possui conhecimento sobre leitura foi substancialmente melhor no reconhecimento dos sons apresentados do que os sujeitos não-alfabetizados.

No entanto, a turma de 1º ano apresentou um resultado mais instável. Por vezes, esse grupo exibiu um bom desempenho, se assemelhando ao 2º ano; em outros contextos,

seu desempenho era regular, próximo ao Pré 2 e, ainda, inferior ao grupo de crianças não alfabetizadas de Pré-escola 2. Os tempos de resposta de resposta do 1º ano também evidenciam a aparente dificuldade desse grupo. Os maiores tempos de resposta registrados foram dos participantes do 1º ano. Em contrapartida, os menores tempos gastos para respostas corretas foram dos alunos do 2º ano.

No que concerne à tarefa de reconhecimento facial, a hipótese aventada de que o desenvolvimento da alfabetização afetaria a percepção mais holística das faces foi atestada. Da mesma forma, nossos resultados validaram a previsão de que as respostas dos participantes alfabetizando e alfabetizados a essas categorias deveriam resultar em um menor nível de acurácia e seu desempenho mobilizariam tempos de resposta mais altos em relação aos sujeitos não alfabetizados.

Nossa principal descoberta foi a de que alterações do tipo *thatcher*, que modificam a forma de elementos pontuais da face, como olhos, nariz e boca, são capazes de prejudicar substancialmente a precisão de crianças em processo de alfabetização na tarefa de reconhecimento desses rostos. Essa dificuldade, ainda que momentânea, não é completamente superada pelo grupo alfabetizado do 2º ano.

Os tempos de resposta também evidenciam o esforço da turma de 1º ano no reconhecimento de faces *thatcherizadas*. Os maiores tempos de resposta foram atribuídos às turmas de 1º e 2º anos. Sugerimos que estes achados estejam relacionados à forte dominância do hemisfério direito para faces em sujeitos alfabetizados, cujo processamento é mais analítico e flexível e, conseqüentemente, menos confiável do que o hemisfério esquerdo em indivíduos não alfabetizados, cuja codificação é mais holística e rigorosa e, porém, mais segura (VENTURA *et al.*, 2013).

Assim, nossas descobertas revelam que o desenvolvimento da leitura parece impactar positivamente o processamento da fala, tornando sua percepção mais acurada e especializada; mas, por outro lado, parece afetar, provisoriamente, a cognição de reconhecimento de faces, especialmente nos casos em que há demanda de mecanismos holísticos.

Por meio de experimentos psicolinguísticos, estruturados ecologicamente a partir de uma metodologia simples e acessível, este trabalho pôde reforçar achados de neurociência de ponta e lançar luz sobre processos cognitivos sincréticos, antes pouco conhecidos, que impactam profundamente os mecanismos de aprendizagem, principalmente no se refere à aquisição da leitura e da escrita.

Referências Bibliográficas

- BRUCE, V.; YOUNG, A. **Face perception**. Hove, East Sussex: Psychology Press, 2012.
- CANTLON, J. et al. Cortical representations of symbols, objects, and faces are pruned back during early childhood. **Cerebral cortex**, v. 21, n. 1, p. 191-199, 2011.
- CASTLES, A. How does orthographic knowledge influence performance on phonological awareness tasks? **The Quarterly Journal of Experimental Psychology**. Section A, v. 56, n. 3, p. 445-467, 2003.
- CHANG, E. F. et al. Categorical speech representation in human superior temporal gyrus. **Nature neuroscience**, v. 13, n. 11, p. 1428, 2010.
- COHEN, L. et al. Language-specific tuning of visual cortex? Functional properties of the visual word form area, **Brain**, ed. 125, p. 1054-1069, 2002.
- COHEN, L.; DEHAENE, S. Specialization within the ventral stream: the case for the visual word form area, **Neuroimage**, n. 22, p. 466-476, 2004.
- COLLISHAW, S. M.; HOLE, G. J. Featural and configurational processes in the recognition of faces of different familiarity. **Perception**, v. 29, n. 8, p. 893-909, 2000.

- DEHAENE, S. et al. How learning to read changes the cortical networks for vision and language. **Science**, v. 330, n. 6009, p. 1359-1364, 2010.
- DEHAENE, S. et al. Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. **Nature Reviews Neuroscience**, v. 16, n. 4, p. 234-244, 2015.
- DEHAENE, S. et al. The neural code for written words: a proposal. **Trends in cognitive sciences**, v. 9, n. 7, p. 335-341, 2005.
- DEHAENE, S. **Os neurônios da leitura**. Porto Alegre: Penso, 2012.
- DEHAENE-LAMBERTZ, G.; MONZALVO, K.; DEHAENE, S. The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. **PLoS biology**, v. 16, n. 3, p. e2004103, 2018.
- DIAMOND, R.; CAREY, S. Developmental changes in the representation of faces. **Journal of Experimental Child Psychology**, v. 23, n. 1, p. 1-22, 1977.
- DIAMOND, R.; CAREY, S. Why faces are and are not special: an effect of expertise. **The Journal of Experimental Psychology: General**, v. 115, n. 2, p. 107, 1986.
- EHRI, L. C.; WILCE, L. S. The mnemonic value of orthography among beginning readers. **Journal of Educational Psychology**, v. 71, n. 1, p. 26, 1979.
- EYSENCK, M. W.; KEANE, M. T. **Manual de Psicologia Cognitiva**. Artmed Editora, 2017.
- FREIRE, A; LEE, K.; SYMONS, L. A. The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. **Perception**, v. 29, n. 2, p. 159-170, 2000.
- KOLINSKY, R.; PATTAMADILOK, C.; MORAIS, J. The impact of orthographic knowledge on speech processing. **Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies**, n. 63, p. 161-186, 2012.
- KOLINSKY, R.; VERHAEGHE, A. How literacy affects vision. Further data on the processing of mirror images by illiterate adults, **Revista linguística**, v. 7, n. 2, dezembro 2011.
- LEAL, K.B.M. **Um estudo do sândi externo: o contato entre o português do Brasil e o mbyá guarani**. Dissertação (Mestrado em Linguística). Programa de Pós-Graduação em Linguística (PPGL), UFRJ, Rio de Janeiro, P.104. 2015.
- LEDER, H.; BRUCE, V. When inverted faces are recognized: The role of configural information in face recognition. **The quarterly journal of experimental psychology. Section A**, v. 53, n. 2, p. 513-536, 2000.
- LI, S. et al. Neural competition as a developmental process: early hemispheric specialization for word processing delays specialization for face processing. **Neuropsychologia**, v. 51, n. 5, p. 950-959, 2013.
- MCCANDLISS, B. D; COHEN, L; DEHAENE, S. The visual word form area: expertise for reading in the fusiform gyrus, **Trends Cogn Sci**, ed. 7, p. 293-299, 2003.
- MCKONE, E. et al. A critical review of the development of face recognition: Experience is less important than previously believed. **Cognitive neuropsychology**, v. 29, n. 1-2, p. 174-212, 2012.
- MCKONE, E.; KANWISHER, N.; DUCHAINE, B. C. Can generic expertise explain special processing for faces? **Trends in cognitive sciences**, v. 11, n. 1, p. 8-15, 2007.
- MESGARANI, N. et al. Phonetic feature encoding in human superior temporal gyrus. **Science**, v. 343, n. 6174, p. 1006-1010, 2014.
- MORAIS, J. **Criar leitores - Para professores e educadores**. Barueri, SP: Manole, 154 p. 2013.

SEIDENBERG, M. S.; TANENHAUS, M. K. Orthographic effects on rhyme monitoring. **Journal of Experimental Psychology: Human Learning and Memory**, v. 5, n. 6, p. 546, 1979.

TANAKA, James W.; FARAH, Martha J. Parts and wholes in face recognition. **The Quarterly Journal of Experimental Psychology**, v. 46, n. 2, p. 225-245, 1993.

VAN OOIJEN, B. et al. Do weak syllables count for newborns? **The Journal of the Acoustical Society of America**, v. 102, n. 6, p. 3735-3741, 1997.

VENTURA, P. et al. Literacy acquisition reduces the influence of automatic holistic processing of faces and houses. **Neuroscience letters**, v. 554, p. 105-109, 2013.

WILSON, R. R.; BLADES, M.; PASCALIS, O. What do children look at in an adult face with which they are personally familiar? **British Journal of Developmental Psychology**, v. 25, n. 3, p. 375-382, 2007.

ZHOU, Z.; WHITNEY, C.; STROTHER, L. Embedded word priming elicits enhanced fMRI responses in the visual word form area. **Plos one** 14(1) p. 1 2019.

Submetido em 22/10/2020

Aceito em 22/01/2021