

# A Fala Sintetizada de Expressões Matemáticas: Um Estudo para melhoria do Processo Cognitivo do Estudante com Deficiência Visual

Adriana Souza<sup>1,2</sup>, Diamantino Freitas<sup>3</sup>

<sup>1</sup>Instituto Federal da Bahia (IFBA) Campus Porto Seguro  
Campus Porto Seguro – Porto Seguro – Ba – Brasil

<sup>2</sup>Programa Doutoral em Media Digitais (PDMD)  
Faculdade de Engenharia – Universidade de Porto – Porto, Portugal

<sup>3</sup>Departamento de Engenharia Eletrotécnica e de Computadores  
Faculdade de Engenharia – Universidade de Porto – Porto, Portugal

adriana@ifba.edu.br, dfreitas@fe.up.pt

**Resumo.** *A fala sintetizada de conteúdos matemáticos ainda apresenta desafios para estudantes que fazem uso dos leitores de tela, entre os quais podemos citar, as pausas inadequadas e as longas saídas auditivas, o que dificulta a memorização desse tipo de conteúdo. Nesse estudo realizamos dois experimentos com o objetivo de identificar e analisar processos capazes de reduzir a sobrecarga cognitiva da fala sintetizada de expressões matemáticas codificadas em MathML. O primeiro experimento visou verificar as dificuldades encontradas pelos estudantes com deficiência visual, além de analisar um modelo de pausas proposto. O segundo experimento buscou compreender os processos cognitivos para memorização de expressões matemáticas, através da técnica de rastreamento ocular de pessoas videntes. Embora alguns resultados não tenham sido conclusivos, a pesquisa mostrou-se relevante, pois aponta direções que podem minimizar a carga mental e, conseqüentemente melhorar o processo cognitivo do estudante na leitura de expressões matemáticas.*

## 1. Introdução

A Matemática possui um estilo de leitura própria que difere da leitura textual. Embora a fala sintetizada tenha evoluído nos últimos anos, torná-la mais natural e cognitivamente eficiente para conteúdos matemáticos ainda é um desafio (Bansal e Balakrishnan 2021; Bates e Fitzpatrick 2010). Entre as limitações apresentadas para o processo de ensino-aprendizagem da Matemática das pessoas com deficiência visual podemos citar as longas saídas auditivas, dificultando a memorização desse tipo de conteúdo e a falta de prosódia correta na leitura, o que gera ambigüidades no significado das expressões matemáticas (Frankel et al. 2016).

Várias abordagens (Bansal e Balakrishnan 2021; Bates e Fitzpatrick 2010; Stevens 1996) têm sido usadas para melhorar a fala sintetizada da Matemática, como dicas lexicais, dicas prosódicas, *earcons* e *spearcons*, controle do fluxo de informações por navegação nas expressões e som espacializado. Embora o uso combinado dessas abordagens tenha sido positivo, a falta de uma boa prosódia e a sobrecarga cognitiva

persistem (Bansal e Balakrishnan 2021; Bates e Fitzpatrick 2010; Souza e Freitas 2018). De acordo com Reich (1980), as pausas durante a fala proporcionam um tempo de processamento vital durante a compreensão pelo ouvinte, ajudando a processar e armazenar as informações divididas em subunidades significativas da linguagem. Teóricos linguísticos (Chomsky 1964, 2009; Kintsch 1978; Rumelhart e Norman 1981) também relatam que estruturas de árvore hierárquica fornecem um modelo para representações cognitivas. Expressões matemáticas representadas em árvores de análise sintática foram apresentadas por Ernest (1987), onde cada nó representava um símbolo, e árvores de análise sintática representavam subexpressões. Para os teóricos, esse tipo de representação pode ajudar a memorizar esse tipo de conteúdo.

Vários estudos têm mostrado que pessoas com deficiência visual e videntes utilizam os mesmos procedimentos básicos para selecionar e reter o conhecimento na memorização (Röder, Rösler, e Neville 2001). Portanto, entender como as pessoas analisam, processam e dividem expressões matemáticas longas em expressões menores, pode ajudar a construir um modelo semelhante para a fala sintetizada da Matemática que reduza a sobrecarga cognitiva. Por outro lado, os estudos sobre pausas na fala da Matemática (Frankel et al. 2016; Frankel e Brownstein 2016; Maier et al. 2014; O'Malley, Kloker, e Dara-Abrams 1973; Stevens 1996) têm utilizado valores estáticos e uniformes ou fazendo uso de pontuação como vírgula, ponto e dois pontos, para forçar o sintetizador a pausar no lugar desejado. Contudo, a fala é um processo dinâmico e as pausas nas expressões também precisam ser dinâmicas para tornar a fala sintética das expressões mais natural, pois isso pode ajudar na melhoria da prosódia e consequentemente na memorização e retenção desse tipo de conteúdo (Reich 1980). Diante da problemática citada, este artigo apresenta dois estudos que visam melhorar o processo cognitivo do estudante durante a sua recepção da leitura sintetizada de expressões matemáticas: o estudo de rastreamento ocular e um modelo de pausas dinâmico para a fala sintetizada de expressões matemáticas.

## 2. Estudos de Rastreamento Ocular na Educação Matemática

O rastreamento ocular é um processo usado para registrar os movimentos oculares de indivíduos em tempo real. Equipamentos específicos permitem registrar evidências significativas sobre como o olhar das pessoas interage com as informações visuais, incluindo padrões de leitura e resolução de problemas (Rayner 1998). As pesquisas científicas baseadas na metodologia de rastreamento ocular aumentaram nos últimos anos. O estudo proposto por André et al. (2015) utilizou a metodologia de rastreamento ocular para analisar como os alunos observam diferentes representações matemáticas, como fórmulas e gráficos. A tecnologia de rastreamento ocular também foi usada por Hartmann (2015) para analisar o desempenho dos alunos em Matemática por meio da resolução de problemas. Eles concluíram que o desempenho na resolução de problemas pode ser um preditor essencial do desempenho em Matemática. A metodologia de rastreamento ocular também foi usado por Schindler et. al (2022) para analisar o processo criativo colaborativo de alunos em Matemática. Todavia, não encontramos em nossa revisão de literatura estudos que utilizem o rastreamento ocular para analisar processos cognitivos para melhorar a fala sintetizada da Matemática. Considerando que tal investigação pode revelar como as pessoas decodificam as informações visuais, observar os movimentos oculares pode oferecer *insights* (pistas) sobre o processamento cognitivo da leitura visual de expressões matemáticas, o que pode ajudar na construção

de um modelo de síntese de fala para leitura de expressões matemáticas que cause uma carga cognitiva aceitável para pessoas com deficiência visual.

### **3. As Pausas na Fala Sintetizada da Matemática**

Stevens (1996) definiu um conjunto de regras para pausas em expressões matemáticas. Foi definida uma pausa de aproximadamente 300 ms para marcar o início/fim de uma subexpressão e, antes dos operadores binários explícitos, foi definida também uma pausa de 250 ms no operador ‘/’ para frações complexas, entre outras regras. No trabalho de Frankel e Brownstein (2016), foi proposto que as pausas poderiam substituir algumas pistas lexicais para reduzir a sobrecarga cognitiva do aluno. Portanto, definiram pausas uniformes de diferentes tamanhos e mudança na velocidade de fala para definir os limites sintáticos das expressões matemáticas, por exemplo, para delimitar o final da raiz quadrada nas expressões, foi definida uma pausa longa de 2750ms. Não encontramos na literatura trabalhos que definam pausas para a fala sintetizada de expressões matemáticas com variabilidade automática, de acordo com o tamanho e tipo da expressão. Todavia, os trabalhos encontrados (Ferreira e Freitas 2004; Frankel e Brownstein 2016; Stevens 1996) definem pausas de duração uniforme ou por meio de sinais de pontuação. Sendo o Audiomath (Ferreira e Freitas 2004) uma aplicação web que converte expressões matemáticas em MathML (Mathematical Markup Language) (Carlisle, Ion, e Miner 2014) para português europeu. Ao criar a versão em linguagem natural da equação, o Audiomath insere as pausas por meio de sinais de pontuação, como vírgula, ponto e dois pontos.

Neste estudo, optamos por comparar o atual modelo para gerar pausas automáticas com o Audiomath, por ser a única ferramenta disponível que fala corretamente expressões matemáticas em português. Este trabalho difere dos trabalhos relacionados por apresentar um modelo que calcula dinamicamente a duração das pausas, aproximando-se da naturalidade da fala, que emprega variabilidade na duração das pausas durante a fala de expressões matemáticas. Segundo (Campioni e Véronis 2002), em estudos de pausas, os cálculos estatísticos devem ser realizados no domínio logarítmico, pois a lei log-normal proporciona um ajuste muito melhor aos dados do que a lei normal. Dada a enorme assimetria da distribuição das pausas, as medições no domínio aritmético não são confiáveis.

### **4. Modelo Dinâmico de Pausas para Expressões Matemáticas**

#### **4.1 Material**

O modelo de pausas foi construído a partir de doze expressões matemáticas selecionadas aleatoriamente do livro de Matemática utilizado no ensino médio, intitulado “Aula de Matemática por turma, volume único”. As expressões foram verbalizadas por cinco professores de Matemática e gravadas em formato de áudio. Todos os professores possuem formação na área e lecionam a disciplina há mais de cinco anos. O estudo utilizou expressões bidimensionais com a estrutura fração. Esse tipo de estrutura exige adição de pistas prosódicas, como pausas e também dicas lexicais que ajudam a desambiguar as expressões, além disso, têm o potencial de causar sobrecarga cognitiva no estudante.

## 4.2 Método

Os áudios com as expressões matemáticas foram etiquetadas no *Praat* (Boersma e Weenink, 2018). A estrutura fração costuma ter um padrão de quatro pausas. A primeira (p1) após a dica lexical “início de fração”, a segunda (p2) após “numerador”, a terceira (p3) antes de falar “denominador” e a última (p4) antes da dica lexical “fim de fração”. Percebemos que a fala natural possui uma variabilidade pessoal. Por isso, assumimos que as pausas são compostas por componentes determinísticos e aleatórios. Os componentes determinísticos são fatores controláveis pela duração média dos caracteres. Os componentes aleatórios são causados pela variabilidade individual, que não podemos controlar. Realizamos uma regressão linear com pausas e duração média dos caracteres para identificar a variabilidade individual. A Tabela 1 apresenta as estatísticas de regressão da estrutura analisada para a pausa relativa 1 (PR1). O  $R^2$  obtido explica 57% da variabilidade das pausas do modelo, representando uma relação moderada de acordo com os parâmetros definidos por (Rumsey 2016).

**Tabela 1. Estatística de Regressão da PR1 e DMC**

Estatísticas de Regressão				
$R^2$	0,578	Observações	26	
ANOVA				
	gl	F	Significância de F	Total
Regressão	1	32,935	0,000	25
Resíduos	24			
		Coefficientes	Desvio padrão	t
Intercept		2,673	0,207	12,887
Duração média do carácter		-0,593	0,103	-5,739
				p-valor
				0,000

No teste ANOVA (Tabela 1), o F, de significância ( $p < 0,05$ ), representa que a variável independente (duração média do carácter) está relacionada à pausa. O teste t também é significativo ( $p < 0,05$ ). Assim, concluímos que a duração média do carácter é um preditor significativo da duração das pausas.

Para obter a variabilidade das pausas, analisamos os resíduos da regressão (a diferença entre o valor real e o valor previsto da pausa). Decompomos os resíduos totais (RT) em resíduos determinísticos (RD) e resíduos pessoais ou aleatórios (RP). Os RD são a raiz quadrada da diferença da variância dos RT e RP. Para obter os RP, realizamos regressão linear com os valores das pausas de uma expressão matemática pronunciada várias vezes pela mesma pessoa. Em seguida, calculamos o valor dos RD. Heuristicamente, definimos um coeficiente C que representa a porcentagem de resíduos determinísticos dos resíduos totais.  $C = RD/RT$  (coeficiente igual a resíduos determinísticos divididos pelos resíduos totais). Multiplicamos os RT pelo coeficiente para eliminar os RP, ou seja, os resíduos da componente aleatória das pausas. Após calcular os resíduos determinísticos (e), criamos a equação de regressão linear (1) com o intercepto (2,67) e o coeficiente (-0,59), conforme mostrado na Tabela 1. Recalculamos a pausa relativa 1 sem a influência dos componentes aleatórios, e obtivemos um  $R^2$  maior, que representa uma relação forte (0,77). Repetimos os procedimentos descritos para calcular as demais pausas da estrutura fração, conforme as Equações (2), (3) e (4).

$$Y = -0,59X + 2,67 + e \quad (1), \quad Y = -0,66X + 2,70 + e \quad (2)$$

$$Y = -0,71X + 2,79 + e \quad (3), \quad Y = -0,58X + 2,56 + e \quad (4)$$

Para avaliar o modelo de pausas proposto, usamos expressões matemáticas que foram utilizadas em outros estudos relacionados, devido às expressões terem sido selecionadas com participação de especialistas e professores que ensinam Matemática

para estudantes com deficiência visual (Frankel e Brownstein 2016; Stevens 1996). No experimento, usamos quatro expressões matemáticas sintetizadas pela ferramenta Audiomath (Ferreira e Freitas 2004) e criamos quatro expressões clones para minimizar os efeitos da memorização sintetizadas com nosso modelo de pausas. A amostra foi composta por dez participantes com deficiência visual, oito portugueses e dois brasileiros, dos níveis fundamental, técnico, graduação, mestrado e doutorado. Criamos um questionário usando o *Google Forms* com treze seções. A primeira seção foi destinada às questões demográficas e antes de iniciar a segunda seção, reproduziam-se dois áudios de expressões matemáticas sintetizadas para que o participante se familiarizasse com a voz antes de iniciar o experimento. Entre as seções 2 a 9, havia uma expressão matemática sintetizada com Audiomath ou com o novo modelo de pausa, inseridas no formulário de forma aleatória para evitar viés. O participante precisava ouvir a expressão matemática e responder a algumas questões que tinham por objetivo verificar se o participante entendeu a expressão e se as pausas ajudaram. Nas seções 9 a 12, havia dois áudios com a mesma expressão matemática, mas um sintetizado com Audiomath e outro com o modelo de pausa, inseridos também de forma aleatória no formulário. O participante deveria selecionar a opção na escala que melhor representasse a sua compreensão ao comparar os dois áudios.

## 5. Experimento de Rastreamento Ocular

### 5.1 Amostra e Material

A amostra foi composta por doze participantes, com idades entre 21 e 67 anos, recrutados para participação voluntária entre alunos e professores da Faculdade de Engenharia da Universidade do Porto em setembro de 2021.

Foram utilizadas sete expressões matemáticas do ensino fundamental e médio que possuíam operações como frações, raízes quadradas, potências, logaritmos e funções trigonométricas. Optamos por utilizar expressões matemáticas com os símbolos usuais das operações algébricas, com complexidade em escala crescente e natureza bidimensional. As expressões foram escritas com a fonte Cambria Math em tamanho 60 com letras pretas sobre um fundo branco e apresentadas em um monitor de LED de 24 polegadas com a resolução 1920px x 1080px.

### 5.2 Método

Para iniciar o rastreamento ocular, o participante se sentou a uma distância de 50 cm da tela e o equipamento (Gazepoint GP3 eye tracker) foi calibrado para cada participante. No primeiro momento o participante deveria olhar para a expressão apresentada na tela para tentar memorizá-la, sem limite de tempo, sinalizando quando pudéssemos retirar a expressão da tela. Essa tarefa foi desenhada para responder às questões de pesquisa 1: Como a quantidade de nós MathML se relaciona com o processamento cognitivo da expressão matemática? e a questão 2: Como a complexidade de uma expressão matemática se relaciona com o processamento cognitivo da expressão? Após o processo de memorização, o participante deveria apresentar a expressão em voz alta sem acesso visual à tela. Caso os participantes tivessem dificuldade em lembrar a expressão, poderiam pedir para colocar na tela novamente e tentar memorizá-la outra vez. Esta etapa teve como objetivo responder às questões de pesquisa 3: Quantos nós da árvore MathML podem ser armazenados na memória de curto prazo? e questão 4: Quantas

repetições são necessárias para memorizar a expressão ou subexpressão matemática? Ao final da apresentação de cada expressão matemática, perguntamos aos participantes o quanto eles estavam confiantes que a expressão estava correta em uma escala de 1 a 5. Embora não tenhamos definido uma questão de pesquisa para esta tarefa, o objetivo foi identificar se a confiança do participante aumentou com as repetições da expressão e relacionar o grau de confiança com a apresentação correta da expressão.

## 6. Análises e Resultados

Para análise dos resultados do experimento da Pausa, foram consideradas as seguintes variáveis: (I) tipo de síntese, como a expressão matemática foi sintetizada, com o novo modelo ou Audiomath; (II) expressão resposta 1, representa se o participante acertou ou errou a primeira questão sobre partes da expressão matemática escutada; (III) expressão resposta 2, representa se o participante acertou ou errou a segunda questão sobre partes da expressão matemática; (IV) confiança é uma variável ordinal (escala de 1 a 5) que representa o quão confiante o estudante está em sua resposta, (V) familiaridade que representa o quão familiar a expressão matemática é para o aluno.

Realizamos o teste de McNemar para analisar se havia diferença significativa entre os erros e acertos dos alunos nas questões em que perguntamos partes específicas de expressões. Escolhemos este teste porque a amostra é pareada e as variáveis são nominais. Embora as diferenças percentuais foram pequenas (Modelo: acerto 87,5% e erro 12,5%, Audiomath: acerto 85% e erro: 15%), houve mais acertos e menos erros nas expressões matemáticas sintetizadas com o modelo de pausa proposto. O teste de McNemar mostrou que existem diferenças significativas nas proporções de acertos e erros entre a expressão matemática sintetizada com o modelo de pausa e o Audiomath para a questão 1 ( $\chi^2(1) = 21,6$ ;  $p < 0,001$ ) e para a questão 2 ( $\chi^2(1) = 11,3$ ;  $p < 0,001$ ). Analisamos também no experimento das pausas o nível de confiança dos participantes para ver se eles marcaram a resposta correta porque sabiam a resposta ou tentaram adivinhar. Dos 61 acertos, em 59 os alunos tinham certeza. Analisamos a frequência de familiaridade das expressões relacionando-a com as respostas certas e erradas. A familiaridade é ordinal, onde 1 significa nenhuma familiaridade e 5 é totalmente familiar. Observamos que a maioria das respostas erradas está relacionada com a falta de familiaridade com as expressões matemáticas. Constatamos que o número de acertos com o modelo de pausa foi maior entre os participantes cujas expressões matemáticas não eram familiares, ao mesmo tempo, o número de erros foi maior com Audiomath. No teste de percepção, o participante ouviu duas expressões matemáticas idênticas: uma sintetizada com o modelo de pausa e outra com o Audiomath. Das 36 respostas, 22 relataram que o áudio sintetizado com o modelo de pausa auxiliou mais na compreensão. Em comparação, 8 respostas preferiram o AudioMath e 6 não perceberam diferença entre os dois áudios.

Para a análise do experimento de rastreamento ocular seguimos a escolha das variáveis de rastreamento dos trabalhos de André et al. (2015), Plummer et al. (2017) e Kohlhase e Kohlhase (2021), que analisam as variáveis: quantidade de fixações, duração da fixação e duração média da fixação do olhar.

As seguintes variáveis foram então definidas neste estudo: (V1) quantidade de nós da árvore MathML; (V2) repetição, representa o número de vezes que os participantes pediram para repetir a expressão durante o processo de memorização;

Complexidade cognitiva, dividida em (V3) complexidade cognitiva indexada e (V4) complexidade cognitiva simplificada, calculada com base no modelo proposto por Bitter (2013); (V5) duração média das fixações, calculada a partir da duração total da fixação dividida pela quantidade de pontos de fixação durante observação da expressão pelo participante; Grau de confiança, na escala de 1 a 5, representando o quão confiante o participante estava da sua resposta, onde 1 significa nada confiante e 5 totalmente confiante, nesse caso consideramos (V6) grau de confiança antes de repetir a expressão e (V7) grau de confiança depois de falar a expressão; Correção para avaliar o quanto corretamente o participante falou a expressão, definimos uma escala de 1 a 5, com base na porcentagem de acertos, onde 1 representa 20%, 2 - 40%, 3 - 60%, 4 - 80% e 5 representa 100%. Também analisamos (V8) correção antes do participante solicitar a repetição e (V9) correção geral (após a repetição).

Para executar os testes estatísticos descritos a seguir, utilizamos *Real Statistics* (Zaiontz 2022) por ser gratuito e adequado às necessidades de nossa análise. Para investigar a força da relação e como as variáveis deste estudo se relacionam com o processamento cognitivo de expressões matemáticas, criamos uma matriz de correlação. Usamos o coeficiente de correlação de Spearman (CS) em nossa análise porque nosso estudo tem variáveis contínuas e ordinais.

**Tabela 2. Coeficientes de Correlação de Spearman**

Coeficiente de correlação	V1-V2	V1-V5	V3-V5	V4-V5	V5-V7	V2-V7	V3-V2	V2-V9	V7-V9
Spearman	0,37	0,54	0,82	0,71	-0,98	-0,91	0,70	-0,85	0,98
<b>Teste de coeficiente de correlação</b>									
t-stat	0,89	1,41	3,22	2,28	-11,620	-4,78	2,2	-0,49	10,17
Valor p	0,41	0,21	0,02	0,07	0,00008	0,005	0,08	0,001	0,000

Para responder à questão 1, “Como a quantidade de nós MathML se relaciona com o processamento cognitivo da expressão matemática?”, analisamos a relação entre a quantidade de nós (V1) e a duração média das fixações (V5), mas os resultados não foram conclusivos (CS = 0,54, valor p 0,21 – Tabela 2). Analisamos também a relação entre a quantidade de nós (V1) e a repetição (V2), mas a correlação foi fraca (CS = 0,37) e o valor  $p > 0,05$ , portanto os resultados obtidos não foram suficientes para obter evidências que respondam a essa questão.

Para responder à questão 2 “Como a complexidade de uma expressão matemática se relaciona com o processamento cognitivo da expressão?”, analisamos os coeficientes da correlação complexidade indexada (V3) e a duração média da fixação (V5). Os resultados mostraram que há uma relação forte, positiva e estatisticamente significativa (CS = 0,82, valor p 0,02 – Tabela 2), indicando que conforme a complexidade indexada de uma expressão matemática aumenta, o processamento cognitivo também aumenta, portanto, podemos concluir que a complexidade indexada sempre deve ser considerada no processo de síntese da fala de expressões matemáticas. Para complexidade simplificada (V4), o valor de p foi maior que 0,05 ( $p = 0,07$ ), portanto, não foi possível obter evidências significativas e conclusivas.

Para a questão 3, “Quantos nós da árvore MathML podem ser armazenados na memória de curto prazo?”, verificamos na questão 1 que não há correlação estatisticamente significativa entre as variáveis quantidade de nós (V1), duração da fixação (V5) e repetição (V2), entretanto, ao analisar os dados de algumas expressões, observamos que havia expressões com complexidade indexada maior em que os participantes não solicitaram a repetição, o mesmo aconteceu para o número de nós.

Verificamos também que foi possível armazenar na memória de curto prazo dez nós de única vez, quando a complexidade indexada da expressão foi em torno de 22, o que pode ser considerada baixa em relação a complexidade das demais expressões analisadas. Não foi possível quantificar as repetições necessárias para armazenar uma expressão matemática na memória para responder à questão 4, pois os valores encontrados para a correlação quantidade de nós (V1), repetição (V2) e complexidade indexada (V3) não foram significativos (Tabela 2). Entretanto, observamos que o grau geral de confiança (V7) e a repetição (V2) possuem uma relação muito forte, inversa e significativa ( $CS=-0,91$ ;  $p\text{-valor} = 0,005$  - Tabela 2). Portanto, podemos concluir que há evidências de que, à medida que o participante se sente mais confiante, é necessária menos repetição para memorizar a expressão matemática.

Observamos também que numa repetição, os pontos de fixação eram direcionados para partes específicas da expressão, não para a expressão inteira. Logo, percebemos que a repetição funciona como um reforço da memória, que pode variar de uma pessoa para outra. Semelhantemente, pessoas com deficiência visual também precisam de controles para implementar a eventual decisão sobre qual parte da expressão matemática precisam reforçar em sua memória para mentalizar totalmente a expressão. Analisando as demais variáveis, verificamos que os coeficientes de correlação entre o grau de confiança geral (V7) e a duração média da fixação (V5) mostraram que essas duas variáveis possuem uma relação muito forte, inversa e estatisticamente significativa ( $CS=-0,98$ ;  $p= 0,00008$  - Tabelas 2). O acerto geral (V9) também tem uma relação extremamente forte e significativa com o nível geral de confiança do participante (V7) ( $CS = 0,98$ , valor de valor  $p = 0,00$  - Tabela 2). Portanto, há indícios de que o participante fala a expressão mais corretamente quando o nível de confiança é maior.

## 7. Conclusão e Trabalhos Futuros

Este artigo apresentou um estudo de rastreamento ocular e um modelo para gerar pausas dinamicamente para a fala sintetizada de expressões matemáticas. Esse trabalho teve como objetivo analisar e estabelecer relações entre variáveis que podem ser utilizadas para melhorar o processo cognitivo do estudante com deficiência visual na recepção da fala sintetizada das expressões matemáticas.

Embora algumas respostas não tenham sido conclusivas no estudo de rastreamento ocular, essa pesquisa aponta direções para trabalhos futuros que podem revelar como minimizar o processamento cognitivo e a carga mental da recepção da fala sintetizada das expressões matemáticas: (I) observamos que algumas expressões de menor complexidade foram repetidas mais vezes do que outras de maior complexidade, o mesmo fenômeno foi observado para o número de nós. Acreditamos que os tipos de simbologia matemática (diversidade de elementos) envolvidos em uma expressão devem ser investigados, pois a complexidade e o número de nós não são os únicos fatores que contribuem para a dificuldade de memorização; (II) é necessário criar um modelo de navegação na árvore MathML que permita o estudante expandir ou reduzir a expressão matemática de acordo com a necessidade de reforço da memória. As pessoas com deficiência visual precisam ter autonomia para gerenciar qual parte da expressão matemática gostariam de repetir sem ultrapassar os limites de complexidade que a memória pode armazenar.

Com relação ao modelo de pausas podemos concluir que, embora o modelo apresentado não seja perfeito e não elimine erros, alguns avanços foram observados no experimento realizado: (I) o teste de McNemar mostrou que houve uma pequena vantagem nos acertos e erros com o modelo de pausa em questões relacionadas a expressões matemáticas; (II) o teste de percepção mostrou que a maioria dos alunos prefere a expressão matemática sintetizada com o modelo de pausa do que a pausa Audiomath; (III) constatamos que um dos fatores para erros no estudo de expressões matemáticas pode ser a baixa familiaridade com o conteúdo, portanto, nessas situações, o modelo de pausa teve melhor desempenho.

Limitações nesse estudo criam oportunidades para trabalhos futuros. 1) No estudo das pausas, o modelo proposto foi testado para frações, portanto expandir para outras estruturas matemáticas é necessário. 2) A amostra utilizada para a construção do modelo baseou-se em cinco falas; portanto, aumentar a amostra pode melhorar o modelo de regressão linear. Com relação ao estudo do rastreamento ocular, 3) O experimento foi realizado com poucas expressões matemáticas, portanto, para trabalhos futuros, o estudo pode abranger uma quantidade maior e mais diversificada de expressões, incluindo outros tipos de expressões de cálculo como somatório, derivada e integral.

## Referências

- Andrá, Chiara, Paulina Lindström, Ferdinando Arzarello, Kenneth Holmqvist, Ornella Robutti, e Cristina Sabena. 2015. “Reading Mathematics Representations: An Eye - Tracking Study”. *International Journal of Science and Mathematics Education* 13:237–59.
- Bansal, Akashdeep e M. Balakrishnan. 2021. “Evaluating Cognitive Complexity of Algebraic Equations”. *The Journal on Technology and Persons with Disabilities* 170.
- Bates, Enda e Donal Fitzpatrick. 2010. “Spoken Mathematics using Prosody , Earcons and Spearcons 2 Visual and Audible Representations of Mathematical Material”.
- Bitter, Marc. 2013. “Braille in Mathematics Education”. Radboud University Nijmegen Braille.
- Boersma, Paul e David Weenink. 2018. “Praat: doing phonetics by computer”.
- Campione, Estelle e Jean Véronis. 2002. “A Large-Scale Multilingual Study of Silent Pause Duration”. *Speech Prosody 2002*, abril, 4.
- Carlisle, David, Patrick Ion, e Robert Miner. 2014. *Mathematical markup language (MathML) version 3.0 2nd edition, W3C recommendation, W3C*.
- Chomsky, Noam. 1964. “Formal discussion of the development of grammar in child language by W. Miller & S. Ervin”. *The Acquisition of language*.
- Chomsky, Noam. 2009. *Syntactic structures*. De Gruyter Mouton.
- Ernest, Paul. 1987. “a Model of the Cognitive Meaning of Mathematical Expressions”. *British Journal of Educational Psychology* 57(3):343–70.
- Ferreira, Helder e Diamantino Freitas. 2004. “Enhancing the accessibility of mathematics for blind people: The audiomath project”. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 3118:678–85.
- Frankel, Lois e Beth Brownstein. 2016. “An Evaluation of the Usefulness of Prosodic and Lexical Cues for Understanding Synthesized Speech of Mathematics”. *ETS Research Report Series* 2016(2):1–19.
- Frankel, Lois, Beth Brownstein, Neil Soiffer, e Eric Hansen. 2016. “Development and

- Initial Evaluation of the ClearSpeak Style for Automated Speaking of Algebra”. *ETS Research Report Series* 2016(2):1–43.
- Hartmann, Matthias. 2015. “Numbers in the eye of the beholder: What do eye movements reveal about numerical cognition?”. *Cognitive Processing* 16:245–48.
- Kintsch, Walter. 1978. “Comprehension and memory of text”. *Handbook of learning and cognitive processes* 6:57–86.
- Kohlhase, Andrea e Michael Kohlhase. 2021. “Conceptual Design for an Eye-Tracking Experiment on Formula Linebreaking”. in *14th Conference on Intelligent Computer Mathematics (CICM 2021)*.
- Maier, H. R., Z. Kapelan, J. Kasprzyk, J. Kollat, L. S. Matott, M. C. Cunha, G. C. Dandy, M. S. Gibbs, E. Keedwell, A. Marchi, A. Ostfeld, D. Savic, D. P. Solomatine, J. A. Vrugt, A. C. Zecchin, B. S. Minsker, E. J. Barbour, G. Kuczera, F. Pasha, A. Castelletti, M. Giuliani, e P. M. Reed. 2014. “Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions”. *Environmental Modelling & Software* 62:271–99.
- Nazemi, Azadeh, Iain Murray, e Nazanin Mohammadi. 2012. “Mathspeak: An Audio Method for Presenting Mathematical Formulae to Blind Students”. P. 48--52 in *5th International Conference on Human System Interactions*.
- O’Malley, Michael H., Dean R. Kloker, e Benay Dara-Abrams. 1973. “Recovering Parentheses From Spoken Algebraic Expressions”. *IEEE Transactions on Audio and Electroacoustics* 21(3):217–20.
- Plummer, Patrick, Melissa DeWolf, Miriam Bassok, Peter C. Gordon, e Keith J. Holyoak. 2017. “Reasoning strategies with rational numbers revealed by eye tracking”. *Attention, Perception, and Psychophysics* 79(5):1426–37.
- Rayner, Keith. 1998. “Eye movements in reading and information processing: 20 years of research”. *Psychological Bulletin* 124:372.
- Reich, S. S. 1980. “Significance of pauses for speech perception”. *Journal of Psycholinguist* 379–389.
- Röder, Brigitte, Frank Rösler, e Helen J. Neville. 2001. “Auditory memory in congenitally blind adults: a behavioral-electrophysiological investigation”. *Cognitive Brain Research* 11(2):289–303.
- Rumelhart, David e Donald A. Norman. 1981. “Analogical processes in learning”. *Cognitive skills and their acquisition* 335–59.
- Rumsey, Deborah J. 2016. “How to interpret a correlation coefficient r”. *Statistics For Dummies*.
- Schindler, Maike, · Achim, e J. Lilienthal. 2022. “Students’ collaborative creative process and its phases in mathematics: an explorative study using dual eye tracking and stimulated recall interviews”. 54:163–78.
- Souza, Adriana e Diamantino Freitas. 2018. “Tecnologias Assistivas para Apoiar o Ensino e Aprendizagem de Pessoas com Deficiência Visual na Matemática: Uma Revisão Sistemática da Literatura”. P. 923 in *Anais do XXIX Simpósio Brasileiro de Informática na Educação (SBIE 2018)*. Vol. 1. Brazilian Computer Society (Sociedade Brasileira de Computação - SBC).
- Stevens, Robert David. 1996. “Principles for the design of auditory interfaces to present complex information to blind people”. University of York.
- Zaiontz, Charles. 2022. “Real statistics using Excel. 2015”. in *[Computer software]* <http://www.realstatistics.com/free-download/> (accessed 2 May 2022).