

19 99291-9777



# Programa de Pós-Graduação em **LINGUÍSTICA**

## **EFEITOS DA FREQUÊNCIA LEXICAL NA INTERPRETAÇÃO DE SITUAÇÃO-PROBLEMA EM MATEMÁTICA**

São Carlos 2024



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

EFEITOS DA FREQUÊNCIA LEXICAL NA INTERPRETAÇÃO DE  
SITUAÇÃO-PROBLEMA EM MATEMÁTICA

Tainara Duro Agostini

Tese apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutora em Linguística.

Orientador: Prof. Dr. Dirceu Cleber Conde  
Co-orientadora: Prof. Dra. Amanda Post da Silveira

São Carlos  
2024

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Educação e Ciências Humanas  
Programa de Pós-Graduação em Linguística

**Folha de aprovação**

Assinatura dos membros da comissão examinadora que avaliou e aprovou a Defesa de Doutorado da candidata Tainara Duro Agostini, realizada em 27/set/2024:

---

Prof. Dr. Renato Miguel Basso  
Universidade Federal de São Carlos (UFScar)

---

Profa. Dra. Renata Prensteter Gama  
Universidade Federal de São Carlos (UFScar)

---

Prof. Dr. Thiago Oliveira de Motta Sampaio  
Universidade Estadual de Campinas (UNICAMP)

---

Profa. Dra. Janaina Weissheimer  
Universidade Federal do Rio Grande do Norte (UFRN)

*Aos meus pais que sempre me apoiaram incondicionalmente em todas as etapas e as escolhas da minha vida.*

## AGRADECIMENTOS

Primeiramente, agradeço a Deus e a todas as entidades de luz por terem me dado forças para continuar a ler e a escrever em um período extremamente conturbado para mim.

Agradeço, em especial, ao Cleber e à Amanda. Ao Cleber, por me acompanhar desde o mestrado sendo sempre respeitoso, cuidadoso e muito amigo. Agradeço por sempre me incentivar e questionar sobre tudo – questionamentos que, em um primeiro momento pareciam obscuros, mas depois eram as respostas das quais precisava. Obrigada pelo amparo, pelas discussões, pelas oportunidades e por ser exemplo de um ótimo professor e orientador que eu quero seguir. À Amanda por ter aceitado encarar este projeto com a gente e me fazer enxergar meu tema para além da semântica, caminho que jamais imaginaria percorrer. Agradeço pelo apoio, pelas leituras enriquecedoras e pelo olhar carinhoso e exigente ao mesmo tempo, o que me desafiou e me fez sentir capaz. Só tenho a agradecer.

Agradeço aos meus pais que sempre me apoiam e me dão todo o suporte necessário para qualquer caminho que eu escolha trilhar. Sou imensamente grata por eles me mostrarem o lado da luz quando vejo só a escuridão, por me darem carinho ao mesmo tempo sabedoria e por me ensinarem a ser grata por tudo. Muito, muito obrigada. Em especial agradeço à minha mãe que, até nos piores momentos desses últimos anos, mostrou-se sempre positiva e crente de que daria conta, mesmo quando eu mesma não acreditava.

Agradeço aos meus irmãos, Taciane e Tainan, por serem quem são: parceiros, riso em todos os momentos e força naqueles turbulentos.

Agradeço ao meu namorado, Cauê, por todo suporte, incentivo, apoio e por acreditar no meu potencial. Além disso, também sou grata por não me deixar esquecer que existe vida fora da academia e que também é preciso vivê-la, o que me fez ter dias mais leves e felizes.

Agradeço a todos os meus amigos que me animaram e estiveram comigo nas alegrias e nas tristezas, sempre me lembrando de quem eu sou e aonde quero chegar.

Agradeço imensamente ao Isaac, expert em tecnologia, que foi luz diante do desespero quando precisei trabalhar com os dados. Sempre muito solícito e ágil, serei sempre muito grata.

Agradeço aos membros da banca de defesa e da banca de qualificação, prof. Marcus, prof. Janaina e prof. Renato, prof. Thiago, prof. Renata por terem aceitado ler o meu texto e contribuírem com a minha formação como linguista. Obrigada por também embarcarem com a gente nesse mundo de situações-problema.

Por fim, agradeço ao PPGL da UFSCar pela minha formação e por todas as oportunidades. E, também, à CAPES pela concessão da bolsa, principalmente por ter sido contemplada durante um governo contrário à educação.

Agradeço a todos que fizeram parte desta jornada, que tenhamos uma vida longa, bem vivida e próspera.

## RESUMO

O presente trabalho possui como objetivo averiguar os fatores linguísticos influenciadores na dificuldade de interpretação de situações-problema em matemática vivenciada por alunos de 6º ao 9º anos. Essa busca se dá, pois, a fim de que se chegue a um resultado ou ao entendimento de um texto matemático, principalmente exercícios, torna-se necessário ler e compreender, isto é, existe um processo de tradutibilidade da língua natural para linguagem matemática, um intercâmbio de códigos. Entretanto, pelo fato de palavras incomuns aos alunos serem utilizadas na composição das situações-problema, os estudantes enfrentam dificuldades para conseguirem resolvê-las. Nesse sentido, observamos que o vocabulário é deveras importante, uma vez que o aluno precisa conhecer o objeto linguístico com o qual está tendo contato para que seja possível a tradução da linguagem natural para a linguagem matemática. Partindo da hipótese de que o desconhecimento vocabular possa ter efeitos sobre a interpretação, construímos um corpus do gênero situação-problema para analisar a escolha e o uso das palavras, a fim de averiguar quais palavras estão sendo utilizadas e sua frequência, de modo a investigar se a escolha dessas palavras influencia no processo interpretativo do aluno. A fim de avaliar a frequência desses termos, trabalhamos com análise comparativa entre alguns corpora: de situações-problema), de palavras frequentes às crianças brasileiras e de termos comum a adultos. Após o trabalho com os corpora, desenvolvemos um experimento para testar a familiaridade dos alunos com as palavras que constam no livro de matemática e, depois, se eles conseguiriam encaixar a palavra dentro do contexto do enunciado. Com isso, averiguamos que de fato as palavras incomuns atrapalham o processo interpretativo dos alunos, isto é, a escolha lexical foi identificada como um desafio significativo que afeta o processo interpretativo dos alunos e o desempenho escolar.

**Palavras-chave:** Interface Linguagem-Matemática. Escolha Lexical. Frequência De Palavras.

## **ABSTRACT**

The objective of this study is to investigate the linguistic factors that influence the difficulty in interpreting math word problems faced by students from 6th to 9th grade. This inquiry is necessary because, to reach a solution or understanding of a mathematical text, especially exercises, it is essential to read and comprehend. In other words, there is a process of translating from natural language to mathematical language, an exchange of codes. However, due to the use of words unfamiliar to the students in the composition of word problems, they encounter difficulties in solving them. In this context, we observe that vocabulary is indeed important, as the student needs to know the linguistic object they are dealing with to enable the translation from natural language to mathematical language. Based on the hypothesis that a lack of vocabulary knowledge might affect interpretation, we constructed a corpus of word problem texts to analyze word choice and usage, aiming to determine which words are being used and their frequency to investigate whether the choice of these words influences the student's interpretative process. To assess the frequency of these terms, we performed a comparative analysis between different corpora: word problems, frequent words for Brazilian children, and common terms for adults. After working with the corpora, we developed an experiment to test students' familiarity with the words in the mathematics book and then to see if they could place the word within the context of the statement. This led us to confirm that indeed unfamiliar words hinder students' interpretative processes, indicating that lexical choice is a significant challenge that affects students' interpretative process and academic performance.

**Keywords:** Language-Mathematics Interface. Lexical Choice. Word Frequency.

## LISTA DE FIGURAS

Figura 1 Coleção Bianchini para 6° anos.....	45
Figura 2 Coleção Bianchini para 7° anos.....	45
Figura 3 Coleção Bianchini para 8° anos.....	45
Figura 4 Coleção Bianchini para 9° anos.....	45
Figura 5 Formulário de teste de familiaridade .....	57
Figura 6 Teste de leitura e sentido (6° anos).....	59
Figura 7 Correlação entre CSPM e LexPorBr Infantil.....	70
Figura 8 Gráfico de correlação entre CSPM e Corpus Fonotático PB.....	70
Figura 9 Correlação entre Corpus Fonotático PB e LexPoxBr Infantil.....	71

## LISTA DE TABELAS

Tabela 1 Quantidade de enunciados .....	47
Tabela 2 Quantidade de palavras .....	48
Tabela 3 Referência de valores de frequência .....	53
Tabela 4 Frequência alta no CSPM e baixa no Corpus Infantil .....	53
Tabela 5 Baixa frequência no CSPM e alta frequência no Corpus Infantil.....	54
Tabela 6 Baixa frequência em ambos corpora .....	55
Tabela 7 Frequentes em ambos corpora.....	56
Tabela 8 Palavras dos enunciados por série.....	58
Tabela 9 Correlação de Pearson e valor de P.....	66
Tabela 10 Correlação do coeficiente R entre os corpora CSPM e LexPorBr Infantil.	67
Tabela 11 Correlação do coeficiente R entre os corpora CSPM e Corpus Fonotático do Português Brasileiro .....	68
Tabela 12 Correlação do coeficiente R entre os corpora Corpus Fonotático do Português Brasileiro e LexPorBr Infantil .....	69
Tabela 13 Termos e suas frequências.....	71
Tabela 14 Porcentagem por categoria.....	73
Tabela 15 Frequência das palavras categorizadas no CSPM .....	73
Tabela 16 Indicação de conhecimento ou desconhecimento de um termo por ano escolar.....	74
Tabela 17 Palavras que erraram.....	75
Tabela 18 Palavras desconhecidas .....	77

## LISTA DE ABREVIATURAS

SiP	Situação-problema
CSPM	Corpus de Situação-Problema em Matemática
PB	Português Brasileiro
LD	Livro didático

# SUMÁRIO

<b>INTRODUÇÃO.....</b>	<b>12</b>
<b>1 PRESSUPOSTOS TEÓRICOS .....</b>	<b>17</b>
<b>1.1 LÉXICO E COMPOSICIONALIDADE .....</b>	<b>17</b>
1.1.1 A relação da semântica lexical com a aquisição léxico-semântica de primeira língua/ língua materna .....	21
1.1.2 A relação entre aquisição e competência lexical.....	28
1.1.3 O léxico: entre o letramento e a literacia .....	30
1.1.4 “Ler é em partes sobre palavras” .....	32
<b>1.2 SENTIDO E REFERÊNCIA: MEANDROS DO LÉXICO E DA COMPOSICIONALIDADE .....</b>	<b>34</b>
<b>2 A SITUAÇÃO-PROBLEMA.....</b>	<b>39</b>
<b>2.1 A TAL DA SITUAÇÃO-PROBLEMA.....</b>	<b>39</b>
<b>2.2 QUANDO A SiP SE TORNA UM "PROBLEMA"? .....</b>	<b>41</b>
<b>3 MATERIAIS E MÉTODOS .....</b>	<b>43</b>
<b>3.1 OBJETOS DE ESTUDO: O CORPUS E OS TESTES .....</b>	<b>43</b>
3.1.1 A coleção .....	44
<b>3.2 O CORPUS .....</b>	<b>47</b>
3.2.1 Caracterização do corpus.....	49
<b>3.3 Análise contrastiva de corpus .....</b>	<b>50</b>
3.3.1 Corpus fonotático do português brasileiro.....	51
3.3.2 LexPorBr Infantil.....	51
<b>3.4 OS TESTES .....</b>	<b>52</b>
3.4.1 Composição e aplicação do teste de familiaridade .....	53
3.4.2 Composição e aplicação do teste de leitura das SiPs .....	57
3.4.3 <i>Participantes</i> .....	59
<b>4 ANÁLISE DOS DADOS .....</b>	<b>60</b>
<b>5 RESULTADOS .....</b>	<b>66</b>
<b>6 REFLEXÕES E ENCAMINHAMENTOS.....</b>	<b>79</b>
<b>CONSIDERAÇÕES FINAIS .....</b>	<b>83</b>
<b>BIBLIOGRAFIA .....</b>	<b>85</b>
<b>ANEXO I.....</b>	<b>88</b>
<b>ANEXO II.....</b>	<b>90</b>
<b>ANEXO III.....</b>	<b>91</b>
<b>ANEXO IV.....</b>	<b>93</b>

## INTRODUÇÃO

Para apresentarmos a temática desta tese, faz-se necessário partirmos da seguinte afirmação: a matemática, como qualquer disciplina ou campo do conhecimento, vale-se de uma linguagem e, por isso, falamos de “linguagem matemática”. Essa afirmação pode soar não muito familiar para o senso comum, haja vista a noção prototípica de linguagem. Isso se deve ao fato de que raramente o estudo da matemática seja encarado como uma incursão em uma linguagem. Contudo, podemos usar a seguinte lógica: se a linguagem é um conjunto de sinais ou códigos convencionados que permitem a transmissão de uma mensagem, logo a matemática pode se valer de uma linguagem, dado que é composta por códigos e sinais convencionados que possuem o objetivo de transmitir uma mensagem.

Ao nos voltarmos para as linguagens matemáticas, podemos empreender uma visada linguística fazendo analogias, por exemplo, assim como os grafemas de um alfabeto fazem parte de uma linguagem criada para a comunicação escrita – conjunto de sinais escritos convencionados – a linguagem matemática também possui um conjunto de sinais e símbolos que representam elementos de um determinado universo de referência, ou seja, um mundo tanto para determinar quantidade quanto para indicar algum processo lógico-dedutivo. Assim como uma língua natural promove a comunicação-interação entre indivíduos em que os proferimentos permitem deduções (causas e consequência), implicações de subentendidos (implicaturas, p. ex.) dentre muitas outras funções, a linguagem matemática também permite aos indivíduos comunicar relações dedutivas e outras inferências.

A incompreensão da matemática como linguagem (que permeia o senso-comum), por vezes, provoca oposições muito artificiais que se refletem na educação, como na crença de que Língua Portuguesa e Matemática são conhecimentos opostos e inconciliáveis. Lorensatti (2009) e Granell (2003) têm abordado exatamente essa ausência de diálogo entre as áreas e esse efeito nos educandos, pois, a fim de que se chegue a um resultado ou ao entendimento de um texto matemático, principalmente quando se trata de exercícios, torna-se necessário ler e compreender, isto é, existe um processo de tradutibilidade da língua natural, no nosso caso o português brasileiro<sup>1</sup>, para a linguagem matemática, um intercâmbio de códigos.

---

<sup>1</sup> Utilizaremos o termo PB para se referir a português brasileiro

Diante de duas línguas(gens) intercambiáveis, vemos a necessidade de um letramento e uma literacia eficientes em ambas, entendendo-as como objetos complementares e não como códigos opostos. Para além disso, acreditamos ser importante o entendimento do enunciado em língua natural para que ele seja um facilitador no momento em que será “traduzido” e interpretado da língua natural para a linguagem matemática, como também destaca Lorensatti (2009, p. 91):

A linguagem matemática pode ser definida como um sistema simbólico, com símbolos próprios que se relacionam segundo determinadas regras. Esse conjunto de símbolos e regras deve ser entendido pela comunidade que o utiliza. A apropriação desse conhecimento é indissociável do processo de construção do conhecimento matemático. Está compreendido, na linguagem matemática, um processo de ‘tradução’ da linguagem natural para uma linguagem formalizada, específica dessa disciplina, segundo Granell (2003). Os enunciados emitidos em língua natural passam a ser escritos para o equivalente em símbolos matemáticos. Essa tradução ‘é o que permite converter os conceitos matemáticos em objetos mais facilmente manipuláveis e calculáveis’.

Partindo desse ideário, observamos que o vocabulário é deveras importante no entendimento de problemas matemáticos, uma vez que o aluno precisa conhecer o objeto linguístico com o qual está tendo contato para que seja possível a tradução da linguagem natural para a linguagem matemática. Nesse sentido, só existe comunicação se os elementos utilizados são conhecidos por todos, isso é o que consta como significado de linguagem. Entretanto, em pesquisa anterior (AGOSTINI, 2019), investigamos que um dos obstáculos vivenciados diariamente pelos alunos em relação ao processo interpretativo matemático é justamente o acesso a determinados vocábulos.

Agostini (2019), movida pelos baixos índices do Brasil no PISA, relata que, dentre outras variantes, os participantes da pesquisa indicaram o vocabulário como um problema, dado que havia palavras não-comuns às suas realidades, por exemplo, ‘prestação’, ‘montante’ e ‘capital’. Tal indicação incomoda, pelo fato de que o gênero escolhido para análise, assim como o do presente trabalho, é o denominado situação-problema, que possui como característica assuntos do cotidiano pessoal, escolar e profissional e sempre contextualizado de acordo com a faixa etária do aluno para que faça sentido o desafio proposto no mesmo gênero. Em virtude disso, é um pouco incoerente um gênero que tem como definição ser o mais próximo da vida do aluno ser composto por palavras desconhecidas por eles.

Embora o material do corpus de análise atual não seja o mesmo utilizado em Agostini (2019), pelo fato de termos criado um corpus somente de situação-problema<sup>2</sup> para este atual trabalho, é possível notarmos que a questão vocabular se mantém, uma vez que encontramos usos que parecerem ser pouco familiares a alunos da faixa etária adequada a um livro de 6 ano do ensino fundamental II, como podemos ver a seguir:

1. Sabendo que, com um trator, Lúcio **ara** 3/20 de um terreno em um dia, responda:
  - a) De segunda-feira a sábado, que parte do terreno Lúcio consegue **arar**?
  - b) Considerando que no domingo ele descansa, quanto faltará **arar** na semana seguinte?
  - c) Ele conseguirá terminar na segunda-feira? Justifique sua resposta.  
(Bianchini, 2019)

O leitor adulto pode não encontrar dificuldade na leitura da SiP<sup>3</sup> acima, porém, uma criança urbana de 11 anos talvez desconheça o verbo ‘arar’. Será que um aluno de 11 anos sabe o que é arar? Faz parte de seu contexto? Não é um verbo muito específico a uma comunidade, por exemplo, rural?

O estudante precisa de muita familiaridade com a terminologia trabalhada na matemática e uma ampla referência de sentidos para que seja possível percorrer o processo interpretativo e conseguir acessar o que determinada SiP provoca. Todavia, ao não levar o vocabulário em consideração, pode ocorrer de o aluno possuir o conhecimento matemático, mas não conhecer as palavras que compõem a SiP (por exemplo, o verbo “arar” do exemplo (1) pelo fato de serem habilidades distintas, uma vez que a primeira é sobre saber calcular, enquanto a segunda é conhecer a referência de determinado objeto o mundo. Em outras palavras, um aluno pode possuir conhecimento matemático, devido ao estudo ou às situações do dia a dia às quais ele foi exposto, porém não estar familiarizado com a linguagem matemática ou com um léxico mais técnico ou específico de outras situações.

A interpretação está diretamente conectada ao referencial linguístico e matemático para que se realize a resolução, logo, ao não estabelecer essa conexão,

---

<sup>2</sup> O detalhamento sobre esse corpus será feito na seção Metodologia

<sup>3</sup> A partir deste momento, utilizaremos SiP para nos referirmos a “situação-problema”.

o aluno não compreende, prejudicando o que deve fazer e como deve proceder a operação. Sendo assim, quanto mais sentidos e conhecimentos sobre o mundo, mais relações podem ser feitas, o que justifica utilizarmos, como uma das bases teóricas deste trabalho, uma abordagem sobre sentido e referência que é bastante útil para pensar em ensino e em resolução de situações-problema em matemática, por um princípio interpretativo, de modo a associar mais sentidos às referências.

Após essa breve exposição do problema em questão, principalmente ao levar em consideração as constatações de pesquisa anterior (Agostini,2019), fica claro, então que as crianças enfrentam dificuldade em compreender os enunciados de SiP pelo fato de o vocabulário não ser comum a elas, de que os sentidos passam a não compor uma referência mínima para sua compreensão. Sendo assim, questionamos: as palavras de baixa frequência no corpus PB adulto - as quais aparecem nas SiP - são empecilhos? A frequência dessas palavras no corpus infantil em comparação ao adulto e ao de SiPs tem alguma relação? Como a frequência nos ajuda a entender a dificuldade de interpretação de SiP por parte dos alunos? Será que a posição na qual elas se encontram (contexto, situação ou comando)<sup>4</sup> interfere em algo?

Pensando, de início, no fator frequência, nossa hipótese é que sim, as palavras de baixa frequência podem interferir na leitura e interpretação das SiPs, uma vez que, ao se deparar com esses termos, o aluno, devido à falta de referência, não atingirá o sentido necessário para entender o contexto completo para a tarefa a ser realizada.

Para a análise desta tese, construímos um corpus composto por itens lexicais extraídos somente por SiPs em matemática, todas retiradas da coleção Bianchini (2018), focada no público-alvo do Ensino Fundamental II, ou seja, alunos de 10 a 15 anos. O corpus foi constituído de acordo com as classes gramaticais de cada palavra – morfologia –, pois nosso foco é somente em palavras de conteúdo, como substantivos, adjetivos e verbos, e, para isso, era necessário, antes, classificá-los para escolhermos quais termos seriam importantes para nossa análise.

Como o nosso objetivo é averiguar a frequência desses termos, trabalharemos com mais dois corpora para a análise: um corpus infantil – LexPorBr Infantil – e um corpus adulto de palavras do português – Corpus do Português Brasileiro com variáveis Fonéticas e Psicolinguísticas – os quais descreveremos com mais detalhes no capítulo de Materiais e Métodos. Consideramos o recurso a esses dois corpora

---

<sup>4</sup> Esses termos serão mais bem explicados no capítulo 2

importante para termos uma referência de frequência mais ampla e geral da língua, assim teremos uma boa base de comparação.

Para que consigamos alcançar nossa finalidade e compreender melhor o objeto de estudo, no capítulo 1, apresentaremos uma ampla discussão que percorre desde o processo de aquisição de linguagem até o conceito de composicionalidade, passando pela aquisição e competência lexical e suas conexões com o tema, além da relação de sentido e referência, conforme proposta por Frege (1978).

No capítulo 2, indicaremos o objetivo principal desta tese e nossa hipótese.

No capítulo 3, pelo fato de lidarmos com um gênero escolar muito específico, a situação-problema em matemática, faremos uma breve discussão sobre esse gênero, descrevendo suas características, contextos de uso, objetivos e de que forma ela pode se tornar, de fato, um problema no cenário estudantil.

No capítulo 4, apresentaremos quais foram os materiais e os métodos utilizados para o desenvolvimento desta pesquisa, evidenciando como o corpus foi construído, analisado, assim como quais são os outros corpora utilizados.

Nos capítulos 5 e 6, respectivamente, apresentaremos os dados obtidos, por meio de tabelas, a fim de visualizarmos melhor os corpora e o quanto, se de fato estão correlacionados; após isso, discutiremos os resultados que encontramos, apresentando-os por meio de tabelas e gráficos, com o fito de facilitar a leitura e interpretação.

# 1 PRESSUPOSTOS TEÓRICOS

## 1.1 LÉXICO E COMPOSICIONALIDADE

A significação de uma sentença só pode ser plenamente alcançada se o falante é capaz de estabelecer relações entre o signo linguístico - o sinal na concepção de Frege (2009) - e o seu referente no mundo.

Para ilustrar, Ferreira (2019), traz o exemplo de uso da seguinte sentença: “Está chovendo em Paris”. Segundo ele, se o falante souber o significado dessa sentença, terá aprendido algo sobre o mundo. Ou seja, saber o significado de uma sentença nos permite transitar da linguagem para o mundo e do mundo para a linguagem (FERREIRA, 2019, pág. 2), pelo menos em sentenças contingentes. Todavia, para que isso seja possível, torna-se essencial conhecer a língua em que o enunciado está escrito. No entanto, voltando o exemplo ao objeto desta tese, a lógica proposta por Ferreira (2019) não se faz realidade em muitos cenários educacionais brasileiros, dado que crianças do ensino fundamental do nosso país estão se deparando com termos que não conhecem, gerando dificuldades de interpretação das SiPs em matemática, logo não conseguindo acessar o sentido e a representação da realidade dentro desses enunciados. Assim, ao crermos que o significado tem um papel crucial na mediação entre linguagem e realidade, precisamos discutir isso.

“O significado de uma sentença é equivalente às suas condições de verdade” (FERREIRA, 2019, pág.3), essa afirmação para nós é importante, pois o léxico compõe as condições de verdade e é sobre ele a nossa investigação. Claro que saber o significado de uma sentença não significa saber se a sentença é verdadeira ou falsa, pois se trata de uma questão empírica. Porém, o significado de uma sentença poderá depender da posição na qual se encontra a palavra e o significado dela, ou seja, depende das relações que as palavras estabelecerão na sentença e o seu próprio significado.

Tomemos como exemplo a situação-problema a seguir:

2. Uma piscina está com 35.750 litros de água. Colocando-se outros 12.250 litros, ela ficará cheia. Quantos litros de água cabem nessa piscina? (BIANCHINI, 2018)

A SiP (2) foi retirada de um dos livros da coleção que analisamos, volume destinado a alunos com 11 anos. Em relação às palavras que compõem esse enunciado, são todas comuns aos alunos e é um enunciado tido como simples, pelo fato de requerer uma operação de soma entre dois números. Com certeza, qualquer falante de português fluente, incluindo crianças de 11 anos, não teria problema com o léxico desse problema. Agora vejamos a SiP (3):

3. Um cobizarro está com 62 plots de janil. Colocando-se mais 5 plots, ele ficará cheio. Quantos plots de janil azem no cobizarro?

Ao contrário da leitura facilitada da SiP (2), a da SiP (3) deve ter gerado algum estranhamento, uma vez que "plots", "janil", "cobizarro" e "azem" não são termos conhecidos, isso porque não fazem parte do português brasileiro. Por mais experiente que nós possamos ser em matemática, por sermos adultos e termos passado pelo processo de escolarização, sentimo-nos inseguros ao resolvê-la. Os termos não são conhecidos, e isso gera incerteza porque precisamos assumir alguns sentidos cuja referência não é clara.

Esse exemplo que elaboramos acima é uma breve simulação de como as crianças devem se sentir ao se depararem com enunciados compostos por termos que elas não conhecem. Isso quer dizer que elas podem até compreender a maior parte das palavras, entretanto um termo desconhecido, mesmo que não seja um elemento-chave para a resolução, já gera essa insegurança e faz com que a resolução fique comprometida.

O Princípio de Composicionalidade, atribuído a Frege<sup>5</sup>, se encontra em muitos manuais de semântica e lógica:

O sentido de uma expressão composta é a função do sentido de suas partes e o modo como ela foi combinada sintaticamente. (PARTEE, 1984, 281, tradução nossa)<sup>6</sup>

Segundo esse princípio [princípio de composicionalidade], o significado de uma expressão complexa depende de modo regular do significado das expressões que a compõem e do modo como elas são combinadas sintaticamente. (CHIERCHIA, 2003, p. 39)

---

<sup>5</sup> Para um maior aprofundamento da história do Princípio de Composicionalidade, incluindo suas controvérsias, recomendamos a leitura de Janssen (2012).

<sup>6</sup> No original: "The meaning of a compound expression is a function of the meanings of its parts and of the way they are syntactically combined".

(...) o chamado Princípio da Composicionalidade, ou Princípio de Frege: o significado de uma expressão complexa é uma função do significado de suas partes e do modo como elas se combinam. (MORTARI, 2001, p. 125)

Mesmo sendo um princípio transparente e operante, é mister observarmos que ele já supõe pelo menos dois pré-requisitos para operá-lo:

- i) que o operador do princípio (humano ou máquina) domine a extensão das partes;
- ii) que o operador domine as regras de composição;

Vamos focar apenas no primeiro princípio, e por isso podemos utilizar o mesmo exemplo de Mortari (2021, p. 121):

(1) Tübingen ist eine schöne Stadt.

Não é possível calcular se essa sentença é verdadeira ou falsa sem saber alemão. No entanto, imaginemos essa sentença semi traduzida:

(2) Tübingen é uma bela Stadt.

Parece que somos capazes de supor algum tipo de sentido sem, contudo, termos certeza.

Até chegarmos:

(3) Tübingen é uma bela cidade.

que nos permite dizer que a sequência fonológica /'ty:bɪŋən/ é o nome de uma cidade, embora ainda não se saiba onde ela fica ou se tenha uma noção de como ela é.

Conseguimos traduzir que Tübingen é o nome de uma cidade pelo significado das demais palavras e pelas regras de composição próprias do português. Mais uma vez, podemos afirmar que obviamente o sentido estabelecido pelos termos de uma sentença também são componentes do significado. Especificamente no caso de crianças, cujo vocabulário ainda depende de amadurecimento, a depender de como

um enunciado é composto e das seleções lexicais, pode-se ter um efeito de indeterminação do sentido e por isso uma dificuldade muito grande para se chegar a interpretações esperadas. Nesse sentido, a composicionalidade permite aos falantes de línguas naturais que possam, através do vácuo da fórmula, preencher de significados estruturas até então operativas, mas sem referência. A dependência do referente das partes para o referente do todo parece ser a condição e a operação para que o sentido seja calculado, muito embora, nas línguas naturais, algumas expressões idiomáticas possam parecer excluídas do Princípio de Frege<sup>7</sup>.

Segundo Ferreira (2019, pág. 7), "para obter o significado de um constituinte, basta olhar para o significado de seus constituintes imediatos, sem se preocupar com o significado das partes que compõem esses constituintes". Contudo, de acordo com pesquisa anterior (Agostini, 2019), os alunos não conseguem acessar o sentido da sentença somente pelas palavras que conhecem, dado que, ao se depararem com as desconhecidas, mantêm-se focados nelas e se confundem na resolução, logo o processamento da composicionalidade, para eles, tem muito mais chances de ser bem-sucedido se de fato todos os termos do enunciado forem conhecidos por eles, dada a necessidade de uma referência mínima.

Sendo assim, o léxico compõe as condições de verdade em uma visada empírica considerando a criança e sua fase de aprendizagem. É mister observar que a abordagem referencial neste momento tenta atingir o conceito de referência de alguns sintagmas. Como o objeto SiP não é uma situação conversacional das línguas naturais e por isso possui uma estrutura baseada em três etapas (contexto + situação + comando), o leitor desse tipo de enunciado não estabelece o juízo de condição de verdade, mas recebe a contextualização (conteúdo temático), situação e comando (Agostini, 2019), como ocorre a seguir:

4. (i) Severino aplicou um **capital** de R\$ 320,00, durante 2 meses, à taxa de juros simples de 0,7% ao mês. (ii) Mara aplicou um **capital** de R\$ 300,00 composta de 1% ao mês. (iii) No fim dos 2 meses, qual deles apresentou um **montante** maior? (DANTE, 2015, p. 254)

---

<sup>7</sup> Sobre composicionalidade de expressões idiomáticas, consultar Westerståhl (2002).

Em um enunciado como o do exemplo acima, as etapas base para um enunciado de SiP são presentes e indicados no próprio exemplo com (i) para contexto, (ii) para situação e (iii) comando.

No exemplo, o leitor não é levado a avaliar o valor de verdade, mas a atender, em última instância, o comando da SiP que é acionado pela interrogação ou por um imperativo. Isso, no entanto, não implica que determinados conteúdos não sejam relevantes, principalmente aqueles que indicam alguma referência ao leitor, por isso, entendemos que o vocabulário pode ser, em certa medida, um dificultador para a compreensão de enunciados.

Ao se deparar com o termo “capital” – não usual com o sentido monetário – no contexto, há um obstáculo interpretativo, o que se torna problemático, pois o contexto é como se fosse uma “ambientalização” para o leitor e, quando há algum empecilho, a leitura é comprometida desde então. Seguindo essa lógica, já há um termo sem referência. Continuando a leitura, “capital” aparece novamente na etapa situação, para, por fim, haver o comando, que é a etapa característica das SiPs por propor um desafio ao aluno. Nessa etapa, ele se deparará com o termo “montante”, palavra que, além de não constar em nenhum dos corpora analisados, também não é usual ao público infantil. Assim, o que temos é uma SiP que, para os alunos, é composta de palavras que para eles parecem pertencentes a outro idioma. Da mesma forma que o exemplo (2) de Mortari não faz sentido para quem não é falante de alemão, pois há composicionalidade, mas não há compreensão do sentido devido ao não domínio de alemão, logo, se não domina o código, não há como calcular o valor de verdade. Não é possível a criança preencher esse vácuo, desse modo, sem referência das partes não há referência do todo.

Por fim, o princípio de composicionalidade é intrínseco a esta tese, visto que o comprometimento do sentido das SiPs é gerado ao não fornecer aos alunos termos conhecidos por eles os quais façam sentido, fazendo com que não consigam combinar as partes durante a leitura, conseqüentemente engendrando como se fosse uma lacuna entre outros termos que são conhecidos.

### *1.1.1 A relação da semântica lexical com a aquisição léxico-semântica de primeira língua/ língua materna*

Quando falamos sobre o processo interpretativo, o vocabulário se torna um dos termos em destaque, pelo fato de a interpretação ser eficiente quando há conhecimento de uma palavra lida. Ao ler um enunciado matemático, principalmente por se tratar de uma área tida pelo senso comum como difícil, os alunos se apegam às palavras contidas nele, assim sua escolha se torna fundamental, pois é uma forma de apoio para o entendimento do que é pedido. Diante disso, compreender como se dá o processo de aquisição desses termos compostos de sentido, para nós, é imperioso.

Os estudos da aquisição da linguagem implicam em compreender como um indivíduo passa de não-falante a falante de sua língua, sendo assim são pautados na explicação e entendimento desse fenômeno. Segundo estudos da Linguística Cognitiva, o desenvolvimento linguístico de uma criança está relacionado ao desenvolvimento cognitivo, isso quer dizer que quanto mais a criança desenvolver sua linguagem, melhor comunicar-se-á e compreenderá as formas de comunicação dos outros indivíduos. Diante disso, a aquisição lexical é um processo que perdura por toda a vida do indivíduo, uma vez que se inicia com a produção das primeiras palavras e permanece, pois o falante sempre está aprendendo palavras novas. No entanto, para que esse processo seja contínuo, primeiro, precisamos compreender que as crianças chegam às escolas com bagagens cognitivas diferentes, claro que pode ser por questões de origem genética, mas, independente dessa questão, as condições sociais e econômicas estão diretamente relacionadas, principalmente em se tratando das situações às quais esse público é exposto. Em outras palavras, entendemos que a formação do repertório lexical das crianças e dos jovens vai depender das situações às quais serão expostos.

Nesse sentido, nos apoiaremos nos pressupostos de Vygotsky pelo fato de propor que o desenvolvimento da linguagem se dá com base na interação interpessoal, mais especificamente, entre adulto e criança. Sendo assim, fatores sociais, a saber, a comunicação, o meio, a cultura, são levados em consideração, ou melhor, são pré-requisitos nesse processo. Quanto mais a criança é exposta a diferentes contextos, mais linguagem e conhecimento de mundo adquire. É com base nessa linha de raciocínio que Xatara et al. (2008) acreditam que seja adquirido e ampliado o léxico, de acordo com as relações sociais e a escola, dessa forma, é um processo completamente social e dependente de contatos e do acesso à informação.

Esse contato é fundamental de ser levado em consideração, pois, segundo Sargiani & Maluf (2018), os seis primeiros anos de vida de uma criança são fundamentais, pois é nesse período em que elas desenvolvem diversas habilidades linguísticas e cognitivas, as quais, inclusive, são primordiais para o processo escolar, uma vez que

O cérebro das crianças praticamente quadruplica de tamanho nesse período atingindo cerca de 90% do volume do cérebro adulto por volta dos seis anos. Entretanto, a quantidade de neurônios é praticamente a mesma desde o nascimento, as conexões entre eles é que se modificam, aumentam ou são eliminadas em virtude das experiências ambientais e culturais (Brown & Jernigan, 2012 apud Sargiani & Maluf, 2018, pág. 482).

É nesse período que esse público começa o processo de aprendizado de leitura e, para que esse processo seja eficiente, existem algumas habilidades que precisam ser bem desenvolvidas, como a leitura, dado que nesse período de intenso desenvolvimento a arquitetura cerebral se forma de acordo com as experiências vivenciadas pelos indivíduos e o ambiente em que vivem. A leitura requer identificação de palavras escritas, o que só se concretiza se o aluno teve um bom processo de aquisição de conhecimento para que ele consiga relacionar a palavra lida com o seu sentido. Claro que ao ler não há somente esse processo em foco, mas também outras habilidades como atenção, memória operacional, conhecimento lexical (vocabulário) e da gramática da língua, raciocínio, conhecimento semântico e enciclopédico (conhecimento de mundo), e capacidades de análise e de síntese (Sargiani & Maluf, 2018, pág. 480).

Para nós, essa perspectiva de Sargiani & Maluf (2018) é importante, pois aborda a situação da educação infantil no país e como ela afeta todas as outras etapas de vida e da educação. Anteriormente, tratou-se de teorias cognitivas sobre como as crianças adquirem a linguagem, a língua e desenvolvem o conhecimento; contudo, pouco ou nada se falou sobre a sistematização e a prática dessas ações, uma vez que a realidade do Brasil, em especial, é heterogênea, ou seja, não há como pensar em uma criança prototípica, pois vive-se num país completamente desigual e repleto de realidades diferentes, sendo muitas inclusive precárias de conhecimento e qualidade de vida. Com isso, observa-se o quão essencial é falar sobre a Educação infantil, uma vez que é essa etapa que oferece a base para todo o desenvolvimento de leitura e, posteriormente, interpretação dos alunos. Embora a condição social seja favorecedora na aquisição de léxico, a escola também possui forte influência nessa

questão, como também ser o lugar no qual muitas crianças passarão mais tempo que em suas casas.

Ademais, a partir dessas pesquisas e exposições científicas, é notável que se uma criança de 6 anos não é limitada no sentido de conhecimento, como defende o artigo de Sargiani & Maluf (2018), pouco será também o público-alvo de nossa pesquisa: jovens de 11 a 15 anos. Nesse sentido, embora já tenham passado pelo processo inicial de aquisição, ainda estão em constante aprendizado lexical, por isso a importância de esse contato ser sempre respaldado de significados e contextos, para que não sejam somente palavras sem sentido a eles.

Leffa (2000) defende que “o ensino do vocabulário, tanto na aprendizagem da língua estrangeira como da língua materna, oscila entre o interno e o externo”. Essa afirmação é relevante, pois a abordagem externa é relacionada aos estudos de frequência, abordagem utilizada por nós. O estímulo para o estudo da frequência de ocorrências é o fato de que a maior parte do vocabulário é formada pelas palavras mais comuns da língua, partindo da ideia de que há graus diferentes de familiaridade do falante com as palavras, algumas são mais conhecidas que outras. Neste trabalho, não focaremos na perspectiva interna de Leffa (2000) por não ser interessante à temática.

É claro que, pensando em textos de áreas específicas, palavras pertencentes àquele léxico aparecerão com muito mais frequência, por exemplo, se pegamos uma SiP, termos como “soma”, “dobro”, “verifique” e “determine” aparecem bastante, porque fazem parte desse universo. Todavia, espera-se que substantivos e adjetivos apareçam muito mais, dado que se trata de um gênero que possui uma narrativa, um (supostamente) contexto do cotidiano que servirá de pano de fundo para o problema.

No processo de aquisição lexical, o intuito é fornecer ao aluno as diversas possibilidades de sentido de uma palavra para que ele tenha mais sentidos para um mesmo referente. Nesse contexto, Leffa (2000) propõe analisar o processo do desenvolvimento lexical por meio de três dimensões simplificadas, que foram definidas como quantidade, profundidade e produtividade.

A dimensão da quantidade é quando o desenvolvimento lexical ocorre somente quando há uma continuidade de palavras conhecidas do aprendiz. Nesse sentido, um falante no início de sua aquisição lexical terá um baixo inventário de palavras, no entanto isso irá aumentando conforme a sua experiência e sua exposição ao léxico.

A dimensão da profundidade consiste em avaliar a qualidade de um conhecimento que pode ir de superficial a profundo. Esse processo se dá, inicialmente, pelo reconhecimento, por exemplo, de uma sequência de letras como sendo ou não uma palavra. Assim, conforme desenvolve sua competência lexical, o falante “se torna capaz de estabelecer as relações paradigmáticas (sinônimos, antônimos, etc.) e sintagmáticas (que palavras podem acompanhar determinadas palavras) (LEFFA, 2000, pág. 17)”.

Por fim, a dimensão da produtividade leva em consideração o conhecimento receptivo *versus* o conhecimento produtivo de léxico. Segundo Leffa (2000), ao ler ou ouvir, temos capacidade de reconhecer muito mais palavras do que quando escrevemos ou falamos, desse modo à medida que conhecemos mais palavras, “aquelas que já eram conhecidas, tornam-se mais profundamente conhecidas e o vocabulário receptivo, com o uso constante, pode também se tornar produtivo” (LEFFA, 2000, pág. 17). No entanto, esse ganho vocabular pode ocorrer de duas formas: aprendizagem incidental e/ou intencional. Ao contrário da incidental que ocorre predominantemente pelo input – quando o falante não está preocupado em aprender novas palavras, mas adquirir significado quando ouve – a intencional, mais comum na aquisição de L2<sup>8</sup>, ocorre de modo planejado, formal, geralmente a partir de um professor.

Leffa (2000) destaca vantagens em relação à aprendizagem incidental do léxico:

(a) é contextualizada, fornecendo ao aprendiz toda a riqueza que envolve o sentido e o uso da palavra;

(b) é pedagogicamente eficaz na medida em que possibilita a ocorrência simultânea de duas atividades: compreensão do léxico e compreensão de leitura;

(c) é mais individualizada porque o léxico que está sendo adquirido vem de textos selecionados pelo próprio aluno (Huckin e Coady, 1999 *apud* Leffa, 2000).

Analisando esses processos, podemos perceber que talvez o trabalho com o ensino de vocabulário específico da matemática não esteja sendo feito de forma contextualizada, uma vez que o léxico que compõe as SiPs nem sempre é composto

---

<sup>8</sup> Apoiamo-nos em teorias da L2 em virtude de o processo de letramento matemático ser semelhante ao processo de aquisição e aprendizado de uma segunda língua.

somente de palavras frequentes aos alunos. Leffa (2000) defende que quanto mais experiências vividas pelo indivíduo envolvendo a palavra, mais profundo será o processamento. Essas experiências podem ser das mais variadas: repetição, escrita e reescrita, tradução, uso do contexto, paráfrase, etc. Isso quer dizer que um termo ouvido ou lido uma vez pode ser facilmente esquecido, ao contrário das palavras que são contextualizadas e das que mais aparecem, entrando numa rede lexical mais ampla e permanecendo na memória de longa duração (Schmitt e Schmitt, 1995; Craik e Lockhart, 1972; Craik e Tulving, 1975; Lawson e Hogben, 1996 *apud* Leffa, 2000).

Quando falamos especificamente da linguagem matemática, o que pode acontecer é somente uma instrução específica atrelada ao enunciado que será trabalhado, e isso não garante a compreensão de leitura, “o aluno deve aprender as palavras novas dentro de um contexto significativo” (LEFFA, 2000). Sendo assim, é importante destacar que as disciplinas curriculares podem tornar a aprendizagem mais próxima do aluno e autoral, pois representam o contexto ideal para que o léxico seja desenvolvido, de modo, inclusive, a retirar a total responsabilidade da disciplina de Língua Portuguesa e aliar os conhecimentos linguísticos das várias áreas, uma vez que os termos ganham conotações diferentes em cada área, como é o caso do verbo no imperativo “verifique”. Enquanto no sentido comum, “verificar” significa “averiguar” e “investigar”, na matemática demanda um cálculo a ser desenvolvido ou demonstrado. É importante esse procedimento, pois, assim como defende Leffa (2000), concordamos que:

Conhecer uma palavra não é apenas estabelecer a conexão rígida entre forma e conteúdo, como se fossem dois monolitos que se encaixassem um no outro, impossíveis de serem analisados. Conhecer uma palavra é despi-la de sua embalagem, descobrir as partes que a compõem e ver como cada uma dessas partes tem repercussões lá fora, com elementos internos de outras palavras

Como dizia Saussure (1916), o ponto de vista cria o objeto, logo não há como criar o objeto se não houver ponto de vista. Partindo desse ideário, é possível estabelecer uma analogia, uma vez que ter tido contato frequente com uma palavra torna-se essencial nesse processo de busca de sentido, principalmente, ao se deparar com esses termos em contextos mais formais como a escola. Pelo fato de neste presente estudo focarmos em SiPs, as quais precisam ser compreendidas para serem resolvidas, o vocabulário é deveras essencial.

Até aqui tratamos da aquisição de modo geral, sem diferenciar se estávamos falando especificamente de fala ou escrita. Neste momento, observamos ser importante tratar do processo de aquisição de escrita, já que a fala depende de outros processos de aquisição. Ademais, normalmente, quando as crianças ingressam no processo de alfabetização, elas já dominam a fala, já a escrita, por ser um sistema simbólico criado e desenvolvido como uma tecnologia humana, requer um ensino por outros meios diferentemente do que acontece com a fala.

A escrita é um objeto simbólico, representa algo, isto é, ela representa determinado elemento do mundo. Seguindo essa lógica, a ausência de conhecimento de uma palavra faz com que o falante não consiga acessar o significado dessa representação, tornando-se somente um conjunto de letras, como os teóricos descrevem:

Segundo Byrne (1995), as crianças desde muito cedo já são falantes bem sucedidos de sua língua nativa, porém, quando se deparam com a tarefa de ler e escrever, a situação é muito diferente. Isso porque a fala ocorre de modo relativamente contínuo e a grande tarefa do aprendiz de leitura é descobrir os elementos da fala contínua que correspondem aos elementos discretos da escrita alfabética. Essa tarefa relaciona-se diretamente com a habilidade da criança em consciência fonêmica. (LORANDI et al 2011, pág. 160)

O que é interessante observar é que, durante a fala, muitas vezes nos deparamos com palavras que não conhecemos, mas aparentemente não interferem tanto na compreensão do significado quanto na escrita. Esta é um processo diferente de contato com a língua, formalizada, não adquirida de modo natural como a fala, e é exatamente por isso que há essa dificuldade, pois o aprendiz precisa passar pelo processo de aquisição e de formação de vocabulário, para que a partir dele consiga compreender os enunciados da língua.

Enfim, embora não seja uma pesquisa da área de aquisição, acreditamos ser importante compreender como o processo de aquisição se dá para, assim, podermos relacionar com o processo de interpretação e entender de que forma a escolha lexical interfere nessas etapas e como ela é apreendida pelo cérebro. Ademais, o nosso objeto é o termo e como a ausência de significação dele impacta no conhecimento linguístico e matemático. Com isso, o próximo tópico abordará de forma mais específica o processo de aquisição e a competência lexical, com o objetivo de

conseguirmos ter uma perspectiva geral de como processamos e relacionamos o léxico.

### *1.1.2 A relação entre aquisição e competência lexical*

Para a relação entre esses processos, é importante, em um primeiro momento, compreendermos de que forma a aquisição afeta ou auxilia a competência lexical e o que essas questões têm a ver com os enunciados matemáticos.

Levando em consideração que enxergamos que a matemática se utiliza de uma linguagem que lhe é específica, podemos compreender que o aluno, para estudá-la, também passa, ou precisaria passar, por um processo de aprendizado, uma vez que ela, como disciplina, ocorre em contexto formal. Essa afirmação se dá pela necessidade de ele aprender e ter acesso ao vocabulário frequente utilizado por essa área a fim de que tenha conhecimento depois para resolver os problemas.

Falamos bastante sobre vocabulário e léxico, mas até agora não distinguimos a diferença entre esses dois termos, então partiremos da discussão de Dos Santos Araújo (2020), que afirma que vocabulário se refere ao conjunto de palavras que o sujeito utiliza no seu discurso, o que faz com que seja reduzido em relação ao conjunto de palavras existentes. Por sua vez, o termo “léxico” remete ao conjunto de todas as palavras portadoras de significado numa língua. Logo, léxico é algo muito mais amplo, enquanto vocabulário mais específico ao contexto de uso, ou seja, as palavras que o sujeito usa no seu discurso pertencem ao seu vocabulário, porém o léxico inclui essas mesmas palavras, as que são usadas, mas também as que já não são usadas (DOS SANTOS ARAÚJO, 2020, pág. 38). O conhecimento lexical é importante, pois quanto mais se conhece, mais se compreende o mundo

léxico é um conjunto aberto que os indivíduos levam longos anos para adquirir, o falante não-adulto ainda está numa das etapas ascendentes de aquisição do vocabulário; a rigor, se considerarmos o léxico da língua na sua totalidade (o que constitui uma utopia), qualquer indivíduo, mesmo adulto, estará sempre aprendendo novos elementos léxicos, pois o tesouro vocabular da língua se expande continuamente (BIDERMAN, 1981 pág. 138 *apud* XATARA, 2008).

Sabendo agora da diferença entre eles, podemos partir da ideia de que um falante tem um léxico amplo, no entanto nem sempre está em contato constante com

esses termos ou até mesmo já ouviu alguma palavra, mas pelo fato de não a ter utilizado, não conhece o sentido em um contexto, assim não faz parte de seu vocabulário. Sob esse viés, é notório que, pensando no contexto escolar, e uma vez que cada disciplina requer um vocabulário específico, o objetivo do professor é apresentar e fomentar o uso de determinadas palavras que não fazem ainda parte do repertório do aluno, ou seja, estimular que o vocabulário usado pelo aluno contenha léxico não usado, desconhecido ou novo, assim aos poucos o aluno adicionará esses termos ao seu vocabulário.

Levando em consideração essa discussão ao contexto das SiPs, principalmente porque estamos trabalhando com a linguagem usada especificamente pela matemática, o aluno aprende-a como se fosse uma nova língua, dessa forma Dos Santo Araújo (2020) indica sugestões de como privilegiar esse processo de aprendizado i. mostrando a importância da relação da unidade lexical com a realidade; ii. incentivando o uso mediante o contexto da unidade lexical mais indicada, por exemplo, recorrendo às noções de sinonímia e antonímia. Com isso, o aluno saberá o sentido de uma palavra de acordo com o emprego dela em um contexto, facilitando a resolução de problemas e o entendimento do próprio componente curricular.

Além disso,

“el vocabulario posee unas características próprias, relacionadas tanto con su extensión, composición y estructura como con su almacenamiento en la memoria y su recuperación cuando se necesita” (Peris, 2013, p. 15 *apud* Dos Santo Araújo, 2020),<sup>9</sup>

Nesse sentido, uma vez conhecida a palavras, o aluno consegue retomá-la quando necessário, ação denominada “léxico mental”. Esse conjunto de léxico mental «conjunto organizado de palavras que conhecemos e usamos» (Duarte, 2011, p. 9) é o que auxiliará na busca por sentido ao ler um enunciado matemático, por exemplo.

Todavia, o que se questiona é o quanto desse vocabulário de fato foi aprendido e o quanto o aluno irá utilizá-lo em sala. É claro que não há como ter incorporado tudo o que se foi colocado em contato, mesmo adultos que já conheceram muitas palavras não vão incorporar todas elas, mas espera-se que se aprenda pelo menos os termos que integrarão com frequência os livros didáticos de matemática, mais

---

<sup>9</sup> Tradução do excerto: “el vocabulario posee unas características próprias, relacionadas tanto con su extensión, composición y estructura como con su almacenamiento en la memoria y su recuperación cuando se necesita” (Peris, 2013, p. 15 *apud* Dos Santo Araújo, 2020) para o português “o vocabulário possui características próprias, relacionadas tanto com sua extensão, composição e estrutura quanto com seu armazenamento na memória e sua recuperação quando é necessário.”

especificamente as situações-problema. Essa nossa preocupação foi exposta por Xatara et al. (2008), ao dizer que “resultado do convívio com professores das diversas disciplinas, todos são unânimes em dizer que os alunos do Ensino Fundamental leem os textos essenciais de suas matérias com dificuldade de compreensão, em diferentes graus”.

Esse fato é recorrente em discussões sobre dificuldades em interpretação de textos e enunciados, o que também foi comprovado por pesquisa anterior (Agostini, 2019), em que se observaram alguns termos que compõem o universo matemático, como ‘prestação’, ‘verifique’, ‘identifique’, que se tornaram obstáculos para a compreensão completa de determinado enunciado, pelo simples fato de não fazerem parte do universo etário do estudante.

Diante disso, o que se observa é que há, em primeiro lugar, um senso comum de que é somente os professores de português que deveriam trabalhar léxico em sala de aula, e isso é um equívoco, visto que a aquisição acontece em todos os momentos da vida escolar e social do aluno, inclusive em aulas de matemática; em segundo lugar, há pouca prioridade para com o incentivo e valorização do léxico em sala de aula, fala-se muito da falta de leitura dos alunos, mas, aparentemente, não há estímulo a um novo conhecimento lexical por parte de outras áreas senão português. O que precisa ser compreendido por todas as áreas é que o léxico pertence a todas elas e não especificamente à área de português, em outras palavras, todas as disciplinas escolares terão palavras cujo significado será específico.

Diante disso, é importante o incentivo à leitura e ao aprendizado dos termos de cada área de estudo, como em matemática, para que os alunos ampliem o vocabulário e tenham uma gama lexical ampla para que, a longo prazo, conheçam mais palavras.

### *1.1.3 O léxico: entre o letramento e a literacia*

É muito comum dentro da área da educação ouvirmos muito sobre alfabetização – capacidade de ler e escrever – e o processo que a envolve, porém, há alguns anos, com o avanço das discussões e do entendimento sobre essa questão, surgiram conceitos designados por outra terminologia. Esses outros conceitos não excluem a alfabetização, mas invocam outros fenômenos envolvidos com a leitura e a escrita para além da decodificação. Nesse sentido, não há como falar sobre

resolução de problemas matemáticos sem nos debruçarmos sobre dois conceitos distintos, contudo complementares: *literacia* e *letramento*.

Correntemente utilizado em Portugal, França e Canadá, entre outros, o conceito de literacia pode ser entendido como competências linguísticas e habilidades cognitivas para a leitura e a escrita desenvolvidos ao longo da vida. No início de seu uso, literacia era sinônimo de alfabetização. Todavia, houve uma mudança no conceito que passou a englobar mais variáveis de habilidades, incluindo a capacidade de utilizar a leitura e a escrita de forma eficaz e crítica em diferentes esferas da vida, como na educação, no trabalho, e na participação cidadã. Porém, esse conceito não tem tanta popularidade no Brasil, esclarece Moraes (2020), que defende que “Para todos os seres humanos, a aquisição da literacia, da condição de letrado, deve ser não um direito vazio, mas uma realidade” (MORAIS, 2020, pág.5). Ou seja, um indivíduo letrado não irá somente ler o jornal, mas conseguirá compreender o que está sendo comunicado, reconhecerá os tipos de texto utilizados, além de que conseguirá discutir sobre o assunto com outras pessoas.

Enquanto *literacia* aborda competências e habilidades cognitivas para leitura e escrita, o *letramento* se volta para práticas de uso social da leitura e da escrita, palavras, como menciona MORAIS (2020, p.3):

“Letramento é entendido como a “prática social” da leitura e da escrita, adquirida através do contato com a escrita e em contextos que põem em evidência questões e relações sociais, sem que, no entanto, se especifiquem aspectos como os processos cognitivos que supostamente conduzem a saber ler e escrever, se realmente se aprende a ler e escrever de maneira eficiente, quais e como são precisamente as questões e relações sociais subentendidas, que representações se desenvolvem dessas questões e relações, e em que medida estas representações ajudam a melhorar a inserção social dos indivíduos e as relações entre eles”.

Assim como também defende Magda Soares (2004, p.7) que promove uma discussão sobre o progresso que

“[...] após alguns anos de aprendizagem escolar, o indivíduo terá não só aprendido a ler e escrever, mas também a fazer uso da leitura e da escrita, verifica-se uma progressiva, embora cautelosa, extensão do conceito de alfabetização em direção ao conceito de letramento: do saber ler e escrever em direção ao ser capaz de fazer uso da leitura e da escrita”.

Partindo dos significados do que significam *literacia* e *letramento* nos contextos já delimitados, é possível que o Brasil possa estar passando por uma carência de

letrados, uma vez que muitos alunos, também devido ao nível escolar ou financeiro para além de outros fatores, estão com dificuldade de ler um enunciado e não somente lê-lo decodificando-o, mas compreendê-lo, assimilar as partes e, no caso de SiPs, resolvê-las.

Isso se mostra um problema, pois afeta diretamente o âmbito escolar e também situações cotidianas que vão desde não interpretarem um exercício até não interpretarem um *post*, por exemplo, nas redes sociais.

A construção do léxico mental é *literacia*, enquanto a aplicação no gênero é *letramento*. O uso desse léxico para realizar problemas matemáticos envolve *literacia* *letramento* e se um desses processos não está sendo desenvolvido como deveria, temos mais um problema que refletirá na dificuldade interpretativa do aluno.

Diante disso, concordamos com Morais (2020) ao dizer que o uso do termo *letramento* em variadas áreas do conhecimento tem causado um esvaziamento de sentido no que diz respeito aos aspectos cognitivos mais estritamente. Como o termo *letramento* está indiscutivelmente mais associado a fatores socioculturais e discursivos, Morais (2020) propõe que se passe a usar o termo *literacia* mais precisamente para nos referirmos à dimensão cognitiva, aos processos de conhecimento, por exemplo, ao se estudar o que se passa no cérebro ao se alfabetizar. O fato é que documentos mais recentes da Política Nacional de Alfabetização (PNA 2019), deixaram de usar *letramento* para usar *literacia*, por isso aqui vimos como importante retomarmos e diferenciarmos esses termos para deixarmos claro que, embora complementares, devem ser discutidos como processos diferentes.

#### 1.1.4 “Ler é em parte sobre palavras”<sup>10</sup>

“Ler é em parte sobre palavras. Ou, para começar o argumento de forma mais enfática, é principalmente sobre palavras.” É com essa afirmação que Perfetti e Hart (2002, pág. 1) iniciam o artigo no qual discutem a “Hipótese de Qualidade Lexical”. A partir disso, eles argumentam que a habilidade em compreensão de leitura depende quase que inteiramente do conhecimento de palavras.

---

<sup>10</sup> Perfetti e Hart (2002, pág. 1)

Perfetti e Hart (2002) defendem que o léxico é visto como o elemento crucial para o nível de compreensão do texto. O processamento lexical abrange a forma das palavras (ortografia), os sons (fonologia) e seus significados (semântica). A habilidade de recuperar rapidamente e com precisão as palavras durante a leitura permite ao leitor integrar os significados dessas palavras, construindo assim o entendimento do texto. Isso significa que a compreensão do texto é sustentada pelos conhecimentos que possibilitam ao leitor ler e interpretar as palavras, incluindo a precisão e a abrangência com que as representações ortográficas, fonológicas e semânticas estão armazenadas na memória de curto e longo prazo. Nesse sentido, o vocabulário funciona como uma ferramenta semântica cujo objetivo é auxiliar na leitura das palavras e, conseqüentemente, na compreensão do texto. Logo, a habilidade de compreensão depende da habilidade de leitura de palavras.

Isso quer dizer que, quanto mais formos expostos a textos, mais conhecimento lexical adquirimos; é essa experiência de leitura que aumentará a qualidade de representações lexicais. Em outras palavras, a qualidade lexical possui relação direta com a frequência das palavras para determinado leitor, ou seja, uma palavra de alta frequência para um indivíduo que já a encontrou inúmeras vezes será acessada facilmente, pois esses encontros contribuirão para especificar os constituintes lexicais. Entretanto, ao contrário disso, as palavras de baixa frequência gerarão confusão e interromperão o processo de leitura.

Tais afirmações se dão a partir de estudos de Perfetti e Hart (2001; 2002) os quais identificaram efeitos de frequência em pares de palavras homófonas, que têm a mesma pronúncia, mas significados e grafias diferentes. Indivíduos considerados leitores habilidosos – frequente contato e boa compreensão de texto – demonstraram confusão apenas ao ler a palavra de menor frequência dentro do par homófono, enquanto leitores menos habilidosos foram confundidos por ambas as palavras. Essas confusões relacionadas à forma das palavras, seja fonológica, seja ortográfica são influenciadas pela frequência com que essas palavras aparecem (Perfetti, 2007). Essa lógica pode ser comprovada pelas palavras "gate" (portão) e "gait" (andar), as quais têm a mesma pronúncia, mas a segunda palavra é menos frequente e pode causar confusão em leitores com menos experiência. Além da frequência de ocorrência das palavras, a familiaridade ou frequência relativa, que se refere à experiência de um leitor com uma palavra específica, também deve ser considerada. No entanto, essa

experiência varia entre indivíduos e "depende da quantidade de leitura realizada" (Perfetti & Hart, 2001, p. 7).

De fato, essa linha defendida por Perfetti e Hart (2001; 2002; 2007) faz muito sentido ao olhar para o nosso trabalho, uma vez que quando estamos lendo e o texto é composto de palavras conhecidas – frequentes a nós – a leitura flui, por outro lado, quando há palavras desconhecidas – pouco frequentes – a leitura é interrompida, isto é, palavras de baixa frequência prejudicam o processo de leitura e, automaticamente, o de compreensão. Não seria diferente com os alunos. Ao voltarmos aos dados do PISA, um dos pontos destacados negativamente é que nossas crianças possuem dificuldade de leitura e interpretação. Uma das justificativas, para além da qualidade educacional e da condição socioeconômica, pode ser o pouco contato com a leitura, não por baixa exposição, pois os alunos têm contato diário com textos, mas o pouco interesse e, talvez, a pouca prática de ler realmente um texto, inclusive de modo independente. Porém, além disso, há também outra problemática: palavras de baixa frequência que compõem os enunciados com os quais eles têm contato.

Portanto, o léxico é extremamente importante para a compreensão de um texto, logo não há como esperar que uma criança de 10 a 11 anos, por exemplo, não se sinta confusa ao ler um enunciado no livro didático de matemática no qual há um termo “escalímetro” ou “hodômetro”. A qualidade lexical é formada pela frequência de contato com a leitura, sim, mas também pela compreensão de textos que só acontecerá se ele for composto de palavras frequentes.

Como já diziam Perfetti e Hart (2002, pág. 1), “Ler é em parte sobre palavras. [...] é principalmente sobre palavras”, é seguindo nessa linha que iniciaremos no próximo tópico um aprofundamento sobre sentidos, uma vez que se ler é sobre palavras e palavras são repletas de sentidos e representações, precisamos compreender como os acessamos, a começar com Frege (1978).

## **1.2 SENTIDO E REFERÊNCIA: MEANDROS DO LÉXICO E DA COMPOSICIONALIDADE**

A clássica compreensão da relação entre sentido e referência é fundamental para esta pesquisa, por isso é essencial retomarmos alguns conceitos para interligá-los a nossos objetivos.

Começando por G. Frege (1848-1925), é preciso observar que, em seus artigos 1891 e 1892 (1978), intitulados *Função e conceito* e *Sobre Sentido e Referência*, abordou o conceito de igualdade ou equivalência entre termos, o que quer dizer que duas expressões podem ter o mesmo referente, por exemplo, é possível referir-se ao numeral 3 por meio de distintas expressões como  $5-2$ ,  $2+1$  ou até mesmo pelo próprio '3' como algarismo. Por isso, a importância de se pensar que a igualdade entre objetos e sentidos, com seria o caso de  $a = a$  acarretaria uma tautologia e portanto nenhum conhecimento novo ou atitude avaliativa seriam necessárias. Enquanto que em  $a = b$ , temos um valor cognitivo diferente por colocar não a identidade entre dois sinais, mas dizer que os dois têm o mesmo valor, embora sejam distintos em si mesmos. O ganho cognitivo é perceber que por caminhos distintos (sentidos) podemos chegar ao mesmo objeto (referente). Para o lógico alemão, sentido é a forma ou aspecto do que ou de que se fala, o modo como é apresentado na frase, ou seja, não é de fato o objeto ou pessoa, mas sim somente sua representação. Seguindo essa lógica, podemos ter mais de uma expressão que indicará o mesmo referente:

- a) Mônica tem um coelhinho de pelúcia chamado Sansão.
- b) A filha mais velha de Maurício de Souza tem um coelhinho de pelúcia chamado Sansão.
- c) A melhor amiga de Magali tem um coelhinho de pelúcia chamado Sansão.

Note-se nos exemplos acima que há, além do próprio nome "Mônica", outras duas formas de se referir à personagem dos quadrinhos de Maurício de Souza, como são ilustrados nos exemplos (b) e (c). Seguindo o pensamento de Frege, as frases acima apresentam três sentidos para se referir ao mesmo referente. Embora existam várias formas de se referir à Mônica, só é possível compreender o enunciado aquele que tem conhecimento sobre Mônica e sua história, ou seja, que saiba que é a filha mais velha de Maurício de Souza e que é a melhor amiga de Magali; quem não tem ciência disso, não conseguirá compreender a referência. É esse cenário que estamos problematizando aqui: palavras incomuns a determinada faixa etária que aparecem em enunciados matemáticos, como ilustrado abaixo pela SiP a qual consta no volume dedicado a alunos de 7º ano, em torno de 12 anos:

5. Com um **escalímetro**, a planta de uma casa foi desenhada na escala 125. Nessas condições, responda as questões.
- Qual é o comprimento real, em metros, de uma sala que, nessa planta, foi representada por um segmento de 5,2 cm?
  - Os quartos dessa casa medem 3m por 4m. Quais são as medidas dos quartos nessa planta?
  - Na planta, o terreno mede 6,4cm por 28cm. Quais são as medidas reais desse terreno, em metro? (BIANCHINI, 2019)

É um enunciado que parece ser bem comum ao cenário matemático, pois aborda contextos reais de geometria, e parece-nos, em um primeiro olhar, que a SiP é clara. Todavia, há o termo "escalímetro" que pode soar estranho sem um glossário ou contextualizado, porque não deve ser um termo comum a uma criança dessa faixa etária. O que seria um escalímetro para uma criança de 12 anos? Assim, devido à dificuldade saber o sentido, existe uma dificuldade de compreender a referência, pois quanto mais sentido sobre o mesmo objeto, mais meios se têm para ganhos cognitivos, como defende Oliveira (2001):

[...] o sentido só nos permite conhecer algo, se a ele corresponder uma referência. Em outros termos, o sentido permite alcançarmos um objeto no mundo, mas é o objeto no mundo que nos permite formular um juízo de valor, isto é, que nos permite avaliar se o que dizemos é falso ou verdadeiro. A verdade não está, pois, na linguagem, mas nos fatos do mundo. A linguagem é apenas um instrumento que nos permite alcançar aquilo que há, a verdade ou a falsidade. (OLIVEIRA, 2001, p. 22)

Seguindo essa lógica, um dos pontos que acreditamos ser importante é o entendimento do enunciado em língua natural para que seja feita a tradução para a linguagem matemática, como destaca Lorensatti (2009, p. 91):

A linguagem matemática pode ser definida como um sistema simbólico, com símbolos próprios que se relacionam segundo determinadas regras. Esse conjunto de símbolos e regras deve ser entendido pela comunidade que o utiliza. A apropriação desse conhecimento é indissociável do processo de construção do conhecimento matemático. Está compreendido, na linguagem matemática, um processo de 'tradução' da linguagem natural para uma linguagem formalizada, específica dessa disciplina, segundo Granell (2003). Os enunciados emitidos em língua natural passam a ser escritos para o equivalente em símbolos matemáticos. Essa tradução 'é o que permite converter os conceitos matemáticos em objetos mais facilmente manipuláveis e calculáveis'.

Se o aluno não entender o enunciado, ele terá dificuldade de resolver o que é pedido a ele, dado que, embora o léxico seja o mesmo, os sentidos mudam de acordo com o contexto da situação-problema. Nesse caso, 'escalímetro' não diz respeito apenas a um determinado objeto ou conceito, mas ao cerne da resolução da SiP, pois se trata da escala relativa, ou seja, quais os padrões proporcionais das medidas representadas por uma planta. Imagine que em lugar de 'escalímetro' tivéssemos a palavra 'asekale' (escala em iorubá). Como se trata de uma língua estrangeira, muito embora presente na cultura afro-brasileira, não é um termo transparente para falantes do português brasileiro. Se 'escalímetro' e 'asekale' podem ter a mesma falta de referência, para um adulto, imagine para uma criança!

Desse modo, o estudante precisa de muita familiaridade com a terminologia trabalhada em sala e uma ampla referência de sentidos. A interpretação é diretamente conexa ao referencial linguístico e matemático para que auxilie na resolução. Sendo assim, quanto mais sentidos e conhecimentos sobre o mundo, mais relações podem ser feitas, o que justifica afirmarmos que a lógica fregeana e a sua abordagem sobre sentido e referência sejam bastante úteis para se pensar em ensino e resolução de SiP que promova movimentos interpretativos de modo a associar mais sentidos às referências.

Todavia, para além das palavras pertencentes ao cotidiano que integram os enunciados, por exemplo, "casa", "piscina", "terreno", "escola"; há também aquelas que fazem parte especificamente do campo da matemática. Há uma barreira considerável que se ergue entre o sentido e a referência no que tange a termos como: "divida", "subtraia", "adicione", "some", "multiplique", ou até mesmo formas nominais que representem as operações, como é o caso de "dobro", "triplo" e "quádruplo". Essa dificuldade foi explorada por Agostini (2019) que averiguou o fato de os alunos não entenderem que termos "emprestar um lápis" poderia ser uma subtração de um conjunto de lápis; ou que "cortar o pedaço da pizza" poderia ter o mesmo sentido de dividir ou fracionar uma unidade, ou até mesmo que o dobro equivale a multiplicação de algum número por vezes 2. A esse respeito Oliveira (2001, p. 94) afirma:

Uma mesma referência pode, pois, ser alcançada por diferentes sentidos. Quando descobrimos um novo sentido, aprendemos algo novo sobre um objeto no mundo, porque sabemos como chegar a ele por um novo caminho. Sentido é o caminho que nos leva à referência

Já sabemos que realmente os termos importam para o processo interpretativo, então o que neste trabalho buscamos averiguar é se essa escolha vocabular que consta nas SiPs está de acordo com a faixa etária, porque não tem como mantermos a ideia de que o aluno que não interpreta devidamente se o campo lexical for incomum àqueles que têm contato frequente com esses textos.

A partir disso, no próximo capítulo, discutiremos a definição de situação-problema, seu uso no ambiente escolar pela matemática e de que forma ela pode ser tornar problemática ao estudante.

## 2 A SITUAÇÃO-PROBLEMA

É fundamental que compreendamos de modo detalhado a que se presta o gênero explorado por nós, e este é o objetivo do presente capítulo.

### 2.1 A TAL DA SITUAÇÃO-PROBLEMA

Todos que passaram pela escolarização se depararam com desafios na aula de matemática textualizados em forma de "situação-problema". Esse gênero textual presente e predominante não é apenas um exercício, mas é também um instrumento de avaliação, que aparece também em concursos e situações de verificação de aprendizagem. Abaixo, apresentamos um outro exemplo de SP:

6. Para a festa de aniversário de seu filho, Cauê estimou que 60 copos de refrigerante seriam suficientes. Ele sabe que em cada copo cabe  $\frac{1}{5}$  do refrigerante de um litro. Quantos litros Cauê deve comprar? (BIANCHINI, 2019)

Geralmente, a SiP tem como contextualização algo comum ao dia a dia, como uma festa de aniversário, é conhecida por ser desafiadora ao aluno. No entanto, sua estrutura não pode ser resumida somente por ter essa parte desafiadora, pois, conforme notou Agostini (2019), esse tipo de gênero textual possui traços linguísticos e estéticos particulares, pois contém palavras relacionadas às ações do dia a dia e nomes tidos como comuns. As SiPs são usualmente formadas por 3 partes, que são: contexto + situação + comando, assim como descreve mais detalhadamente:

Em relação ao plano composicional, podemos dizer que a situação-problema [...] é composta por um conteúdo temático, geralmente com situações do cotidiano, como ações comerciais que envolvem dinheiro, momentos escolares, compras, circunstâncias comuns em aniversários, entre outros momentos. Seguido do contexto, temos as proposições que serão importantes para a solução do problema. Elas são as situações relacionadas ao contexto descrito. [...] O comando pode ser dado por meio de uma sentença interrogativa iniciada por um WH ou uma sentença no imperativo, a qual indica, pelo menos, a ação que deve ser desenvolvida. (AGOSTINI, 2019, p.50)

A SiP é composta por um conteúdo temático, geralmente com situações do cotidiano, como ações comerciais que envolvem dinheiro, momentos escolares, compras de objetos e imóveis, circunstâncias comuns como em aniversários, entre outros momentos. Seguido do contexto, temos as proposições que são importantes para a solução do problema, as quais são as situações relacionadas ao contexto descrito. Como forma de exemplificação de tal lógica, seguem os enunciados retirados do livro para alunos do 6º ano da coleção Bianchini (2019)<sup>11</sup>:

7. A capacidade do tanque do meu carro é de 50 litros. O combustível que uso é composto de  $\frac{4}{5}$  de gasolina e  $\frac{1}{5}$  de álcool. Vou abastecer o carro com 30 litros de combustível. Quantos litros de gasolina colocarei?
  
8. Tenho um terreno retangular cujo comprimento é igual ao triplo da largura. Pensando em colocar um muro ao redor desse terreno, consultei um pedreiro para saber quantos tijolos deveria comprar. Ele me disse que, para cada metro de muro no comprimento, seriam necessários 130 tijolos. Então, comprei 10.000 tijolos. Sabendo que a largura desse terreno é de 10,8 m, sobraram ou faltaram tijolos? Quantos?

Podemos observar que os exemplos seguem exatamente a estrutura da SiP descrita por Agostini (2019), uma vez que são iniciados por (1) contexto - história inicial que contextualiza em que cenário irá ocorrer o problema -, (2) situação - o que ocorre no problema que fará com que exista uma "situação-problema" - e (3) comando - parte final do enunciado, composta por uma estrutura de verbo no imperativo ou de pergunta em que há algum direcionamento que o aluno deverá solucionar.

Para ilustrar, em (8) o contexto (parte 1) é a existência de um terreno com medidas específicas; a situação (parte 2) é a construção de um muro nesse terreno; por fim, há o comando (parte 3) que questiona o aluno se a quantidade de tijolos comprados é suficiente ou não para a construção do muro.

Ainda sobre sua composição, as SiPs, geralmente, aparecem com verbos imperativos que indicam o que fazer, como *calcule*, *indique*, *descubra*, entre outros termos recorrentes. Ao nos depararmos com problemas, já sabemos que não se trata

---

<sup>11</sup> Coleção base para a formação do nosso corpus.

de um simples enunciado matemático, mas sim um texto particular, devido às características apresentadas acima. Assim, ainda que existam vários tipos de textos, estruturas, sentido, objetivo e público-alvo, o que há em comum é a tentativa de simular contextos reais nos quais a matemática é utilizada, até para mostrar aos estudantes que é uma disciplina que não ficará somente na sala de aula; pelo contrário, utilizar-se-ão dela nos mais diversos contextos, desde compras parceladas até uma organização de um evento.

Então, a SiP é um gênero que tem como objetivo emular condições reais do cotidiano, com o fito de aproximar o estudante à realidade. Essas situações cotidianas sempre são retratadas com o surgimento de problemas, de algum aspecto que deve ser pensado para ser solucionado. Acreditamos que o desenvolvimento e a presença dela em aulas e livros de matemática seja, para além de desafiar o aluno, ambientá-lo a cenários não tão abstratos e distantes da vivência comum e real, por isso é tão importante se ter cuidado ao compô-la, pois precisa ser bem escrita a fim de que o aluno consiga realmente aproveitá-la da melhor forma possível: interpretando-a e resolvendo-a.

## **2.2 QUANDO A SiP SE TORNA UM "PROBLEMA"?**

Os enunciados de SiPs, para além das características já citadas, têm em comum a presença de números que se destacam dentre as palavras. O fato de haver esses números desperta no aluno a sensação de que algo deve ser feito. Com base nisso, Braguil *et al.* (2017) apontam que o aluno não interpreta o enunciado, pois só lê buscando por dados relevantes que auxiliarão na resolução do problema.

Contudo, essa leitura seletiva não garante que os dados coletados sejam estabelecidos e executados de forma coerente, já que não há o processo interpretativo, isto é, a relação semântica desses números com as palavras que constam também no enunciado faz com que não ocorra uma leitura aprofundada e crítica, mas superficial, dado que evita uma conexão com suas necessidades. Tal ação potencializa a falta de indagações e leitura crítica, que entendemos como sendo aquela leitura na qual se entende o sentido das palavras que compõem a SiP e principalmente o que se pede nela. Por exemplo, existem SiPs nas quais há o verbo "emprestou" que na matemática podemos ler como "subtrair", porém que naquele

determinado momento está ali somente para contextualizar algo e não que o aluno precise subtrair os números que constam no enunciado. Vejamos a seguir:

9. Paula tem um estojo repleto de canetas coloridas. Ela **emprestou** 5 para Laura, 3 para Lucas e 8 para Luana. Quantas canetas coloridas **tinham** no estojo de Paula antes de emprestá-las?

Note que, embora exista o verbo "emprestar", o ponto importante dessa SiP é indicar a quantidade total, então a operação matemática pedida é adição e não subtração. Porém, muitos alunos, consoante a defesa de Agostini (2019), confundem-se porque querem "fazer algo" com os números, mas sem compreender a relação deles com as palavras.

Em suma, a SiP é um gênero utilizado na escola com o objetivo de aproximar o contexto estudantil ao da disciplina de Matemática. Nessa esteira, utiliza-se de situações corriqueiras e das mais variadas ações do dia a dia a fim de demonstrar ao aluno que a matemática será utilizada para além da sala de aula. No entanto, ao não ser escrita com um vocabulário adequado ao público-alvo, de um texto educativo e simples, transforma-se em um obstáculo dificultoso, pois o estudante não a compreende, logo não a soluciona ou a faz de modo qualquer. Diante disso, no próximo capítulo, descreveremos quais foram os materiais e os métodos que utilizamos neste trabalho para compreender e discutir quais são esses termos problemáticos e de que forma isso gera efeitos nos aprendizes.

### **3 MATERIAIS E MÉTODOS**

Nesta seção, discorreremos sobre as etapas do processo investigativo. Como a metodologia é a parte fundamental para que uma investigação seja adequada e obtenha êxito, serão pontuadas todas as etapas cumpridas nesta pesquisa.

Este trabalho objetiva-se a investigar a influência dos termos que compõem enunciados de situações-problema em matemática durante o processo interpretativo do estudante. Para isso, focaremos na análise da frequência desses termos a fim de averiguar se são usos comuns ou não para as crianças – público-alvo dos livros utilizamos em nossa análise.

Nossa hipótese é que termos comuns nos livros didáticos de matemática, em especial com os que trabalharemos, não são frequentes às crianças, por isso os efeitos na interpretação dos alunos, dado que, ao se depararem com termos com os quais não possuem contato frequente, não os compreendem, conseqüentemente não conseguindo resolver a SiP.

#### **3.1 OBJETOS DE ESTUDO: O CORPUS E OS TESTES**

Partido da hipótese de que o desconhecimento vocabular possa ter efeitos sobre a interpretação de SiP, construímos um corpus somente com textos desse gênero para analisar a escolha e o uso das palavras. Se considerarmos que o acesso de boa parte dos estudantes ao universo da leitura se concentra no livro didático<sup>12</sup>, então podemos inferir que tanto o reconhecimento de palavras quanto a aquisição delas pelos estudantes se dá no LD.

Para isso, elegemos uma coleção de livros de matemática utilizada na rede pública de São Carlos, que faz parte do Programa Nacional do Livro Didático de 2020 a 2023, denominada "Matemática Bianchini", do ano de 2018. Tal coleção é composta por 4 livros do 6º ao 9º anos, correspondentes aos anos do Ensino Fundamental II, para serem utilizados no componente curricular de matemática. Nesse sentido, enfatizamos que o foco é no material linguístico, sendo assim, não é interessante o número exato de escolas que utilizam o material eleito, mas só pelo fato de ele ser

---

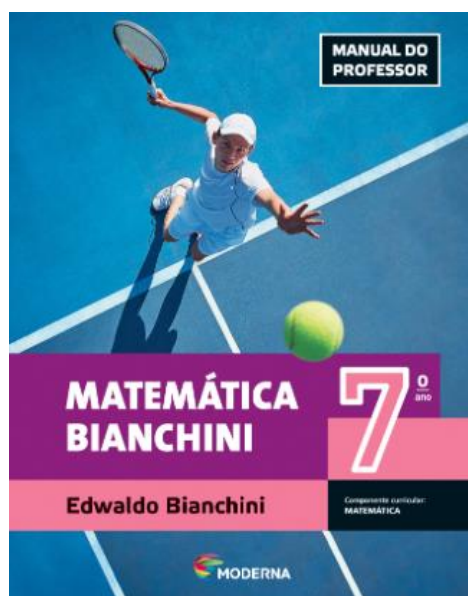
<sup>12</sup> A partir daqui, utilizaremos LD para nos referirmos a "livro didático".

uma das opções de escolha das escolas e a escola parceira o utilizar, já evidencia que existem, no mínimo, 618<sup>13</sup> alunos do Ensino Fundamental II em contato com a obra - quantidade de discentes da Escola Estadual Doutor Álvaro Guião, a qual nos forneceu o material para análise.

### 3.1.1 A coleção

"Matemática Bianchini" foi produzida por Edwaldo Bianchini – Licenciado em Ciências pela Universidade da Associação de Ensino de Ribeirão Preto, com habilitação em Matemática pela Faculdade de Filosofia, Ciências e Letras do Sagrado Coração de Jesus, Bauru (SP). Professor de Matemática da rede pública de ensino do estado de São Paulo, no Ensino Fundamental e Médio, por 25 anos, segundo informações da própria editora do livro, a Moderna. A coleção aborda os conceitos matemáticos por meio de recursos como textos, imagens e fatos históricos, promovendo uma estética atraente e colorida ao aluno.

A coleção é composta por 4 livros, cada um destinado a um ano escolar, sendo eles do 6º ao 9º ano, respectivamente<sup>14</sup>:



<sup>13</sup> Dado fornecido pelo site cujo foco é educação QEdU o qual se denomina como um projeto inédito idealizado pela Meritt - na pessoa dos Srs. Ricardo Fritsche e Alexandre Oliveira - e pela Fundação Lemann em 2012. Esses dados, embora publicados pelo site, foram retirados do Censo Escolar, feito pelo INEP, de 2020. Disponível em: <https://novo.qedu.org.br/escola/35024512-alvaro-guiao-doutor>. Acessado em 30.mai.2022.

<sup>14</sup> Imagens retiradas do site da editora Moderna. Disponível em: <https://pnld.moderna.com.br/matematica/matematica-bianchini/>. Acessado em 30.mai.2022.

Figura 1 Coleção Bianchini para 6º anos.



Figura 2 Coleção Bianchini para 7º anos.



Essa coleção tem como proposta pedagógica a diversidade na abordagem de temas e inter-relações entre as unidades, além de ser adaptada às diretrizes da BNCC. Para o projeto, essas informações são importantes pelo fato de trabalharmos com o vocabulário do aluno, ou seja, uma vez adquirido, facilitará a interpretação dos enunciados das próximas coleções, para isso é de fundamental importância que os conteúdos abordados sejam familiares ao contexto do estudante e de acordo com a faixa etária, pois o não reconhecimento de um termo impede as crianças de encontrarem sentido no enunciado. Caso isso ocorra, o processo interpretativo será comprometido, uma vez que o discente depende do sentido das palavras para compreender o que deverá ser calculado a

Figura 4 Coleção Bianchini para 9º anos.

Figura 3 Coleção Bianchini para 8º anos.

partir dos enunciados. Além disso, a

diversidade de temas e situações se tornam importantes para levar em consideração diversos contextos, ou seja, quanto mais exposição do estudante a diferentes tipos de vocábulos, possivelmente maior será a oportunidade de aprendizagem vocabular, visto que, a partir dessa exposição, terá contato com mais palavras e contextos.

Em relação aos contextos abordados, estes foram os mais recorrentes: horas, tamanhos de casas, terrenos e piscinas, comprimento e áreas residenciais e de objetos do dia a dia, festas de aniversário, comemorações de feriados nacionais, esportes, divisão de alimentos, situações cotidianas (compras, vendas, compartilhamento de objetos e comidas, empréstimos e ações de rotina), construção de ambientes, investimentos, recebimento de dinheiro, pagamento salarial e situações no trânsito. Levando em consideração o gênero do enunciado (situação-problema), cujo foco é abordar contextos comuns, estão de acordo com o previsto, pelo fato de serem vivências esperadas por jovens de 10 a 14 anos, público pertencente ao Ensino Fundamental II em contato com a coleção.

Nos volumes, um recurso muito utilizado é a elaboração de cenários comuns aos alunos, ideário refletido, por exemplo, nas capas que ilustram esportes, a fim de fazer com que o público leitor se interesse e se identifique com o livro utilizado. Essa abordagem é importante, pois o gênero analisado neste trabalho utiliza-se de contextos comuns aos alunos como pano de fundo para os desafios propostos por ele, dessa forma, o livro abordar esse recurso de levar o mundo do jovem às aulas de matemática torna-se um elemento positivo a ser destacado.

Para selecionarmos as SiPs, levamos em consideração a pesquisa anterior (Agostini, 2019), na qual defendemos que situações-problema configuram-se como um gênero, visto que seguem características linguísticas e estruturais como: contexto, situação e comando e são utilizadas como função social, ou seja, para ser definido como tal, precisa conter em seu plano composicional conteúdo temático, situação problema relacionada ao conteúdo (algum evento problemático que demanda alguma ação) e o comando composto ou por imperativo ou por uma questão a qual desafiará o aluno.

As situações-problema aparecem em 5 seções diferentes na coleção: "exercícios propostos", "pense mais um pouco", "trabalhando a informação", "exercícios complementares" e "agora é com você!", sempre seguidas da parte teórica para a aplicação. Portanto, foi fácil encontrá-la durante a leitura do material, dado que apareceu com frequência em praticamente todos os capítulos de todos os volumes.

### 3.2 O CORPUS

Para a construção do corpus, primeiramente, foram selecionados enunciados que apresentavam características que os classificavam como situação-problema<sup>15</sup>. Em relação à quantidade e aos anos, ilustramos na tabela a seguir:

*Tabela 1 Quantidade de enunciados*

<b>Série</b>	<b>6ºano</b>	<b>7ºano</b>	<b>8ºano</b>	<b>9ºano</b>
<b>Quantidade de enunciado por série</b>	120	86	94	90
<b>Total de situações-problema</b>	390			

Após a coleta dos enunciados, todos os exercícios foram digitalizados, por meio do escaneamento digital e catalogados por número de exercício, capítulo, volume e assunto, a fim de facilitar a busca futura desses enunciados caso fosse necessária. Depois da digitalização, os enunciados foram convertidos, por meio do software OCR (Optical Character Recognition), pelo qual é possível converter tipos diferentes de documentos digitalizados em texto real, formato TXT<sup>16</sup>, de modo que fosse possível, a posteriori, a leitura no software de estatística lexical escolhido para o trabalho, o Léxico 5.

A ferramenta escolhida, Léxico 5, programa de análise lexical, foi eleita, dado que, a partir dela, é possível saber a quantidade de vezes que um termo é utilizado

---

<sup>15</sup> Em trabalho anterior (AGOSTINI, 2019), foi constatado que a situação-problema possui uma estrutura prototípica integrada por contexto + situação + comando. A combinação desses três blocos de informação resulta no desafio e na transposição do que foi dito em línguas naturais à linguagem matemática. Em relação ao plano composicional, pode-se dizer que a situação-problema possui determinada estrutura sem muitas modificações de um enunciado para outro. Ela é composta por um conteúdo temático, geralmente com situações do cotidiano, como ações comerciais que envolvem dinheiro, momentos escolares, compras, circunstâncias comuns em aniversários, entre outros momentos. Seguindo do contexto, temos as proposições que são importantes para a solução do problema, é o momento em que se elabora um evento problemático. Elas são as situações relacionadas ao contexto descrito. Por fim, o comando, composto por uma oração encabeçada por verbo no imperativo ou por uma pergunta encabeçada por pronomes interrogativos, como qual, quais, quanto parte a qual indicará o que o aluno deverá calcular.

<sup>16</sup> Extensão de arquivo de texto que não admite qualquer formatação.

na base de dados e quais os contextos sentenciais do termo em questão, por exemplo, quais as palavras que o antecedem e o sucedem. Assim, poderíamos com facilidade acessar as partes do problema (contexto + situação + comando) com a finalidade de averiguar o contexto no qual determinado vocábulo foi empregado, como também se a parte composicional em que a palavra se encontra influencia no sentido do termo.

Após a conversão em TXT, os dados, 390 enunciados, foram inseridos no software Léxico 5 com o objetivo de termos uma lista de frequência total de palavras que constam nas situações-problema da coleção "Matemática Bianchini". Essa lista foi composta pelas palavras, pela frequência de cada uma delas e pelo contexto em que se encontravam. No total, foram 3.756 termos, sem ainda a limpeza e a seleção do corpus. No entanto, observamos que seria necessário também gerar listas separadas de cada um dos livros para que fosse possível analisarmos a frequência vocabular em cada ano escolar, pois é um dado que pode surtir resultados interessantes ao nosso questionamento se a palavra de fato influencia na interpretação de um enunciado matemático. Essa prática fundamenta uma de nossas hipóteses: conforme o aluno avança no ano escolar, surgem termos menos usuais ou até mesmo desconhecidos a eles, por se tratar de séries mais avançadas.

A partir desse pensamento, foram geradas mais 4 listas de frequência, ilustradas abaixo:

*Tabela 2 Quantidade de palavras*

<b>Série</b>	<b>6° ano</b>	<b>7° ano</b>	<b>8° ano</b>	<b>9° ano</b>
<b>Quantidade de formas</b>	1.654	1.359	1.511	1.594
<b>Quantidade total de formas</b>	6.118			

Nessa esteira, é importante ressaltar que a diferença de valor em relação à quantidade de formas se dá, pois, quando unimos todos os enunciados no software, a frequência se dá por palavra, independente de em qual livro se encontra, por esse motivo que foi também necessário gerar a frequência separada para que tenhamos mais possibilidade de combinações do corpus.

Depois desse processo, extraímos as listas de frequência de cada um dos livros e as unimos no Excel para que fosse possível dar sequência à organização, limpeza e classificação dos dados.

Logo após a extração do Léxico 5 da lista com as palavras e suas frequências, ela foi aberta no Excel para que pudéssemos começar a limpeza do corpus, para, posteriormente, fazer a análise. Primeiramente, foi necessário limparmos todas as entradas que eram compostas por letras sem relação de sentido entre elas, pois, ao converter a imagem digitalizada dos enunciados ao texto, por meio do OCR, o software faz algumas leituras equivocadas, por exemplo, nota de rodapé, títulos e subtítulos do livro, como também palavras separadas pela translineação que ele não compreende como sendo um único termo. Além disso, algumas fontes de texto podem causar problemas, como a confusão entre “n” e “r”. Dessa forma, são gerados muitos dados incompletos e sem sentido, por isso a necessidade de uma limpeza antes do início da classificação.

Em relação à limpeza de termos específicos, como não utilizamos números em nossa análise, todos os que não estavam por extenso foram excluídos da lista, junto com palavras mal formuladas e letras que não tinham função semântica no contexto gerado pelo Léxico 5. Nesse primeiro momento, não sabíamos ainda exatamente o que encontraríamos, então optamos por não excluir nenhuma palavra, número por extenso ou letra, somente símbolos que de fato não contribuiriam com o processo analítico ou não tinham conteúdo semântico.

Destarte, o corpus final, após a organização, tem 5.052 palavras dos 4 livros analisados da coleção "Matemática Bianchini". Assim, a próxima etapa consistirá na caracterização dos dados e os filtros utilizados para uma melhor análise.

### *3.2.1 Caracterização do corpus*

Depois de excluir os dados que não seriam utilizados, tais como números cardinais, símbolos e pontuação, porque não são termos e elementos importantes para nossa análise, o corpus foi organizado em 4 colunas que apresentam 4 variáveis: 1) volume no qual a palavra se encontra, 2) o item lexical em ordem alfabética, 3) sua classificação morfológica e 4) sua frequência no corpus. Pelo fato de trabalharmos

somente com palavras de conteúdo, achamos melhor classificá-las morfológicamente para já termos de modo claro e indicado a qual classe cada termo pertence.

O corpus a ser analisado – depois de feitas as exclusões necessárias – contém 5052 tokens, dentre eles, substantivos, artigos, pronomes, adjetivos, verbos, preposições, conjunções e numerais. É importante ressaltar que dentre esses 5052 tokens, há termos repetidos, principalmente aqueles que ocorrem em letra maiúscula e minúscula, plural e singular, pois foram coletados dos 4 livros de matemática. Em um primeiro momento, achamos necessário não haver contabilização de palavras independente do livro para que, posteriormente, pudéssemos cruzar dados e ver se uma palavra é mais frequente em um nível estudantil em relação ao outro, ou até mesmo outras possibilidades de respostas ao nosso problema.

Para que fosse possível chegarmos a essas classes de palavras encontradas, caracterizamos manualmente cada um dos termos, com base nas características morfológicas que determinam a qual classe pertence cada uma das palavras. Todavia, em virtude de haver vocábulos que podem se encaixar em mais de uma classe gramatical, tornou-se imprescindível, em alguns casos, buscar o contexto de itens como "o" e "a" que podem ser artigo, pronome e, no caso de "a", até mesmo preposição. Assim, ressaltamos a importância de estar em contato com o texto o tempo todo, de modo a salientarmos a praticidade que foi ter utilizado a ferramenta Léxico 5, pois, através dela, foi possível acessarmos facilmente o contexto de um determinado problema.

### **3.3 Análise contrastiva de corpus**

Como o nosso objetivo é trabalhar com a frequência de palavras, principalmente daquelas que são pouco frequentes em corpus do português, mas aparecem nas SiPs, acreditamos na necessidade de estabelecer comparações entre nosso corpus com outros dois, um de adulto – Corpus fonotático em português (POST DA SILVEIRA et al, 2018) – e outro de criança – LexPorBr Infantil (ESTIVALET et al, 2019) – todos do português brasileiro. É importante ressaltar que ambos os corpora estão disponibilizados de modo gratuito em plataformas digitais<sup>17</sup>.

---

<sup>17</sup> LexPorBr Infantil: <http://lexicodoportugues.com/infantil>

### 3.3.1 *Corpus fonotático do português brasileiro*

Corpus Fonotático do Português Brasileiro<sup>18</sup> (adulto), com 228.766.402 tokens e 123.826 types, foi criado para averiguar a distribuição fonotática e acento das palavras, levando-se em consideração seu tamanho (POST DA SILVEIRA et al, 2018). Para a composição do corpus, foram utilizados outros corpora e algumas ferramentas disponíveis gratuitamente, tais como a Avaliação Sonora do Português Atual (ASPA), o dicionário de pronúncia de leitura automática para o português brasileiro, Aeiouadô e The Bank of Portuguese corpus.

O Corpus Fonotático do Português Brasileiro provê informações sobre o número de grafemas e fonemas em uma palavra, número de sílabas, padrão de acentuação da palavra, distribuição fonotática por posição silábica e frequência de ocorrência da palavra no corpus. As palavras incluídas variam em comprimento de 1 a 10 sílabas.

Para o nosso objetivo, esse corpus é importante pela completude e tamanho, uma vez que precisamos da maior quantidade possível de palavras e suas frequências para que nosso estudo seja completo, a fim de que a comparação entre os corpora seja coerente e a mais verídica possível.

### 3.3.2 *LexPorBr Infantil*

O LexPorBr Infantil (ESTIVALET et al., 2019) conta com 129.053.297 tokens e 874.887 types retirados de textos escritos, ouvidos e falados por crianças.

Segundo Estivalet et al. (2019), o trabalho com corpus construído somente a partir de textos escritos não é tão seguro para análise infantil, uma vez que não há interação comunicativa. Nesse sentido, trabalhos focados no público infantil têm adotado materiais mais similares à linguagem oral, como o corpus construído pelos autores. Sendo assim, o LexPorBr Infantil foi produzido a partir de legendas de filmes e séries ouvidas por crianças, textos escritos e lidos pelo público infantil, sob a justificativa de que “alguns estudos mostraram que a frequência obtida a partir de legendas prevê com mais precisão os tempos de reação para o reconhecimento de palavras” (Brysbaert and New 2009, Brysbaert et al. 2011a, Brysbaert et al. 2011b,

---

<sup>18</sup> Adulto: <https://osf.io/5u9mf/>

van Heuven et al. 2014 apud Estivalet, 2019, p.191). Com base nisso, achamos importante contrastar nosso corpus, extraído de livros didáticos de matemática, isto é, linguagem escrita, com textos também ouvidos e falados, para que tenhamos uma perspectiva do público infantil mais próxima do que de fato ele conhece.

### 3.4 OS TESTES

Parte de nosso objetivo foi investigar se os termos utilizados nos livros didáticos dificultam a interpretação dos alunos, portanto, somente trabalhar com análise contrastiva do corpus não é suficiente, posto isso, desenvolvemos um experimento com alunos do Fundamental II de uma escola pública de São Carlos para ter uma visão mais próxima de como os alunos lidam com as palavras.

O experimento pautou-se na metodologia *off-line* (FIELD, 2004; GERNSBACKER, 1994; LEITÃO, 2008). Tal método fornece dados do comportamento dos participantes em reação a estímulos linguísticos previamente processados, tais como o julgamento de aceitabilidade. Esse teste foi desenvolvido com os participantes por meio de um formulário no qual constaram palavras para serem julgadas por eles como conhecidas ou não, seguindo como parâmetro a escala Likert<sup>19</sup> de 5 pontos, sendo 0 "não conheço" e 5 "conheço". O outro formulário consistiu num teste de leitura e interpretação o qual foi composto por enunciados de situação-problema em matemática com lacunas que deveriam ser preenchidas pelos participantes ao elegerem quais palavras melhor se encaixavam no contexto. Nesse sentido, concordamos com Grosjean (1998); Gertken et al. (2014) *apud* Oliveira & Sá (2022) ao dizerem que um questionário pode ser uma ferramenta muito útil, por se tratar de um instrumento mais global e permitir a consideração de fatores variados. Dessa forma, para o estudo, consideramos importante podermos averiguar o quanto

---

<sup>19</sup> Desenvolvida pelo psicólogo Rensis Likert em 1932, a escala Likert geralmente consiste em uma série de afirmações que indicam grau de concordância ou discordância do participante sobre alguma questão perguntada. A escala vai de "Discordo totalmente" que pode ser associado ao valor 1 e "Concordo totalmente" ao valor 5, por exemplo. Esse processo permite que os dados sejam analisados quantitativamente, facilitando a interpretação dos resultados e, por isso, é utilizada em pesquisas de satisfação, estudos de opinião pública, e em diversas áreas das ciências sociais, pelo fato de permitir capturar nuances nas respostas dos participantes.

os jovens conhecem os termos utilizados nos enunciados de situação-problema em matemática, a fim de que consigamos avaliar se de fato são ou não termos adequados à faixa etária e se são eles que interferem no processo interpretativo do aluno.

Antes de aplicarmos o experimento, realizamos um teste piloto com 10 alunos, que consistiu em uma lista de palavras para testar a familiaridade desses termos escolhidos e depois, em outro dia, houve a leitura e a interpretação de situações-problema pré-selecionadas. A partir dessa primeira aplicação, analisamos os resultados e realizamos ajustes finos no procedimento, como a inclusão de nome, gênero e idade no formulário, além de configurações estéticas para facilitar a leitura por meio de tablets. Após esses ajustes, prosseguimos para o experimento final que foi dividido em duas etapas: i. teste de familiaridade de palavras; ii. leitura e interpretação de situações-problema. As duas etapas serão detalhadas a seguir.

### 3.4.1 Composição e aplicação do teste de familiaridade

O teste de familiaridade foi construído a partir de 100 palavras retiradas do Corpus de Situação-problema em Matemática (SiPCorp), as quais foram selecionadas de acordo com a seguinte gradação: 25 palavras com frequência alta no SiPCorp e baixa no Corpus Infantil; 25 palavras com baixa frequência no SiPCorp e alta frequência no Corpus Infantil; 25 palavras frequentes em ambos os corpora. Como base de referência para compormos a lista, levamos em consideração os seguintes dados:

*Tabela 3 Referência de valores de frequência*

Alta frequência no SiPCorp- 2,03
Baixa frequência no SiPCorp- 0,3/0
Alta frequência no CI- 6,07
Baixa frequência no CI- 0,3

Abaixo, seguem as listas de palavras utilizadas no teste de familiaridade:

*Tabela 4 Frequência alta no SiPCorp e baixa no Corpus Infantil*

Palavras	Classe de palavras	Freq. SiPCorp	Freq. SiPCorp LOG10	Freq. Infantil	Freq. Infantil LOG10
Retangular	adjetivo	13	1,113943352	57	1,755874856

Justifique	verbo	15	1,176091259	57	1,755874856
Calcule	verbo	24	1,380211242	67	1,826074803
Determine	verbo	41	1,612783857	40	1,602059991
Largura	substantivo	27	1,431363764	257	2,409933123
Construa	verbo	13	1,113943352	265	2,423245874
Percurso	substantivo	12	1,079181246	367	2,564666064
Ângulos	substantivo	14	1,146128036	379	2,57863921
Percorreu	verbo	11	1,041392685	134	2,127104798
Gramas	substantivo	11	1,041392685	490	2,69019608
Soma	verbo	13	1,113943352	527	2,721810615
Gráfico	substantivo	17	1,230448921	470	2,672097858
Represente	verbo	17	1,230448921	139	2,1430148
Fração	substantivo	30	1,477121255	208	2,318063335
Mede	verbo	30	1,477121255	338	2,5289167
Comprimento	substantivo	45	1,653212514	512	2,709269961
Litros	substantivo	57	1,755874856	827	2,91750551
Centímetro	substantivo	82	1,913813852	820	2,913813852
Representa	verbo	24	1,380211242	1786	3,251881455
Altura	substantivo	50	1,698970004	4789	3,680244837
Área	substantivo	37	1,568201724	8509	3,929878524
Quilômetros	substantivo	26	1,414973348	1847	3,266466895
Distribuição	substantivo	13	1,113943352	467	2,669316881

*Tabela 5 Baixa frequência no CSPM e alta frequência no Corpus Infantil*

<b>Palavras</b>	<b>Classe de palavras</b>	<b>Freq. SiPCorp</b>	<b>Freq. SiPCorp LOG10</b>	<b>Freq. Infantil</b>	<b>Freq. Infantil LOG10</b>
Agora	advérbio	5	0,698970004	368701	5,566674316
Vamos	verbo	3	0,477121255	462581	5,665187791
Então	advérbio	6	0,77815125	344521	5,537215699
Mãe	substantivo	2	0,301029996	142495	5,153799626
Pouco	advérbio	1	0	132735	5,122985454
Homem	substantivo	1	0	106873	5,028868001
Hoje	advérbio	5	0,698970004	102393	5,010270268
Precisa	verbo	5	0,698970004	102416	5,01036781
Trabalho	substantivo	4	0,602059991	102841	5,012166291
Grande	adjetivo	4	0,602059991	99219	4,996594846
Dinheiro	substantivo	4	0,602059991	88933	4,949062943
Hora	substantivo	8	0,903089987	79902	4,90255765
Amigo	substantivo	1	0	71318	4,853199156
Pai	substantivo	6	0,77815125	154625	5,189279713
Crianças	substantivo	1	0	42330	4,626648268

Significa	verbo	3	0,477121255	36932	4,567402827
Pequeno	adjetivo	1	0	26253	4,419178939
Metade	substantivo	8	0,903089987	12622	4,101128176
Inteiro	adjetivo	3	0,477121255	9513	3,978317497
Dobro	substantivo	8	0,903089987	2332	3,367728546
Caminho	substantivo	2	0,301029996	33050	4,519171464
Chance	substantivo	1	0	29782	4,473953859
Café	substantivo	2	0,301029996	29632	4,471760965
Comida	substantivo	2	0,301029996	27305	4,436242181
Jogar	verbo	3	0,477121255	26626	4,425305928

*Tabela 6 Baixa frequência em ambos corpora*

<b>Palavras</b>	<b>Classe de palavras</b>	<b>Freq. SiPCorp</b>	<b>Freq. SiPCorp LOG10</b>	<b>Freq. Infantil</b>	<b>Freq. Infantil LOG10</b>
Hodômetro	substantivo	4	0,602059991	12	1,079181246
Arar	verbo	2	0,301029996	97	1,986771734
Triplo	numeral	6	0,77815125	778	2,890979597
Verifique	verbo	2	0,301029996	1469	3,167021796
Gramar	verbo	4	0,602059991	3	0,477121255
Totalizou	verbo	1	0	4	0,602059991
Quíntuplo	numeral	1	0	5	0,698970004
Disposição	substantivo	1	0	955	2,980003372
Mililitros	substantivo	4	0,602059991	9	0,954242509
Polinômio	substantivo	2	0,301029996	9	0,954242509
Traduz	verbo	1	0	110	2,041392685
Divisível	adjetivo	4	0,602059991	10	1
Subtraindo	substantivo	1	0	12	1,079181246
Elabore	verbo	1	0	19	1,278753601
Irredutível	adjetivo	2	0,301029996	22	1,342422681
Algarismos	substantivo	10	1	23	1,361727836
Revestir	verbo	3	0,477121255	25	1,397940009
Proporcionais	substantivo	3	0,477121255	27	1,431363764
Residência	substantivo	5	0,698970004	881	2,944975908
Troque	verbo	5	0,698970004	868	2,938519725
Lote	substantivo	1	0	917	2,962369336
Potência	verbo	2	0,301029996	981	2,991669007
Pares	verbo	5	0,698970004	1064	3,026941628
Aproximadamente	advérbio	11	1,041392685	1080	3,033423755
Unidades	substantivo	8	0,903089987	1097	3,040206628

Escalímetr o	substantivo	1	0	#NUM	#NUM
-----------------	-------------	---	---	------	------

*Tabela 7 Frequentes em ambos corpora*

<b>Palavras</b>	<b>Classe de palavras</b>	<b>Freq. SiPCorp</b>	<b>Freq. SiPCorp LOG10</b>	<b>Freq. Infantil</b>	<b>Freq. Infantil LOG10</b>
Pessoas	substantivo	44	1,643452676	103748	5,015979733
Anos	substantivo	48	1,681241237	113889	5,05648178
Qual	pronome	114	2,056904851	113272	5,054122569
Água	substantivo	68	1,832508913	28084	4,448458964
Número	substantivo	109	2,037426498	24952	4,397105362
Meio	adjetivo	15	1,176091259	51224	4,709473489
Primeira	numeral	14	1,146128036	40832	4,611000653
Razão	substantivo	14	1,146128036	33014	4,518698147
Igual	adjetivo	12	1,079181246	11759	4,07037039
Mês	substantivo	31	1,491361694	12208	4,086644521
Casa	substantivo	28	1,447158031	179035	5,252937941
Menor	adjetivo	27	1,431363764	6406	3,806586934
Pontos	substantivo	27	1,431363764	6338	3,801952235
Parte	adjetivo	25	1,397940009	55359	4,743188236
Tempo	substantivo	24	1,380211242	172786	5,237508551
Cidade	substantivo	22	1,342422681	41918	4,622400553
Ano	substantivo	22	1,342422681	41744	4,620594062
Resposta	substantivo	22	1,342422681	12377	4,092615391
Seguir	verbo	20	1,301029996	12005	4,079362164
Total	substantivo	20	1,301029996	6349	3,802705327
Possível	adjetivo	18	1,255272505	17646	4,246646275
Preço	substantivo	18	1,255272505	7362	3,866995813
Depois	advérbio	17	1,230448921	114962	5,060554311
Conta	substantivo	17	1,230448921	34791	4,541466912
Carro	substantivo	16	1,204119983	72603	4,860954566

As palavras selecionadas foram agrupadas em seções de 5 palavras para que ficasse mais fácil a visualização da escala que mede o grau de conhecimento de cada um quanto ao termo em destaque. Pelo fato de ser uma quantidade alta de palavras e elas não caberem todas na mesma tela, fizemos essa divisão. Quanto a escala, organizamos da seguinte forma:

Não conheço	Já vi alguma vez, mas não sei o sentido	Já vi algumas vezes, mas não sei o sentido	Vejo sempre, mas não sei o sentido	Conheço
-------------	---	---	--	---------

A figura 5 apresenta uma imagem de como ficou o formulário:

Quanto você conhece as palavras a seguir? \*

	Não conheço	Já vi alguma vez	Já vi algumas ...	Vejo sempre, ...	Conheço
Retangular	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Justifique	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Calcule	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Determine	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Largura	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 5 Formulário de teste de familiaridade

É importante frisar que a aplicação do teste foi feita individualmente e por níveis escolares, começando pelos 6ºanos até os mais velhos dos 9ºanos. Eles foram chamados individualmente e responderam o teste em notebook ou tablets, levando cerca de 7 minutos cada um, uma média de 12 palavras por minuto.

Assim que os alunos chegaram à sala, a pesquisadora explicou que cada um deveria ler atentamente as palavras e classificá-las de acordo com o quanto a conhecia. Ela se colocou à disposição para qualquer dúvida, mas deixou claro que só não poderia responder o sentido das palavras da lista, dado que era o objetivo da aplicação.

### 3.4.2 Composição e aplicação do teste de leitura das SiPs

Após a elaboração do teste de familiaridade, consideramos ser estritamente necessário desenvolvermos um teste voltado ao sentido da palavra dentro do contexto linguístico mais amplo e não apenas no sentido isolado das palavras. Assim, selecionamos da lista do teste de familiaridade palavras com alta frequência no CSPM

e baixa no Corpus infantil; baixa frequência no CSPM e alta no Corpus infantil; alta em ambos. Após isso, buscamos pelos enunciados completos<sup>20</sup> nos quais continham esses termos e deixamos no lugar das palavras uma lacuna para que o aluno indicasse qual termo – dentre as opções nas alternativas –, a partir da leitura e interpretação, estava faltando.

Abaixo, constam os termos selecionados:

*Tabela 8 Palavras dos enunciados por série*

	<b>6ºano</b>	<b>7ºano</b>	<b>8ºano</b>	<b>9ºano</b>
<b>Palavras selecionadas por série</b>	Arar Justifique Algarismos Determine Represente Soma	Justifique Determine Represente Escalímetro Comprimento Razão	Justifique Represente Verifique Gramar Razão	Justifique Determine Calcule Percorreu Traduz

Essa seleção foi feita a partir de palavras retiradas dos livros da coleção de cada série, de modo que somente alunos de determinada série pudessem ler e responder o questionário específico para aquela faixa etária. Para nós, foi importante organizar dessa forma para que o contexto seja muito parecido com aquilo que o estudante supostamente encontraria, isto é, um aluno do 6ºano terá muito mais contato com as palavras que estariam presentes em seu contexto escolar. Além disso, também é importante ressaltar que existem palavras que são difícil desde o início escolar e continuam até o final, como é o caso de “determine”.

O teste ficou da seguinte forma:

<sup>20</sup> Todos os enunciados estão no Anexo I.

1. Sabendo que, com um trator, Lúcio vai \_\_\_\_\_ 3/20 de um terreno em um dia, \*  
responda:

- De segunda-feira a sábado, que parte do terreno Lúcio consegue \_\_\_\_\_?

- Ele conseguirá terminar na segunda-feira? \_\_\_\_\_ sua resposta.

- Criar/Criar/Explique
- Arar/Arar/Justifique
- Cozinhar/Cozinhar/Justifique

*Figura 6 Teste de leitura e sentido (6° anos)*

A aplicação do teste foi feita individualmente com a finalidade de que cada um lesse e analisasse de acordo com seu conhecimento e sem interferência de outro colega de sala.

### *3.4.3 Participantes*

A amostra foi composta por 52 participantes, alunos do 6° ao 9°ano de uma escola pública estadual de São Carlos, com faixa etária de 11 a 15 anos. No teste de familiaridade, todos participaram, porém no teste de leitura e interpretação 31 compareceram, sendo 9 alunos do 6° ano, 9 do 7° ano, 6 alunos do 8° ano e 7 alunos do 9° ano. Produzimos 4 testes diferentes, um para cada ano escolar, então o aluno respondia de acordo com a seu grau de escolaridade. A participação foi voluntária, os alunos foram convidados a participar do projeto, com autorização da escola e dos responsáveis seguindo as orientações dos procedimentos autorizados pelo Comitê de Ética em Pesquisa em Humanos da UFSCar, conforme CAAE é 77975224.1.0000.5504

## 4 ANÁLISE DOS DADOS

Para a composição dos dados, dividimos o trabalho em três etapas: (i) trabalho com os corpora: produzimos um corpus de SiP em matemática (SiPCorp) para contrastar com dois grandes corpora – infantil e adulto – com o objetivo de saber a frequência dos termos que aparecem nesses enunciados; (ii) teste de familiaridade: selecionamos palavras com alta e baixa frequência nos corpora, principalmente focando no SiPCorp, com elas produzimos um teste de familiaridade e aplicamos aos alunos para que eles julgassem o quanto conheciam de cada palavra; (iii) teste de leitura e interpretação: elegemos algumas palavras do teste de familiaridade – aquelas que tinham alta e baixa frequência no SiPCorp e produzimos 4 testes, um para cada ano escolar, com 5 SiPs nas quais havia palavras retiradas para que os participantes indicassem qual delas, dentre as opções, estava faltando. É importante ressaltarmos essas partes, pois os dados são resultado de etapas distintas, porém, ainda que tenham sido procedimentos diferenciados, irão convergir para uma explicação, porque possuem relações. Assim, pelo fato de serem processos diferentes, houve também tratamentos diferentes, sendo necessário o detalhamento deles.

### 4.1 TRATANDO OS DADOS

Utilizamos os três corpora: Corpus de Situação-problema em Matemática<sup>21</sup>, LexPorBr Infantil e Corpus Fonotático do Português Brasileiro, a fim de termos uma boa base de quais termos são frequentes ou não para depois construirmos os testes. No Excel, inserimos a planilha com as palavras encontradas nos livros didáticos e a frequência que elas aparecem. Para saber se essas palavras são comuns às crianças, primeiramente, buscamos dentro dos corpora adulto e infantil a frequência de cada um dos termos existentes. Após isso, fizemos uma análise comparativa entre os corpora, a fim de averiguar qual o nível de frequência de cada palavra.

Sabemos que os corpora têm tamanhos diferentes, sendo assim, com a finalidade de equiparmos, primeiramente, realizamos o cálculo Logarítmico 10. Na

---

<sup>21</sup> A partir daqui, referir-nos ao corpus produzido como Corpus de Situação-Problema em Matemática como CSPM.

estatística, o Log de base 10 é utilizado para estabilizar quantidades muito diferentes, com o intuito de torná-las comparáveis, isto é, normalizar a frequência do termo em cada corpus para uma base comum.

Pelo fato de estarmos trabalhando com corpora de diferentes tamanhos, nos apoiamos na discussão proposta por Bestgen (2019), que contribui para a pauta científica sobre a comparação de quantidade de feixes lexicais identificados em corpora que diferem apenas em tamanho. Para esse estudo, foram conduzidas análises em cinco corpora bem conhecidos, dos quais foram extraídos subcorpora de 100.000 a dois milhões de palavras por dois procedimentos complementares de amostragem aleatória. As análises realizadas apontaram que quanto menor o corpus, maior o número de feixes selecionados por um dado limiar normalizado. Essa conclusão é relacionada à natureza Zipfiana da distribuição do vocabulário em corpora de linguagem natural. Isso ocorre porque, em tais corpora, há muito mais sequências de palavras raras do que frequentes.

A Lei de Zipf propicia uma base matemático-linguística para analisar frequência e distribuição das palavras contidas em um texto. Por meio de cálculo, é possível mapear e criar rankings de ocorrência das palavras no texto (Zipf, 1949), evidenciando que existe uma lei do menor esforço (Egghe, 1991; Zörnig; Altmann, 1995; Günther et al., 1996 apud Pinto, 2015), assim como defende Manning & Schütze (1999, p. 24) *apud* Bestgen (2019): “há algumas palavras muito comuns, um número mediano de palavras de média frequência e muitas palavras de baixa frequência”.

Yves Bestgen (2019), além de desenvolver uma abordagem científica mostrando como o tamanho do corpus influencia na frequência, também defende que é mais seguro para a análise trabalhar com corpora de tamanhos semelhantes, uma vez que, ao seu olhar, as equiparações podem interferir na leitura dos dados. Todavia, Marc Brysbaert e Kevin Diependaele (2013) apresentam outra perspectiva, que defende que as medidas de frequência baseadas em corpora muito grandes, embora forneçam estimativas para todas as palavras, não fornecem estimativas muito boas. Duas razões são apresentadas: a primeira é que esses corpora geralmente se baseiam em fontes não muito frequentes aos alunos de graduação, por exemplo, enciclopédias; a segunda razão é que corpora muito grandes tendem a conter relativamente mais palavras raras do que as que os participantes provavelmente encontraram em suas vidas.

Seguindo essa lógica, acrescentamos mais uma justificativa para enfatizar que o uso de corpora de tamanhos diferentes não é negativo à análise, baseada em nossa análise em relação a algumas palavras existentes no CSPM – incomuns ao público-alvo. Alguns termos aparecem com baixíssima frequência no CSPM, o que é até interessante uma vez que são palavras incomuns aos alunos, por exemplo, a forma verbal “desenhá-lo” que possui frequência 1 no CSPM, porém com frequência 21 no LexPorBR Infantil. Claro que ainda é uma baixa frequência, pensando no tamanho do corpus, mas, ainda assim, pensando que é um verbo com complemento direto, não é uma estrutura natural de ser utilizada por crianças, porque não é comum vermos na fala do cotidiano crianças utilizando verbos com complemento direto ou indireto em suas construções. Nesse caso, acreditamos que isso ocorre, pois o LexPorBr Infantil foi construído, para além de outros textos, por meio de legendas de filmes, enunciados que são escritos e mais formais, que não condizem muito com a fala da criança e a audição dela. Sendo assim, embora o nosso corpus seja menor em relação ao infantil, ele parece ser mais coerente em relação à quantidade de frequência.

Para estabelecermos comparação entre os corpora, principalmente com o objetivo de averiguar a relação entre eles, utilizamos o teste de Correlação de Pearson. É um teste que avalia a relação estatística entre duas variáveis contínuas. Se a relação entre os elementos não for linear, o coeficiente não refletirá corretamente essa associação.

Os valores do coeficiente de correlação de Pearson variam de +1 a -1. O valor 0 (zero) indica que não há associação entre as duas variáveis, já um valor positivo indica uma associação positiva, ou seja, quando o valor de uma variável aumenta, o valor da outra também aumenta. Por outro lado, um valor negativo indica uma associação negativa, significando que quando o valor de uma variável aumenta, o valor da outra diminui.

Antes de calcular a correlação de Pearson, deparamo-nos com uma questão importante: como lidar com palavras de frequência 0. Ao equiparar os corpora por meio do Log10, os termos que possuem 0 como frequência em um dos corpora apresentaram um erro (#NÚM!), ou seja, o cálculo não funcionou nesses casos. Assim, precisávamos de alguma alternativa para a tratativa dessas frequências.

Dentre as opções, segundo Brysbaert e Diependaele (2013), há algumas alternativas, sendo a primeira o descarte das palavras com frequência 0. Para nós, não é uma ação válida, uma vez que é importante ao trabalho que tenhamos clareza

sobre os termos que não são frequentes às crianças, dado que é nosso objetivo averiguar exatamente a presença de termos incomuns ao público-alvo.

A segunda alternativa é adicionar 1 a cada valor de frequência. Assim, as palavras não observadas no corpus recebem uma contagem de frequência de 1; palavras observadas uma vez no corpus adquirem frequência 2; palavras observadas duas vezes obtêm uma frequência de 3; e assim por diante. Esta transformação foi originalmente proposta por Laplace e é conhecida como “alisamento de Laplace”. É uma boa alternativa, pelo fato de não excluir nenhum item e também alterar muito pouco a quantidade de dados do corpus, principalmente em corpora maiores, todavia é uma tentativa de simplificar algoritmos de maneira que não sejam totalmente inaceitáveis matematicamente.

Outra opção é adicionar um pequeno valor, prática denominada de “suavização de Lidstone”, porém essa manobra coloca as palavras com frequência zero tão longe das outras palavras que elas se tornam discrepantes para análises estatísticas, o que pode ter um forte impacto no resultado, anulando uma correlação oposta ao efeito dos valores discrepantes ou aumentando a significância de uma correlação insignificante em linha com os valores discrepantes.

Por fim, outra alternativa é utilizar o algoritmo de Good-Turing, que leva em consideração a comparação de palavras com frequência 1 e 2 para que seja dado um número para os termos com frequência zero. As frequências corrigidas podem ser obtidas comparando os tamanhos das várias caixas de frequência. Assim, a frequência correta das palavras vistas uma vez no corpus é estimada com base no número de palavras vistas uma vez em relação ao número de palavras vistas duas vezes. Além disso, leva-se também em consideração que a probabilidade de observações invisíveis é igual à probabilidade das palavras vistas uma vez. Embora seja uma opção viável, o algoritmo de Good-Turing não fornece uma estimativa direta para palavras não observadas.

Refletimos sobre todas as possibilidades estatísticas e decidimos que, embora existam essas alternativas, inclusive duas interessantes, em uma primeira visão, a Laplace e o algoritmo Good-Turing, nenhuma delas consegue de fato nos auxiliar, uma vez que a primeira modifica o valor da frequência e, para nós, o zero é muito importante, e a segunda não nos fornece uma resposta plausível sobre a frequência. A frequência zero é um dado fundamental ao presente estudo, não queremos neutralizá-la, pelo fato de indicar que é um termo com o qual os alunos deparam-se

durante a leitura das SiPs, porém não consta em dois grandes *corpora* (infantil e adulto). Sendo assim, por enquanto, trabalharemos com os termos que possuem pelo menos uma frequência e os que possuem frequência zero deixaremos separados para analisá-los posteriormente.

Após igualar as quantidades por meio do Log10, utilizamos a correlação de Pearson, visto que é necessário saber o quanto a variabilidade da frequência lexical de um corpus (dos livros didáticos de matemática) tem relação com a variabilidade da frequência lexical dos outros corpora (infantil e adulto), a fim de avaliarmos a coerência entre eles e, logo, se os termos utilizamos nas SiPs são adequados. Com isso, no próximo capítulo, apresentaremos os resultados desses levantamentos quantitativos.

Além dos dados obtidos pela composição e análise dos corpora, temos também os dados que conseguimos por meio do desenvolvimento e da aplicação dos testes de familiaridade e de leitura e interpretação.

Para o teste de familiaridade, utilizamos a escala Likert de 0 a 5 – ímpar –, sendo 0 “não conheço” e 5 “conheço”, que é muito utilizada em formulários e pesquisas de opinião. Dentro dessas extremidades, ainda existiam “já vi alguma vez, mas não sei o sentido”, “já vi algumas vezes, mas não sei o sentido” e “vejo sempre, mas não sei o sentido”. Diante disso, entendemos que conhecer um termo pressupõe que o falante saiba o sentido para usá-lo, sendo assim, qualquer resposta diferente de “conheço”, categoria criada por nós como “palavras conhecidas”, classificamos como “palavra medianamente conhecida” e “palavra não-conhecida”. Essa forma de tratar os dados foi essencial para termos uma visão de quais termos seriam potenciais empecilhos aos alunos.

Nos testes de leitura e interpretação, primeiramente, analisamos as respostas para ver quantos alunos acertaram. Após isso, cruzamos as respostas do teste de familiaridade com o teste de leitura e interpretação, levando em consideração a resposta de cada aluno que estava participando do segundo formulário, lembrando que do teste de familiaridade participaram 55 alunos e no de leitura e interpretação 31, de modo que todos esses participaram antes do de familiaridade.

Ainda sobre os dados do teste de leitura e interpretação, transformamos em porcentagem a quantidade de alunos que diziam conhecer o termo e erraram, que

diziam conhecer e acertaram, que diziam não conhecer e erraram e que diziam não conhecer e acertaram<sup>22</sup>.

Na próxima seção, apresentaremos os resultados das análises quantitativas em números absolutos e qualitativas, primeiro, do contraste em os corpora e, depois, dos testes aplicados.

---

<sup>22</sup> O Quadro que ilustra essa prática encontra-se na próxima seção

## 5 RESULTADOS

É importante iniciarmos este capítulo destacando que se trata de um trabalho amostral, dessa forma, seguimos sob perspectiva da estatística inferencial para testar nossa hipótese de que as palavras que constam nas SiPs de livros didáticos de matemática não são frequentes e comuns aos alunos e, por isso, o processo interpretativo pode sofrer prejuízo.

Após a equiparação dos dados, devido aos tamanhos diferentes de corpora, por meio do Log10, verificamos se as diferenças encontradas entre os corpora são significativas, ou seja, se a probabilidade (valor  $p$ ) é maior ( $>$ ) ou menor ( $<$ ) do que 0,05. Para medir a magnitude das correlações, utilizamos o software IBM SPSS Statistics (versão 29.0.2.0 (20)), desenvolvido pela IBM, pelo qual foi possível obtermos o valor de  $p$  e fazer o teste da Correlação de Pearson. O valor de  $p$  representa a probabilidade de a diferença detectada entre os grupos analisados – no nosso caso os *corpora* – ter ocorrido ao acaso. Com isso, um resultado igual ou menor a  $<0.05$  indica que há uma pequena probabilidade de que a diferença observada entre as variáveis seja ao acaso, então, considera-se que há diferença significativa entre as variáveis e o contrário a isso caso o resultado seja  $>0.05$ . Com base nisso, as primeiras comparações entre os corpora demonstraram, em número absolutos, os seguintes dados:

Tabela 9 Correlação de Pearson e valor de  $p$

Corpora comparados	Resultado da correlação R	Valor de $p$
SiPCorp e LexPorBr Infantil	0,3782856	$<0.01$
SiPCorp e Corpus Fonotático PB	0,4017462	$<0.01$
Corpus Fonotático PB e LexPorBr Infantil	0,7985612	$<0.01$

Como podemos observar, os corpora foram analisados de modo comparativo em pares. Lembrando que o valor de  $p$  é dedicado para determinar a significância estatística de um teste de hipótese, a partir de nossos resultados, com o resultado do valor de  $p < 0.01$ , indica-se que a probabilidade de nossos resultados serem ao acaso é de 1%, ou seja, muito baixa. Assim, esse resultado nos sugere que há correlação entre as variáveis SiPCorp e os corpora LexPorBr (infantil) e Corpus Fonotático do

Português brasileiro (adulto), uma vez que o valor  $p$  foi significativo. É claro que existem outros fatores que corroboram a análise e avaliação de uma pesquisa, para além do valor de  $p$ , mas achamos importante realizar o teste.

Ao mesmo tempo que obtivemos o resultado de  $p$ , também acessamos os resultados da Correlação de Pearson. Partindo da comparação entre SiPCorp e LexPorBr Infantil, nota-se que a correlação entre eles é fraca, o que significa que os termos utilizados nos livros didáticos não são tão comuns aos alunos. Outrossim, ao observar a correlação entre SiPCorp e o Corpus Fonotático do Português brasileiro, temos também um valor baixo de correlação (0,378). No entanto, a correlação entre LexPorBr Infantil e Corpus Fonotático do Português brasileiro é forte (0,798), o que indica que a frequência de vocabulário é muito parecida.

Sendo assim, vocabulário que consta nos livros didáticos correlaciona-se de forma bem fraca com o léxico esperado de uma criança, de acordo a amostra do corpus infantil, assim como também acontece com o vocabulário adulto, pois a correlação é baixa ao compará-lo com os dois corpora. Isso quer dizer que as frequências lexicais encontradas no CSPM são distantes dos dois léxicos (infantil e adulto).

Partindo para uma análise mais específica, analisamos também a correlação entre o CSPM e os demais corpora em se tratando das classes gramaticais que constam nas SiPs. Selecionamos somente aquelas que possuem sentidos mais completos e não somente se ligam os termos, a saber, substantivo, verbo, adjetivo, numeral e advérbio. Retiramos, como consta na metodologia, as preposições e artigos, pois são palavras de ligação e determinação, que, na nossa amostra, não causaram estranhamento ao aluno, por isso não dedicamos a análise a essas categorias.

A seguir, apresentamos os resultados utilizando o coeficiente de relação de Pearson:

*Tabela 10 Correlação do coeficiente R entre os corpora CSPM e LexPorBr Infantil*

<b>Classe gramatical (variável)</b>	<b>Resultado da correlação R</b>	<b>Valor de P</b>
<b>Substantivo</b>	0,3572114	<0.01

<b>Verbo</b>	0,2720294	<0.01
<b>Adjetivo</b>	0,2853069	<0.01
<b>Pronome</b>	0,2479919	<0.01
<b>Numeral</b>	0,5556978	<0.01
<b>Advérbio</b>	0,3231099	<0.01
<b>Conjunção</b>	0,6108191	<0.01

Ao analisar os substantivos, nota-se que a correlação entre os corpora é fraca, assim como as classes verbo, adjetivo, pronome e advérbio. Isso nos mostra que a escolha de termos que compõem as SiPs não é tão próxima ao que parece ser o universo de uma criança, dado que o nível de correlação entre o corpus dos livros didáticos e o infantil é baixo. Ademais, as categorias gramaticais conjunção e numeral correlacionam-se de forma moderada, ou seja, ainda que exista uma correlação baixa, é mais coerente que as demais classes, o que pode acontecer pelo fato de serem classes fechadas, invariáveis e menores em número de formas.

A tabela 5 refere-se ao coeficiente de correlação de Pearson entre os corpora adulto (Corpus Fonotático do PB) e de livros didáticos (CSPM):

*Tabela 11 Correlação do coeficiente R entre os corpora CSPM e Corpus Fonotático do Português Brasileiro*

<b>Classe gramatical (variável)</b>	<b>Resultado da correlação R</b>	<b>Valor de P</b>
<b>Substantivo</b>	0,43197222	<0.01
<b>Verbo</b>	0,26627053	<0.01
<b>Adjetivo</b>	0,28600699	<0.01
<b>Pronome</b>	0,38	<0.01
<b>Numeral</b>	0,52067264	<0.01
<b>Advérbio</b>	0,43714986	<0.01
<b>Conjunção</b>	0,24331051	<0.01

Em se tratando da correlação entre as classes gramaticais, ainda há correlações baixas, como as categoriais gramaticais verbo, adjetivo, pronome e conjunção. O substantivo tem uma correlação mais expressiva, embora baixa, que os dados anteriores, mas se enquadra como sendo moderada, assim como o advérbio e o numeral.

Por fim, a correlação do coeficiente de Pearson entre os corpora infantil e adulto:

*Tabela 12 Correlação do coeficiente R entre os corpora Corpus Fonotático do Português Brasileiro e LexPorBr Infantil*

<b>Classe gramatical (variável)</b>	<b>Resultado da correlação R</b>	<b>Valor de P</b>
<b>Substantivo</b>	0,827828485	<0.01
<b>Verbo</b>	0,735051019	<0.01
<b>Adjetivo</b>	0,780832889	<0.01
<b>Pronome</b>	0,814739222	<0.01
<b>Numeral</b>	0,931289429	<0.01
<b>Advérbio</b>	0,905262393	<0.01
<b>Conjunção</b>	0,894147639	<0.01

Ao observarmos os dados, já podemos notar a diferença em relação à correlação dos resultados anteriores. Diferentemente da correlação entre o corpus de livro didático e os corpora de adulto e de criança, a qual obteve como análise baixa e moderada correlação, os corpora LexPorBr Infantil e Corpus Fonotático do Português Brasileiro (adulto) possuem coeficiente de correlação forte nas categorias substantivo, verbo, adjetivo, pronome e conjunção. Nas demais categorias, numeral e advérbio, o coeficiente de correlação é muito forte, ultrapassando de 0,9.

Para ilustrar os coeficientes de correlação, utilizamos gráficos de dispersão para visualizar a relação entre duas variáveis. Os gráficos nos ajudam a entender melhor os números e ilustra a correlação. Vejamos abaixo:

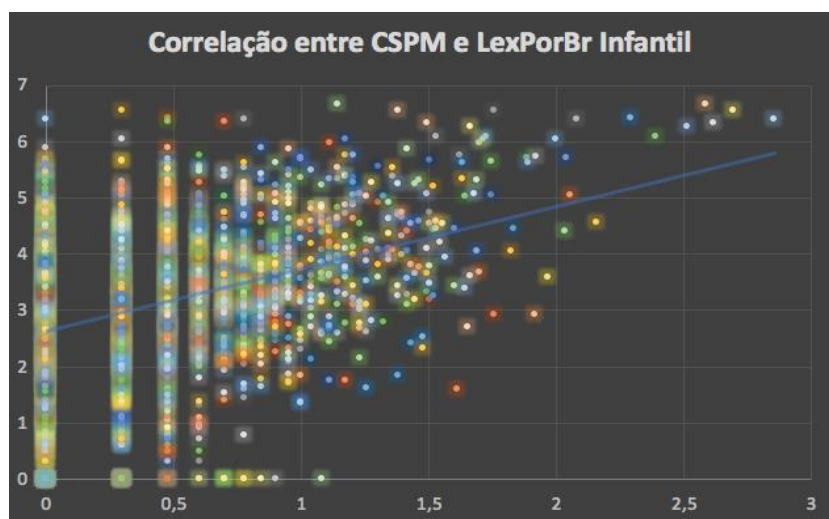


Figura 7 Correlação entre SiPCorp e LexPorBr Infantil

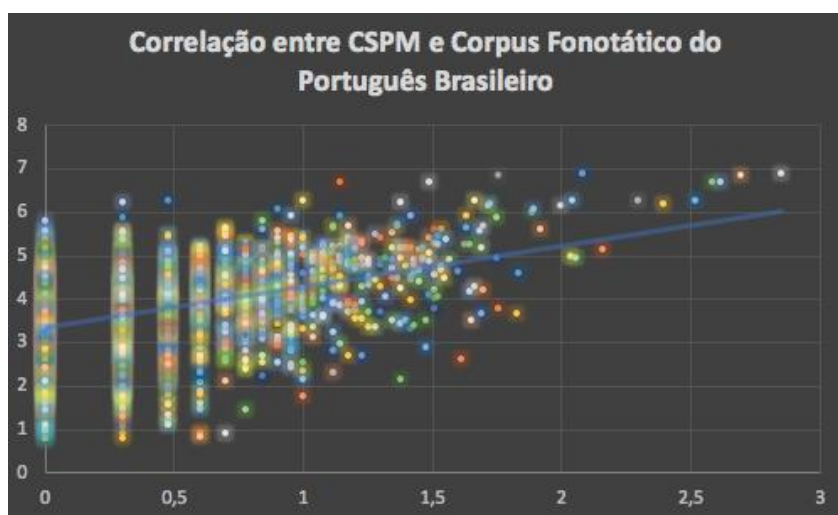


Figura 8 Gráfico de correlação entre SiPCorp e Corpus Fonotático PB

As Figuras 1 e a 2 representam a correlação entre o CSPM e os corpora LexPorBr Infantil e o Corpus Fonotático do Português Brasileiro, assim podemos observar, de início, que se trata de uma correlação positiva, isto é, há uma concentração dos pontos direcionados a uma crescente. Todavia, há muitos pontos dispersos, pelo fato de a correlação ser fraca no primeiro e moderada no segundo, podemos inclusive analisar que não há proximidade entre os pontos e uma tendência crescente que aumenta.

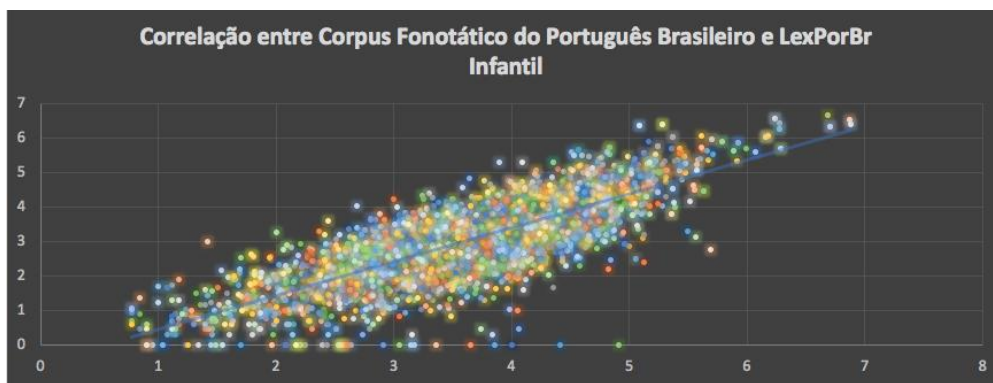


Figura 9 Correlação entre Corpus Fonotático PB e LexPoxBr Infantil

Já na Figura 3, ainda que também seja uma correlação positiva, podemos notar maior proximidade entre os pontos, ou seja, estão menos dispersos. Com isso, concluímos que esses corpora possuem maior correlação, logo são mais coerentes entre si.

Após analisar os dados mais gerais, corpora e as classes gramaticais que constam nesses corpora, proporemos uma análise preliminar e ilustrativa de alguns termos que nos chamaram a atenção durante a construção e observação do corpus SiPCorp pelo fato de serem termos de baixa frequência e outros com alta frequência, porém com sentido não tão claro aos alunos:

Tabela 13 Termos e suas frequências

Termos	Freq. <sup>23</sup> . absoluta (SiPCorp)	Freq. estabilizada (SiPCorp)	Freq. absoluta (LexPor Infantil)	Freq. estabilizada (LexPor Infantil)	Freq. absoluta no (CFPB <sup>24</sup> )	Freq. estabilizada (CFPB)
<b>Determine</b>	47	1,61	40	1,60	429	2,63
<b>Justifique</b>	15	1,17	57	1,75	491	2,69
<b>Retangular</b>	13	1,11	57	1,75	207	2,31
<b>Alameda</b>	10	1	22	1,34	2.089	3,31
<b>Indique</b>	3	0,47	112	2,04	562	2,74
<b>Prestação</b>	4	0,60	114	2,05	6.983	3,84

Na tabela, encontram-se alguns termos lexicais que compõem os enunciados matemáticos da coleção Bianchini (2018), os quais podem causar estranhamento aos

<sup>23</sup> Frequência absoluta.

<sup>24</sup> Corpus Fonotático do Português Brasileiro.

alunos pelo fato de serem pouco frequentes a eles, segundo os dados do LexPorBr Infantil.

O verbo “justifique” tem seu uso frequente com o fito de pedir explicação, alguma justificativa de resolução de SiP. Embora ele seja um termo frequente, com 15 usos, para os estudantes, não se trata de um termo muito conhecido, principalmente se pensarmos que esse tem 57 ocorrências no LexPorBr Infantil. Nesse sentido, para muitos pode ser uma palavra estranha ou que se aplica ao contexto da diagramação.

Ainda pensando nos verbos que se destacaram, o “determine” e suas variantes é o que possui maior frequência no SiPCorp, 47 usos, e o que nos chama a atenção é que esse verbo, ao contrário do “justificar” que tem um sentido mais específico (explicar), parece mais abrangente. O que é, de fato, “determinar”? Vejamos os exemplos a seguir:

10. A professora de Artes distribuiu igualmente 7 cartolinas para 3 grupos de alunos. **Determine** a quantidade de cartolina que cada grupo recebeu na forma de fração e na forma mista. (BIANCHINI, 2018)

Em ambos os casos, o “determine” poderia ser facilmente substituído por “indique”, um termo mais frequente aos alunos, com frequência de 112 no LexPorBr Infantil, o triplo de ocorrências, no entanto optou-se por uma palavra não tão comum ao público-alvo. Ao analisar o verbo “indique” no nosso corpus, ele é menos frequente, apenas com 3 ocorrências, o que parece demonstrar que não há muita preocupação com a escolha desse item. Seguindo nessa linha, “prestação”, termo analisado em pesquisa anterior (AGOSTINI, 2019), repete-se também nessa coleção, com 4 ocorrências no CSPM e 114 no LexPorBr Infantil, ou seja, ainda que “parcela” seja muito mais comum aos alunos, com 178 ocorrências no LexPorBr Infantil, os autores permanecem escolhendo termos que são mais usuais às faixas etárias diferentes daquelas dos estudantes, como adultos, público que lida com prestação em seu cotidiano.

Até aqui os resultados já nos indicavam que a familiaridade com o léxico realmente parece interferir na interpretação do aluno, mas ainda assim achamos necessário realizar alguns testes para que tivéssemos uma visão mais específica em relação ao julgamento do aluno quanto ao conhecimento de algumas palavras.

Assim, para o teste de familiaridade, 55 participantes responderam ao formulário indicando o nível de familiaridade com as 100 palavras selecionadas. A partir disso, chegamos aos resultados:

*Tabela 14 Porcentagem por categoria*

	<b>Conhecidas</b>	<b>Medianamente conhecidas</b>	<b>Desconhecidas</b>
<b>Palavras do teste de familiaridade</b>	85%	11%	4%

Das 100 palavras avaliadas, 85% foram indicadas por eles como conhecidas, 11% medianamente conhecidas, isto é, já viram alguma vez, mas não sabem o sentido e 4% delas foram apontadas como desconhecidas. Além disso, também vimos como é importante saber a frequência dessas palavras julgadas por eles no SiPCorp, tendo em vista que o contato dos discentes com o livro de matemática é diário. Organizamos esses dados na tabela abaixo:

*Tabela 15 Frequência das palavras categorizadas no CSPM*

	<b>Conhecidas</b>	<b>Medianamente conhecidas</b>	<b>Desconhecidas</b>
<b>Baixa frequência no SiPCorp</b>	41%	100%	100%
<b>Alta frequência no SiPCorp</b>	59%	0%	0%

A tabela 15 nos mostra a porcentagem de palavras classificadas como conhecidas pelos alunos e a frequência delas no SiPCorp, sendo 59% das palavras indicadas por eles como familiar e frequentes no corpus de SiPs. Além disso, 41% das palavras conhecidas por eles são de baixa frequência e todas indicadas como medianamente conhecidas e desconhecidas também não são muito frequentes no corpus.

Dentre as palavras medianamente conhecidas estão: “totalizou”, “quíntuplo”, “divisível”, “elabore”, “subtraindo”, “revestir”, “proporcionais”, “lote”, “gramar” e “irredutível”. Em se tratando das desconhecidas, os termos indicados foram “hodômetro”, “arar”, “escalímetro” e “polinômio”. Ainda que nenhuma delas seja frequente no CSPM, são, pelo menos, 15 enunciados que podem ser obstáculos para

o aprendizado do aluno, uma vez que, segundo os resultados do teste de leitura e interpretação, 37% dos alunos do 6º ano, embora conhecessem a palavra, erraram na escolha para o contexto determinado no enunciado, conforme consta abaixo.

*Tabela 16 Indicação de conhecimento ou desconhecimento de um termo por ano escolar*

	6º ano	7º ano	8º ano	9º ano	Todos os anos
<b>Quantidade de vezes, em porcentagem, que as crianças assinalaram que conheciam a palavra, mas erraram</b>	37%	7%	17%	6%	17%
<b>Quantidade de vezes, em porcentagem, que as crianças assinalaram que não conheciam a palavra e acertaram</b>	16%	27%	17%	11%	18%
<b>Quantidade de vezes, em porcentagem, que as crianças assinalaram que conheciam a palavra e acertaram</b>	40%	40%	60%	80%	53%
<b>Quantidade de vezes, em porcentagem, que as crianças assinalaram que não conheciam a palavra e erraram</b>	7%	22%	7%	3%	10%

Achamos importante destacar que, mesmo dizendo conhecerem o termo, isso pode significar que o aluno pode já ter ouvido ou visto, mas não que tenha intimidade para o uso, então o erro no emprego lexical pode ser causado quando selecionamos um item.

Analisando os dados, é possível notar que quanto mais avançado é o nível do aluno, supostamente melhor ele interpreta os contextos dos enunciados, visto que a maior porcentagem na análise comparativa tivemos um aumento de 40% para 80% em relação ao conhecimento e acerto da palavra no contexto de alunos do 6º ao 9º anos. Em relação aos 6ºanos, 37% deles disseram conhecer a palavra, mas não acertaram seu significado dentro do contexto, o que pode ter uma relação com o fato de eles serem mais jovens, com um vocabulário ainda não consolidado quanto ao uso. Já em relação aos dados da quantidade de vezes que as crianças indicaram não conhecer uma palavra e erraram, de forma geral, tivemos o resultado de 10%, isto é, ainda que pareça pouco em relação ao total, é um resultado significativo, pois há então crianças sendo prejudicadas pela escolha lexical dos enunciados.

Ademais, chamou-nos atenção a quantidade de vezes que as crianças acertaram as palavras que classificaram como desconhecidas (18% em relação a todos os anos). Isso nos mostra que, apesar de o contexto auxiliar quando não há conhecimento do termo lexical, somente 18% ter acertado é preocupante, o que

evidencia a importância de um enunciado ser composto por vocabulário comum a eles, pois somente o contexto não dá conta de auxiliá-los. Embora os acertos tenham sido mais frequentes que os erros, o que observamos é que o léxico é extremamente importante, principalmente quando há um termo desconhecido, sendo necessário conhecer os demais termos para que aquela lacuna interpretativa seja preenchida pelo contexto.

Selecionamos as palavras que os alunos erraram no teste de interpretação para analisarmos a frequência dentro do SiPCorp:

*Tabela 17 Palavras que erraram*

Palavras que erraram	Frequência no SiPCorp	Classificação da frequência dentro do SiPCorp
Escalímetro	0	Baixa
Comprimento	1,653212514	Alta
Arar	0,301029996	Baixa
Algarismos	1	Baixa
Determine	1,612783857	Alta
Represente	1,230448921	Alta
Razão	1,46128036	Alta
Verifique	0,301029996	Baixa
Gramar	0,602059991	Baixa
Justifique	1,176091259	Alta

Dos 10 termos que os participantes erraram, 50% são frequentes no SiPCorp e 50% não são frequentes no corpus. Novamente, um dos termos que se destaca é o verbo “determine”, o qual já era um vocábulo analisado por nós desde a análise contrastiva dos corpora e o colocamos exatamente para observar como os participantes lidariam com ele. A seguir, constam os enunciados com “determine” que os participantes do 6º, 7º e 9º ano, respectivamente, tinham que interpretar e preencher a lacuna com: “determine”, “pense” ou “reflita” (6º e 9 anos); “eleja”, “determine” e “justifique” (7ºano):

11. Em uma receita de bolo de chocolate, são necessários 3/3 copos de leite. Sabendo que em um copo cabem 200 mililitros, determine quantos mililitros de leite serão necessários para essa receita. (BIANCHINI, 2018)
12. Em certo país, as eleições para presidente ocorrem a cada 4 anos, e para senador, a cada 8 anos. Em 2014, essas eleições coincidiram.

determine os anos das quatro próximas vezes em que as eleições voltarão a coincidir. (BIANCHINI, 2018)

13. Em um jogo de futebol, Paulo cobra uma falta jogando a bola para Marcos, que a joga para Daniel. A trajetória da bola está representada na figura abaixo. Determine as medidas dos ângulos assinalados em vermelho na figura. (BIANCHINI, 2018)

Dos 9 participantes do 6º ano, 44% indicaram conhecer a palavra, mas erraram a seleção; 22% disseram não a conhecer, mas acertaram; 22% dos que colocaram que a conheciam, acertaram. Em relação ao 7º ano, também foram 9 participantes os quais 78% disseram não conhecer o termo “determine”, porém o acertaram, enquanto somente 22% disseram conhecer o termo e acertaram. Os participantes do 9º ano, 7 alunos, não demonstraram muita dificuldade com essa palavra, uma vez que 71% conheciam e acertaram a escolha, enquanto somente 14% conheciam, mas erraram e outros 14% não conheciam e acertaram.

Outro termo que nos chamou a atenção desde as análises dos corpora foi “verifique”, pelo fato de aparecer bastante em enunciados matemáticos, não necessariamente em SiPs. Ele apareceu no teste dos 7º anos no seguinte enunciado:

14. A tabela abaixo mostra o tempo, em hora, que os meninos e as meninas do oitavo ano do colégio Alegria acessam a internet semanalmente. Tempo semanal de acesso à internet Tempo Classe (de a até -).  
Frequência absoluta: meninos. Frequência absoluta: meninas.  
Construa a tabela de frequências relativas na forma percentual e verifique se sua resposta ao item a está correta.

Ainda que tenha sido uma das palavras classificadas como conhecidas pelos participantes, o que se nota é que a veem bastante, todavia não sabem utilizá-la, uma vez que 67% dos alunos que disseram conhecê-la erraram o teste, somente 33% acertaram. As opções, além de “verifique”, eram “determine” – a mais assinalada – e “relacione” – escolhida por somente um participante. É interessante notar que o fato de os alunos terem escolhido majoritariamente “determine” mostra que eles também não sabem o sentido desse verbo, pois ali naquele contexto não seria possível seu uso. Isso porque, na área da matemática, “determinar” assume um significado

diferente; está relacionado à resolução de problemas que envolvem a descoberta de valores desconhecidos ou até mesmo à definição de relações entre grandezas. No caso do enunciado, o comando é para que o aluno se certifique de que a resposta está correta.

Ainda sobre as palavras que os participantes erraram no teste de interpretação, o termo “grammar” foi uma palavra em que a maioria, 67%, disse não conhecer, entretanto, desse total, 50% acertou, ainda que não conhecesse o termo.

O último enunciado que analisaremos é também do 7 ano, do qual os termos retirados foram “escalímetro” e “comprimento”:

15. Com um \_\_\_\_\_ , a planta de uma casa foi desenhada na escala 125. Nessas condições, responda as questões.

Qual é o \_\_\_\_\_ real, em metro, de uma sala que, nessa planta, foi representada por um segmento de 5,2 cm?

A SiP pede um contexto muito específico em que é necessário conhecer um termo o qual nem mesmo o contexto consegue auxiliar, pois se trata de uma ferramenta cuja função é pode ser específica: auxiliar no desenho de plantas de casa. Dentre as opções, havia pincel – a escolhida por todos os participantes – que também é um utensílio para desenho, mas não no caso de uma residência que necessita de medidas convertidas em escala. Apesar de a frequência de “escalímetro” ser baixa, somente 1 ocorrência em toda a coleção, trata-se de uma palavra que nem o contexto consegue contribuir para que haja uma boa interpretação, pois é um termo específico, técnico. No entanto, essa situação em que nem o contexto contribui não acontece somente com “escalímetro”, mas também com outras palavras que listamos abaixo:

*Tabela 18 Palavras desconhecidas*

Palavras desconhecidas	Frequência no SiPCorp	Não conhecem/ sabem o sentido (%)	Conhecem/ sabem o sentido(%)
<b>Hodômetro</b>	0,602059991	94%	7%
<b>Arar</b>	0,301029996	63%	37%
<b>Escalímetro</b>	0	93%	7%
<b>Polinômio</b>	0,301029996	96%	4%

Por fim, de fato, a escolha lexical das SiPs indica ser um fator realmente problemático à interpretação desses enunciados, uma vez que, ao analisar os corpora como variáveis, a correlação do SiPCorp em relação ao infantil é baixa, o que indica que os termos utilizados não são comuns às crianças. Chegamos a essa conclusão porque, além dessa análise, também fizemos os testes que nos mostraram o quanto importante é a escolha lexical, dado que nem sempre o contexto consegue dar o suporte necessários para que a criança consiga interpretar um enunciado matemático.

Claro que existem muitos outros termos não abordados aqui que também podem ser potenciais obstáculos interpretativos – aqui temos somente uma amostra –, porém acreditamos que, por esse panorama, já seja possível notarmos que o léxico é um ponto crucial na elaboração de problemas para que seja, depois, possível sua resolução sem qualquer dificuldade relacionada às palavras.

## 6 REFLEXÕES E ENCAMINHAMENTOS

O que buscamos com este trabalho, além do que discutimos até aqui, claro, é promover uma reflexão sobre a temática. O senso comum costuma colocar Português e Matemática como disciplinas e conteúdos em oposição e em campos cuja eficiência em uma obrigatoriamente leva ao fracasso na outra e vice-versa. Cria-se o mito de que ninguém pode ter desempenho satisfatório em ambas. Essa visão estereotipada pode (des)motivar. No fundo, esse comportamento cultural é prejudicial à toda a comunidade de aprendizagem.

A interdisciplinaridade que é possível estabelecer entre essas duas áreas do conhecimento é importantíssima para que entendamos que é por meio da língua natural, no nosso caso o português, que entenderemos a linguagem matemática e se o enunciado não estiver claro com palavras frequentes e conhecidas, ou melhor, que possam ser compreendidas, não haverá uma interpretação satisfatória. Para além dessa questão, devemos entender que, na aula de Português, tanto a competência e o desempenho na língua materna, que são ou deveriam ser objetos da aprendizagem, quanto as competências e o desempenho nas operações e interpretações da sua linguagem, na Matemática, também ou deveria ser o seu objeto de aprendizagem. Ou seja, todas as “disciplinas” estão ligadas por uma linguagem que as operacionaliza, mas também que em algumas delas, é o objeto.

Quando questionamos sobre a frequência de determinadas palavras que compõem os enunciados das SiPs, não pretendemos que sua redação passe por um crivo de “esse termo sim” *versus* “esse termo não”, até porque isso faria com que os educandos não tivessem a oportunidade de ampliação de seu vocabulário. O que se deseja com as reflexões deste trabalho é uma ação contextualizada dentro da sala de aula, nas leituras, nas propostas de livros didáticos e seus exercícios: glossários, balões, destaques, imagens ilustrativas etc. Se quanto mais o aluno lê, mais habilidoso ele se torna na leitura, precisamos, claro, que esse ato seja praticado e frequentemente, principalmente com palavras novas para que seu vocabulário seja ampliado, mas isso não precisa acontecer, por exemplo, em um momento avaliativo – como exames semestrais, provas governamentais ou até mesmo as periódicas desenvolvidas pelas escolas – até porque o intuito é avaliar o conhecimento matemático do aluno ou linguístico?

Gostaríamos de retomar o ideário de Perfetti e Hart (2002), os quais defendem que um termo de baixa frequência confunde e atrapalha a leitura do indivíduo. A partir disso e somado ao contexto de avaliação, que é quando os discentes já se encontram inseguros por estarem sendo avaliados, ressaltamos a importância de enunciados de problemas matemáticos serem formulados e pensados para o contexto para os quais eles servirão de modo que contribuam para o processo interpretativo e formativo do discente e não o prejudique.

Ciência para Educação é o que nos objetivamos. Queremos deixar claro que não se trata de criticar metodologias ou a prática dos professores, mas de fornecermos evidências científicas para que a educação possa ser cada vez mais baseada em estudos, em ciência. Nesse sentido, queremos fornecer evidências da análise linguística para auxiliar a elaboração de materiais mais eficazes e contextualizados. E aqui novamente surge a interdisciplinaridade, o quanto que nós linguistas podemos contribuir com os profissionais da área da matemática e vice-versa

Nós defendemos que entender a leitura a partir de seus processos cognitivos é crucial para que possamos intervir e reduzir os impactos da baixa literacia que afeta muitos alunos ao longo de sua vida escolar, muitas vezes sem que eles saibam qual é, de fato, a raiz do seu baixo desempenho acadêmico.

Há vários fatores que podem influenciar nesse panorama, vivemos em sociedade na qual as tecnologias digitais afetam todos e não seria diferente na educação. Modismos, falsos pressupostos, interesses comerciais e ideológicos rondam o ambiente escolar, assim como algumas ideias e boas intenções. O encantamento da tecnologia, por vezes, pode ser saudável. Para surpresa do mundo a Suécia, reconhecida pelo seu alto padrão educacional, – referência mundial em educação de qualidade e país pioneiro desde 1990 a implementar educação 100% digital nas escolas – voltou atrás e decidiu retomar os livros físicos, por causa da queda de rendimento escolar<sup>25</sup>, dentre vários outros motivos. Enquanto o Governo do Estado de São Paulo, em 2023, na contramão, noticiou a retirada dos livros didáticos<sup>26</sup> e voltou atrás por pressão popular, sendo obrigado a fornecer material impresso. A

---

<sup>25</sup> Notícia sobre a Suécia voltar ao uso de livro impresso: <https://g1.globo.com/educacao/noticia/2023/08/07/por-que-a-suecia-desistiu-da-educacao-100percent-digital-e-gastara-milhoes-de-euros-para-voltar-aos-livros-impresos.ghtml>

<sup>26</sup> Notícia sobre o Governo do Estado de São Paulo não aderir livros didáticos do MEC <https://g1.globo.com/sp/sao-paulo/noticia/2023/08/01/governo-de-sp-nao-adere-livros-didaticos-do-mec-para-2024-e-diz-que-usara-material-proprio-e-100percent-digital-a-partir-do-6oano.ghtml>

decisão da Suécia foi embasada em pesquisas e reflete um estado preocupado com a formação de seus cidadãos. Enquanto algumas políticas públicas adotadas no Brasil, nos três níveis (municipal, estadual e federal) são pautadas por outras motivações. Em um caso como esse relatado sobre São Paulo, fica muito difícil se pensar em promover a literacia nos estudantes.

Diante disso, a Ciência para Educação deve estar em duas frentes: para o educado diretamente e para a fundamentação das práticas educacionais. Esta pesquisa buscou, humildemente, se filiar na segunda linha: buscamos indícios científicos e reflexões – trabalhos como este, inclusive, inédito nessa temática, porque busca lançar um olhar que auxilie a promover meios mais eficazes e acessíveis aos alunos, pois eles já estão vivendo em um mundo com muitas mudanças e repleto de estímulos tecnológicos que nem sempre os auxiliam.

Há um paradoxo subjacente, já citado em outros momentos, que é a existência de palavras pouco frequentes aos alunos de determinada faixa etária em enunciados matemáticos, por outro lado, eles precisam conhecer novas palavras; desse modo é necessário que ele, em algum momento, passe a ter contato para aprendê-la e por isso não se deve evitar a frequência baixa, nem estimular um vocabulário muito novo, pois o aluno precisa ter segurança para interpretar e oportunidade para aprender. Assim, temos algumas sugestões, fruto deste percurso que podem ser práticas adotadas por docentes, editoras de livros didáticos, orientadores, planejadores educacionais etc.:

- a) Criar glossários informativos na mesma página em que aparece o termo pouco frequente aos alunos daquela faixa etária;
- b) Trabalhar esses termos em sala de aula com explicações e ilustrações, etc.;
- c) Utilizar os termos em mais de um enunciado para que o aluno não só veja uma vez, mas para que essa palavra se torne conhecida em mais de um uso.

Enfim, sabemos que há muito ainda o que ser pesquisado, analisado e discutido, o que temos feito é somente o início de uma profunda e extensa discussão. Essa área carece de pesquisas que trabalhem de forma conjunta, com especialistas e profissionais da educação, para que o olhar também não seja somente acadêmico, fora das salas de aula. Por isso, acreditamos na necessidade de mais pesquisas que

considerem também a importância prática da interdisciplinaridade, para que prossigamos avançando com a educação.

## CONSIDERAÇÕES FINAIS

Nesta tese, objetivamos investigar a frequência dos termos que compõem os enunciados das SiPs em matemática, a fim de verificar se eles são de fato os obstáculos enfrentados pelos alunos durante a leitura e interpretação. Para isso, buscamos entender, antes de tudo, como se dá a questão lexical, partindo do conceito de composicionalidade.

Entendemos, a partir da discussão proposta no capítulo 1, que a compreensão de uma sentença só acontecerá se o aluno for capaz de estabelecer relações entre o signo linguístico e seu referente, conceito trabalhado de formas diferentes pelas diferentes áreas de Frege (1978) a Charles Perfetti (2001; 2002; 2007). Todavia, essa significação não se dará sozinha, ela precisa das partes do texto para se chegar ao todo, isto é, dependerá também de outros termos ali presentes. Desse modo, o princípio da composicionalidade possibilita às crianças, através do vácuo da fórmula, preencherem de significados. Isso quer dizer que elas até podem conhecer a maior parte das palavras, entretanto uma desconhecida, mesmo que não seja um sentido-chave para a resolução, gera insegurança e faz com que a resolução do problema matemático seja comprometida.

Diante dessa problemática, recorreremos aos teóricos da aquisição lexical e de língua-materna a fim de compreendermos como lidar com essa questão. Nesse sentido, há um consenso de que a exposição da criança a diversos contextos faz com que ela adquira mais linguagem e conhecimento de mundo. A partir disso, Xatara (et al, 2008) acreditam que o léxico seja adquirido e ampliado de acordo com as relações sociais e a escola, dessa forma, é um processo completamente social e dependente de contatos e do acesso à informação. É claro que acreditamos ser importante o conhecimento de novas palavras, mas questionamos se esse tipo de contato precisaria ser enquanto essas crianças estão sendo testadas – como em avaliações periódicas nas escolas ou até mesmo internacionais como o PISA –, pois, ainda que as SiPs sejam mais voltadas ao contexto diário, são desafiadoras, utilizadas recorrentemente como exercícios em livros didáticos, inclusive o tipo de exercício que mais aparece nessas obras.

Seguindo essa lógica, concordamos com Leffa (2000) ao defender que a frequência é extremamente importante na manutenção do léxico de um indivíduo. Sendo assim, há pouca serventia o uso de um termo como “hodômetro”, mesmo que

com glossário, em um enunciado no livro todo e depois ele não ser mais citado, não poderemos dizer que a criança internalizou esse termo.

É claro que é comum em livros de matemática aparecerem termos particulares dessa área, a saber, “dobro”, “retangular”, assim espera-se mesmo que eles sejam frequentes em um corpus específico desses livros, no entanto as SiPs são formadas majoritariamente por termos que são mais conhecidos, para além do cenário matemático. Porém, ao analisar a correlação entre os corpora, especialmente focando como variante principal o SiPCorp, essas classes gramaticais que são majoritárias na composição desses enunciados não possuem uma forte correlação, pelo contrário, sua correlação é fraca. Isso mostra que nossa hipótese de que o estudo da frequência desses termos responderia se de fato é um dos entraves para a dificuldade dos alunos em interpretar situação-problema estava certa.

Como uma forma de testar essa nossa hipótese de que o léxico utilizado no livro didático de matemática não é próximo e bem pensado à faixa etária do aluno, elaboramos um teste de familiaridade – para que os participantes julgassem o quanto conhecem de cada palavra – e um teste de interpretação – a fim de averiguarmos se, para além de só verem as palavras, conseguiriam encaixá-las em contextos de enunciados de SiPs. A partir disso, notamos que realmente o vocabulário importa e pode ser sim um dificultador do processo interpretativo. Ainda que sejam pouco frequentes as palavras desconhecidas por eles, quando utilizadas, podem ser um obstáculo que nem mesmo o contexto consegue driblar, assim como vimos em alguns termos trabalhamos por nós: “gramar”, “escalímetro” e “verifique”.

Gostaríamos de acrescentar que tem sido um enorme prazer poder trabalhar com duas grandes áreas do conhecimento que se complementam e estão presentes em toda a vida escolar das crianças. Dessarte, salientamos que este trabalho não é uma crítica às metodologias adotadas em sala de aula e muito menos aos professores de matemática, mas uma análise linguística para que compreendamos os possíveis obstáculos enfrentados linguisticamente pelos alunos durante o contato com enunciados de SiPs.

## BIBLIOGRAFIA

AGOSTINI, Tainara Duro. Estrutura e interpretação de situações-problema em matemática: uma análise baseada na semântica formal. 2019.

BESTGEN, Yves. Comparing lexical bundles across corpora of different sizes: The Zipfian problem. **Journal of Quantitative Linguistics**, v. 27, n. 3, p. 272-290, 2020.

BIANCHINI, Edwaldo. Matemática Bianchini, 6º ano. 9. edição. São Paulo: Moderna, 2018.

BIANCHINI, Edwaldo. Matemática Bianchini, 7º ano. 9. edição. São Paulo: Moderna, 2018.

BIANCHINI, Edwaldo. Matemática Bianchini, 8º ano. 9. edição. São Paulo: Moderna, 2018.

BIANCHINI, Edwaldo. Matemática Bianchini, 9º ano. 9. edição. São Paulo: Moderna, 2018.

BIDERMAN, Maria Tereza C. Teoria lingüística. São Paulo: Martins Fontes, 2001.

BRYLSBAERT, Marc; DIEPENDAELE, Kevin. Dealing with zero word frequencies: A review of the existing rules of thumb and a suggestion for an evidence-based choice. **Behavior research methods**, v. 45, p. 422-430, 2013.

CANÇADO, Marcia. Introdução à semântica lexical. Petrópolis, Rio de Janeiro: Editora Vozes, 2016.

CHIERCHIA, Gennaro. Partitives, reference to kinds and semantic variation. **Semantics. Critical concepts in linguistics**, v. 3, p. 415-446, 2003.

CRISTÓFARO-SILVA, T., de Almeida, L.S., Fraga, T.: ASPA: A Formulação de um Banco de Dados de Referência da Estrutura Sonora do Português Contemporâneo. In: Proceedings XXV Congress of Brazilian Society of Computing Science, São Leopoldo, RS, Brazil (2005).

PINTO, Adilson-Luiz et al. Comparação da Lei de Zipf em conteúdos textuais e discursos orais. **Profesional de la Información**, v. 24, n. 2, p. 157-167, 2015.

DOS SANTOS ARAÚJO, Sónia Patrícia. Estratégias de aquisição e ampliação do capital lexical em Português e Espanhol. 2020.

ESTIVALET, Gustavo et al. LexPorBr Infantil: uma base lexical tripartida e com interface Web de textos ouvidos, produzidos, e lidos por crianças. In: Proceedings. 2019.

FERREIRA, Marcelo. 2019. Curso de semântica formal (Textbooks in Language Sciences 6). Berlin: Language Science Press.

FIELD, J. Psycholinguistics: The key concepts. New York: Routledge, 2004.

FREGE, Friedrich Ludwig Gottlob. *Lógica e filosofia da linguagem*. São Paulo: Cultrix, 1978.

GERNSBACHER, M. A. *Handbook of psycholinguistics*. San Diego: Academic Press, 1994.

LEFFA, V. J. Aspectos externos e internos da aquisição lexical. In: LEFFA, Vilson J. (Org.). *As palavras e sua companhia; o léxico na aprendizagem*. Pelotas, 2000, v. 1, p. 15-44.

LEITÃO, M. M. *Psicolinguística Experimental: Focalizando o processamento da linguagem*. In: MARTELOTTA, M. E. (org.) *Manual de Linguística*. São Paulo: Contexto, 2008.

LORANDI, Aline; CRUZ, Carina Rebello; SCHERER, Ana Paula Rigatti. *Aquisição da linguagem*. Verba Volant, v. 2, no 1. Pelotas: Editora e Gráfica Universitária da UFPel, 2011.

MORAIS, José (2020). Os caminhos da Literacia. *Cadernos de Linguística*, v. 1, n. 1, p. 01-14.

MORTARI, Cezar A. *Introdução à lógica*. Unesp, 2001.

OLIVEIRA, C. S. F.; SÁ, T. M. M. *Métodos experimentais em psicolinguística [recurso eletrônico]*. 1. ed. - São Paulo: Pá de Palavra, 2022.

PARTEE, Barbara H. Nominal and temporal anaphora. ***Linguistics and philosophy***, p. 243-286, 1984.

PERFETTI, Charles A.; HART, Leslie. The Lexical Bases of Comprehension Skill. In: GORFIEN, D. S. (Ed.) *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*. Washington, DC: American Psychological Association, 2001, pp. 67-86.

\_\_\_\_\_. "The Lexical Quality Hypothesis". *Precursors of functional literacy*, Verhoeven, Ludo; et al Amsterdam: John Benjamins. 2002, pp. 189-213.

PERFETTI, Charles. Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, v. 11, n. 4, 2007, pp. 357-383.

POST DA SILVEIRA, Amanda et al. What weighs for word stress? Big data mining and analyses of phonotactic distributions in Brazilian Portuguese. In: *Computational Processing of the Portuguese Language: 13th International Conference, PROPOR 2018, Canela, Brazil, September 24–26, 2018, Proceedings 13*. Springer International Publishing, 2018. p. 399-408.

SARGIANI, Renan de Almeida; MALUF, Maria Regina. Linguagem, cognição e educação infantil: contribuições da psicologia cognitiva e das neurociências. ***Psicologia Escolar e Educacional***, v. 22, p. 477-484, 2018.

SOARES, Magda. Letramento e alfabetização: as muitas facetas. **Revista brasileira de educação**, p. 5-17, 2004.

XATARA, Claudia Maria et al. A aquisição do vocabulário básico e a competência lexical. *Caderno Seminal*, v. 10, n. 10, 2008.

## ANEXO I

Estímulos utilizados nos experimentos separados por ano escolar

### 6 ano

<q=066> Sabendo que, com um trator, Lúcio **ara**  $\frac{3}{20}$  de um terreno em um dia, responda:

- De segunda-feira a sábado, que parte do terreno Lúcio consegue **arar**?
- Ele conseguirá terminar na segunda-feira? **Justifique** sua resposta.

<q=005> Na recepção de um laboratório, os pacientes preferenciais têm senha com dois **algarismos**; os pacientes agendados têm senha com três **algarismos**; e os demais têm senha com quatro **algarismos**.

- Que senha de quatro **algarismos** sucederá a do painel? Qual a antecedeu?

<q=040> Em uma receita de bolo de chocolate, são necessários  $\frac{3}{3}$  copos de leite. Sabendo que em um copo cabem 200 mililitros, **determine** quantos mililitros de leite serão necessários para essa receita.

<q=067> Isabel dividiu sua horta retangular em 3 canteiros iguais. Em um desses canteiros, plantou couve em uma metade e, na outra, espinafre. Agora, responda:

- **Represente** por meio de uma figura e com uma fração a parte da horta onde foi plantado espinafre.

<q=014> Tatiana jogou dois dados, obtendo uma **soma** de 9 pontos. Quais são os possíveis pares de números para que ocorra essa **soma**?

### 7 ano

<q=0148> Manuela e Luciano resolveram a operação  $3 + 0,5$  de formas diferentes. Reúna-se com um colega para analisar as duas resoluções e respondam à questão: Manuela e Luciano encontraram o mesmo resultado? Se não, o que aconteceu? **Justifiquem** a resposta.

<q=0136> Em certo país, as eleições para presidente ocorrem a cada 4 anos, e para senador, a cada 8 anos. Em 2014, essas eleições coincidiram. **Determine** os anos das quatro próximas vezes em que as eleições voltarão a coincidir.

<q=248> <vol=03> <cap=04> Orlando, dono de uma quitanda, colocou 5 maçãs em um dos pratos de uma balança e nivelou-a colocando no outro prato um peso de 500 g e dois pesos de 200 g. Indicando a massa média de cada maçã por x, **represente** essa situação por meio de uma equação.

<q=0205> Com um **escalímetro**, a planta de uma casa foi desenhada na escala 125. Nessas condições, responda as questões.

- Qual é o **comprimento** real, em metro, de uma sala que, nessa planta, foi representada por um segmento de 5,2 cm?

<q=0200> Uma latinha de refrigerante tem capacidade de 350 ml, e uma garrafa 2 litros. Determine a **razão** entre a capacidade dessa latinha e a capacidade dessa garrafa.

## 8 ano

<q=267> Ana e Renata moram perto de uma lanchonete, conforme mostra o esquema abaixo. Qual delas mora mais distante da lanchonete? **Justifique**.

<q=252> Uma empresa de software lançou um novo programa no mercado. No primeiro mês, ela vendeu determinada quantidade desse novo programa. No segundo mês, foi vendido o dobro do que se vendeu no primeiro mês. No terceiro mês, foi vendido o triplo do que se vendeu nos meses anteriores. **Represente** a quantidade de unidades vendidas nos três primeiros meses.

<q=0221> A tabela abaixo mostra o tempo, em hora, que os meninos e as meninas do oitavo ano do colégio Alegria acessam a internet semanalmente. Tempo semanal de acesso à internet Tempo Classe (de a até -. Frequência absoluta: meninos. Frequência absoluta: meninas.

- Construa a tabela de frequências relativas na forma percentual e **verifique** se sua resposta ao item a está correta.

<q=283> Um jardineiro precisa **gramar** um terreno que tem o formato da figura abaixo. Quantos metros quadrados de grama serão necessários para cobrir esse terreno?

<q=262> Ontem, eu tinha em minha conta poupança certa quantia em reais. Hoje, meu avô depositou nessa conta o quadruplo do que eu tinha ontem mais R\$ 200,00. Represente por uma fração algébrica a **razão** entre o saldo que minha conta poupança tinha ontem e o saldo depois do depósito feito por meu avô.

## 9 ano

<q=366> Num triângulo isósceles, cada lado congruente mede 15 cm. Determine a área desse triângulo, sabendo que sua base mede 24 cm. 7 É possível colocar um lápis de 18 cm num estojo retangular de 12 cm por 15 cm? **Justifique** sua resposta.

<q=268> Em um jogo de futebol, Paulo cobra uma falta jogando a bola para Marcos, que a joga para Daniel. A trajetória da bola está representada na figura abaixo. **Determine** as medidas dos ângulos assinalados em vermelho na figura.

<q=387> Um avião de acrobacias levanta voo formando um ângulo de  $50^\circ$  em relação à pista. **Calcule** a que altura o avião estará do solo após percorrer 3,5 km. (Dado:  $\sin 50^\circ = 0,76$ .)

<q=372> Um gavião, a 700 m de altura, avista uma presa; faz uma descida de  $17^\circ$  em relação à horizontal e consegue capturá-la. Que distância o gavião **percorreu** para capturar essa presa?

<q=354> <vol=04> <cap=07> Daqui a 6 anos, a idade de Daniela será igual ao quadrado da idade dela há 6 anos. Indique a idade atual de Daniela por x para resolver as questões que se seguem.

- Que equação **traduz** a situação do problema?

## ANEXO II

### **Roteiro para o protocolo de aplicação do teste da pesquisa “Efeitos da frequência lexical na interpretação de situações-problema”**

Os testes serão realizados com crianças do sexto ao nono ano do Ensino Fundamental II da escola \_\_\_\_\_. A primeira parte do experimento consistirá em um questionário de palavras que serão avaliadas de acordo com a familiaridade de cada um. O segundo momento consistirá em também um questionário, porém com enunciados de situação problema com lacunas as quais serão preenchidas com o termo que o aluno julgar importante. Seguem o roteiro com os procedimentos e as questões a serem aplicadas:

1) Apresentação em linguagem simplificada para o(a) participante das justificativas, objetivos e os procedimentos que serão utilizados durante a pesquisa. Expor os possíveis riscos e benefícios de sua participação;

2. Lembrar que houve consentimento do responsável para a realização do teste, avisar que a identidade dele(a) será mantida em segredo e que seu nome não constará em nenhum relatório;

3. Deixar claro que todo procedimento será gravado como forma de manter a integridade dos fatos;

4. Iniciada a gravação, os procedimentos serão os seguintes:

1. Informação do nome e da idade do sujeito;

2. Comunicar que o teste contém seis problemas de matemática sem uma das palavras que ele deverá eleger a mais adequada, segundo seu conhecimento, dentre três opções;

3. Comandos a serem realizados durante o teste:

5. Ao longo do processo, será lembrado que o sujeito tem total liberdade para recusar-se a continuar participando do teste, em qualquer fase da pesquisa.

É importante frisar que todas as etapas do teste serão realizadas da forma mais descontraída possível e com um palavreado simplificado, de conhecimento e utilização dos alunos, para proporcionar um relaxamento e uma maior confiança dele(a) na pesquisadora, podendo transmitir, assim, de uma forma mais segura, os dados necessários para a pesquisa.

## ANEXO III

### Termo de consentimento livre e esclarecido (responsáveis dos alunos)

#### PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

Doutoranda: Tainara Duro Agostini

Orientadores: Dirceu Cleber Conde e Amanda Post da Silveira

#### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

O/A menor \_\_\_\_\_, sob sua responsabilidade, está sendo convidado (a) a participar como voluntário da pesquisa *Efeitos da frequência lexical na interpretação de situação-problema*. Nesta pesquisa, pretendemos investigar se os termos que compõem enunciados de situação-problema são adequados à faixa etária e se eles interferem no processo interpretativo. O motivo que nos leva a estudar essa questão é verificar de que forma o processamento linguístico pode influenciar no raciocínio matemático.

Para esta pesquisa, adotaremos os seguintes procedimentos: o jovem, primeiro, responderá a um questionário de familiaridade no qual constará uma lista de palavras sobre as quais ele irá falar o grau de uso e contato. Depois, ele responderá outro questionário, porém com enunciados de situação-problema os quais terão lacunas para serem preenchidas por palavras julgadas pelo aluno como adequada ao contexto. As tarefas estão diretamente relacionadas aos conteúdos pedagógicos previstos no currículo escolar e serão realizadas durante o período escolar, sob a orientação da professora Tainara Duro Agostini.

Para participar deste estudo, o menor sob sua responsabilidade não terá nenhum custo, nem receberá qualquer vantagem financeira. Ele (a) será esclarecido (a) sobre o estudo em qualquer aspecto que desejar e estará livre para participar ou recusar-se a participar. Você, como responsável pelo menor, poderá retirar seu consentimento ou interromper a participação dele (a) a qualquer momento.

Os dados e instrumentos utilizados na pesquisa ficarão arquivados com o pesquisador responsável por um período de 5 (cinco) anos, e, após esse período, serão destruídos. Os pesquisadores tratarão a sua identidade com padrões profissionais de sigilo, atendendo à legislação brasileira (Resolução Nº 466/12 do Conselho Nacional de Saúde), utilizando as informações somente para os fins acadêmicos e científicos.

Eu, \_\_\_\_\_, portador (a) do documento de identidade \_\_\_\_\_, responsável pelo (a) menor \_\_\_\_\_ fui informado dos objetivos da pesquisa *Efeitos da frequência lexical na interpretação de situação-problema* de maneira clara e esclareci todas as minhas dúvidas. Recebi uma via original deste termo de consentimento livre e esclarecido e estou ciente de que a qualquer momento poderei solicitar novas informações e também retirar o meu consentimento se assim desejar. Declaro que concordo em participar. Recebi uma via original deste termo de consentimento livre e esclarecido e me foi dada a oportunidade de ler e esclarecer as minhas dúvidas.

São Carlos, \_\_\_\_\_ de \_\_\_\_\_ de 2023.

Assinatura do (a) responsável: \_\_\_\_\_

Assinatura do (a) pesquisador: \_\_\_\_\_

Em caso de dúvidas, com respeito aos aspectos éticos desta pesquisa, você poderá consultar os responsáveis pela pesquisa – Tainara Duro Agostini (doutoranda) e Dirceu Cleber Conde (orientador) – pelos emails [tainara.agostini@gmail.com](mailto:tainara.agostini@gmail.com) e [cleberconde@ufscar.br](mailto:cleberconde@ufscar.br).

## ANEXO IV

### TERMO DE ASSENTIMENTO INFORMADO

Este formulário de assentimento informado é para jovens de 10 a 15 anos que estamos convidando a participar da pesquisa.

Meu nome é Tainara Agostini e, no meu curso de doutorado, estou realizando uma pesquisa sobre o porquê de você sentir dificuldade em interpretar situações-problema na aula de matemática. Eu gostaria de convidá-lo a participar desta pesquisa. Nessa pesquisa, você vai ler algumas palavras e frases e, se for necessário, me explicar o que entendeu ou se conhece as palavras que eu lhe mostrar. Você pode escolher se quer participar ou não. Já conversei com sua família sobre esse trabalho. Eles sabem que você está sendo convidado e estão de acordo com sua participação. Mas se você não desejar fazer parte da pesquisa, não é obrigado, até mesmo se seus pais concordarem. É você quem decide.

#### **Certificado do Assentimento**

Eu entendi que a pesquisa é sobre a forma como as crianças interpretam algumas frases e que, dependendo da tarefa, eu poderei escolher algumas palavras ou até mesmo dizer minha opinião sobre algumas frases que eu vou ler ou ouvir.

Nome do participante: \_\_\_\_\_

Assinatura do participante: \_\_\_\_\_

Assinatura dos pais/responsáveis: \_\_\_\_\_

Ass. Pesquisador: \_\_\_\_\_

Dia/mês/ano: \_\_\_\_\_