

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
FACULDADE DE LETRAS – FL
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

OLÍVIA FERNANDES BOGO

DIFERENÇAS NA AQUISIÇÃO E PROCESSAMENTO DE NOMES CONCRETOS E
ABSTRATOS POR CRIANÇAS BRASILEIRAS DE 2 ANOS

RIO DE JANEIRO

2025

Olivia Fernandes Bogo

DIFERENÇAS NA AQUISIÇÃO E PROCESSAMENTO DE NOMES CONCRETOS E
ABSTRATOS POR CRIANÇAS BRASILEIRAS DE 2 ANOS

Tese de doutorado apresentada ao Programa de
Pós-graduação em Linguística da Universidade
Federal do Rio de Janeiro como requisito à
obtenção do título de Doutora em Linguística

Orientadora: Profa. Dra. Aniela Imbrota
França

Rio de Janeiro

2025

FERNANDES BOGO, Olívia. Diferenças na aquisição e processamento de nomes concretos e abstratos por crianças brasileiras de 2 anos. **Universidade Federal do Rio de Janeiro, 2025. Tese de doutorado em Linguística.**

Tese de Doutorado submetida ao Programa de Pós-Graduação em Linguística, Faculdade de Letras, da Universidade Federal do Rio de Janeiro – UFRJ, como requisito necessários à obtenção do título de Doutora em Linguística.

Banca Examinadora:

Professora Doutora Maria Cristina Lobo Name – UFJF

Professora Doutora Marília Uchoa Cavalcanti Lott de Moraes Costa – UFRJ

Professor Doutor Márcio Martins Leitão – UFPB

Professora Doutora Janaína Weissheimer – UFRN

Professora Doutora Aleria Cavalcante Lage, Suplente.

Professora Doutora Juliana Novo Gomes, Suplente.

Rio de Janeiro

2025

CIP - Catalogação na Publicação

F665d Fernandes Bogo , Olívia
Diferenças na aquisição e processamento de nomes
concretos e abstratos por crianças brasileiras de 2
anos. / Olívia Fernandes Bogo . -- Rio de Janeiro,
2025.
192 f.

Orientadora: Anielá Improta França.
Tese (doutorado) - Universidade Federal do Rio
de Janeiro, Faculdade de Letras, Programa de Pós
Graduação em Linguística, 2025.

1. aquisição de linguagem. 2. concretude. 3.
abstração. 4. rastreamento ocular. 5.
psicolinguística. I. Improta França, Anielá, orient.
II. Título.

O presente trabalho foi realizado no Laboratório ACESIN (UFRJ), sob a orientação da Professora Aniela Imbrota França, com apoio da Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior – Brasil (CAPES) e da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro, na modalidade Doutorado nota 10, e, em conjunto, no Laboratório LaPsyDé (Université de Paris) com apoio da Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior – Brasil (CAPES), no âmbito do Programa Institucional de Internacionalização (PrInt) – código de financiamento 0001, sob supervisão na França do prof. Dr. Alex de Carvalho (Université Paris-Cité).

Às crianças e aos seus familiares, bem como aos demais participantes, que, ao acreditarem na ciência, tornaram esta pesquisa possível.

AGRADECIMENTOS

Agradecer por cada encontro e por tantas partilhas nestes quatro anos de doutorado é pensar em quantas vidas e histórias perpassam a minha para que eu pudesse fazer este caminho formativo tão longo e realizar este sonho.

À pioneira nos estudos em neurociência da linguagem no Brasil, Anieli França, minha orientadora, pela coragem de ter começado e pela persistência diante de todos os desafios que surgiram no seu percurso para que nós, seus alunos, também pudéssemos realizar os nossos sonhos e nos tornar pesquisadores; pela forma curiosa e apaixonada de fazer ciência, pela sua humildade em compartilhar tanto do seu conhecimento comigo e pela forma humana e sensível com que sempre me tratou. É uma honra e alegria tê-la como orientadora.

Ao Alex de Carvalho, meu supervisor no período de sanduíche na França, pelo interesse na pesquisa que estávamos desenvolvendo e por aceitar me receber no Laboratoire de Psychologie du Développement et de l'Éducation de l'Enfant (LaPsyDé), da Université Paris Cité. Agradeço também pela imensa contribuição na parte experimental, pelas orientações brilhantes, das quais eu saía maravilhada com seu raciocínio científico e inspirada a realizar as minhas tarefas, pela paciência diante das minhas dificuldades (sobretudo com a plataforma experimental) e pela confiança de que eu seria capaz de superá-las e de que teríamos um bom resultado.

Ao Grégoire Borst, diretor do LaPsyDé, em Paris, e aos demais membros, por me receberem tão bem em seu laboratório, por seis meses. Foi uma experiência incrível na Cidade Luz, em que tentei viver cada oportunidade do que é fazer pesquisa experimental em psicolinguística e neurociência na Europa. Participei de um congresso na sede da UNESCO, observei estudos com Ressonância Magnética Funcional no Institut de Psychiatrie et Neurosciences de Paris, no hospital Sainte-Anne, e acompanhei experimentos com rastreamento ocular no Baby Lab da École Normale Supérieure e com adultos no próprio LaPsyDé. Inspirada por grandes pensadores iluministas, como Voltaire, Rousseau e Marie Curie, fiz meu caminho para o LaPsyDé, perto do Panthéon, muitas vezes emocionada por viver esta cooperação acadêmica. C'était fantastique! Je ne l'oublierai jamais!

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); à Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), pela bolsa doutorado nota 10 - e ao programa Capes PrInt, agências de fomento das quais usufruí de bolsa durante o meu tempo de doutoramento e sem as quais seria impossível me dedicar às atividades de pesquisa, sobretudo durante o meu período de sanduíche na França.

Às professoras Cristina Name, Marília Uchôa e ao professor Márcio Leitão, por aceitarem compor a banca de defesa e pelas contribuições valiosas que trouxeram ao texto. À professora Janaína Weissheimer, não apenas pela participação na defesa, que enriqueceu o nosso trabalho, mas também pelo seu comprometimento social, tão humanamente inspirador, em seus projetos de pesquisa.

Aos meus queridos do Laboratório de Acesso Sintático (ACESIN-LAB), da UFRJ, em especial à Ste Martin, amiga querida, que me ouviu chorar por telefone, apesar da distância Rio–São Paulo, às vésperas da sua defesa, dedicando-me seu tempo, sua escuta e seu cuidado em diversos momentos ao longo do doutorado, à Renata Libotti e ao Victor Ramos. Agradeço imensamente à professora Marije Soto, ao Pedro Henrique Santos, à Júlia Cataldo, ao Leonardo Cabral, à Lara Romão, à Larissa Lilge e à Rayssa Pires por compartilharem toda a alegria e também o inerente desespero de todo acadêmico, desfrutado com muito café, em todos os nossos encontros. Agradeço especialmente à Maria Maurell, já tão capaz e inspiradora, ainda na Iniciação Científica, pelo brilho nos olhos e pelo entusiasmo com que vinha até mim falar da sua pesquisa, quando eu estava super cansada e desmotivada com a minha.

À professora Natália Mota, pelas trocas inspiradoras e pelo seu comprometimento social em realizar pesquisas que transformam a realidade, sobretudo a do nosso país. Ao professor Roberto Lent, pela honra de ter sido sua aluna na disciplina de Neurociência Translacional e pelo quanto essa experiência transformou a minha visão de fazer ciência.

Aos queridos pesquisadores e amigos que conheci na UFRJ, em congressos e em eventos científicos, e pelos quais tenho imensa admiração, em especial Moíra Souza, Frederico Ludugerio, Tânia Magalhães, Ludmilla Chaves, Mariana Valença, Luan Mattos, Igor Reis, Rômulo Weckmuller e Regina Barragat, que tanto se interessaram pela pesquisa, trouxeram palavras de incentivo e contribuíram para a divulgação dos pré-testes e/ou dos experimentos, especialmente o de bebês.

Aos queridos do Laboratório de Aquisição, Processamento e Sintaxe (LAPROS), da Unicamp, Fernando Sabatini, Alice Fernandes e Thuany Figueiredo pelas discussões sobre linguística, cognição, aquisição de linguagem, concretude e abstração, emocionalidade e comitês de ética.

Aos coordenadores Gean Damulakis e Karen Sampaio, bem como às secretárias Patrícia Pereira e Carla Santino, do Programa de Pós-Graduação em Linguística da UFRJ, por todo o auxílio prestado ao longo do doutorado.

Às instituições de ensino que abriram suas portas, acreditaram na ciência e nos permitiram apresentar a pesquisa aos responsáveis, recrutar participantes e aplicar o experimento infantil, em especial à Área Bambini; à Associação Espaço Educativo São Charbel; o Centro de Educação Infantil – Profa. Hermínia Matheus; ao Centro de Educação Infantil – Santo Agostinho; ao Centro de Educação Infantil – Tia Alice; ao Centro de Recuperação Infantil, em Ipanema, no Rio; ao Colégio de Aplicação da Universidade Católica de Petrópolis; à Escola Lápis de Cor; ao Lar de Emmanuel; e ao Lar Nossa Senhora das Graças. À clínica médica Docctor Med, em Petrópolis, pelo auxílio na divulgação do experimento infantil junto à pediatria e por ampliar o convite do experimento controle, com adultos, entre seus colaboradores. Reforço meus agradecimentos a essas instituições pela cooperação tão indispensável para a promoção da ciência e pela recepção carinhosa.

Aos amores da minha vida – e de uma vida inteira – Marlene e Patrícia Amorim, Diego Villa, Karlen Ullmann, Sara Gattis, Thaís Paiva (e Nina – nosso mais novo amor), Jayna Cosmo, Gisele Caneschi, Tadeu Furtado, Tatiana e Ivânia Vasconcelos, Maria Mariosa, Marcella Serôa, Ana Luísa da Costa, Késsia Henrique, Bárbara e Evelyn Delgado, Juliana Fernandes, Isabella Pederneira, Ludmilla Fernandes, Bárbara Ramos, Pedro e Janaína Cunha, Inês Damasceno, Maria Eduarda Costa, Rafaela Medeiros, Samarane Barros, Suellen Lima e sua família, dedico meus agradecimentos mais amorosos. Como escreveu Clarice Lispector: “Dar a mão a alguém sempre foi o que esperei da alegria”, agradeço por *todos* os momentos em que tive a mão de vocês por perto para *nunca* esquecer o que é ter alegria.

À minha avó, Nelma, e às minhas irmãs, Mima e Vic, pela alegria de compartilhar meus sonhos e realizações e, acima de tudo, por fazerem valer, em nossa história, o que Freud escreveu: “Juntamente com as exigências da vida, é o amor o que mais educa”. Agradeço por serem razões para eu gostar da vida e por trazerem ao mundo as crianças mais lindas, Luiza, Noah, Stella, Sienna e Louis, que me concederam o título de tia.

À Cristina e Isabelle Azalim, à Silvia Nogino, à Leana Villa, ao Pedro Araújo, à Claudia Schittini, à Renata Silva, à Andreia Garcia, à Fabiana Beutler, à Azussa Matsuoka e família, à Eliana Ramos e família, ao Ulisses Doors, à Thaísa Souza, à Natália Bibikow, à Maria Galli, à Maria Andrea Stehling, ao Leandro Nascimento, ao Victor Rosado, à Letícia e Lorraine Ullmann, ao Eduardo Birman, ao Amilcare Zanoni, à Daniela e Bruno Hafner, pela amizade e pela torcida constante pela minha realização.

Às minhas maiores referências na música, João Batista Ramos, meu professor de uma vida inteira e João Vicente Corrêa, meu parceiro de cantorias e de palco, os quais perdi nos dois últimos anos. Deixo meu agradecimento, em memória, pelo grande privilégio que foi

fazer arte e ser sensibilizada pelo talento dos dois. A falta *imensa* que sinto de ambos é proporcional à alegria que me trouxeram. Levo-os nas minhas melhores lembranças e sou grata ao tempo que tivemos para fazermos o que mais amávamos: música.

À minha gatinha Luna, tão linda e amável, pela presença constante em suas loongas tardes de soneca enquanto eu escrevia esta tese, e por me lembrar da importância das pequenas pausas, como quando eu interrompia a escrita para fazer carinho em sua barriga gigante e me distraía com o seu ronronar.

À minha psicóloga, Caroline Basílio, por todo o acolhimento psicoterapêutico.

A todos que tornaram este trabalho possível, cuja contribuição coletiva perpassa cada etapa desta pesquisa, deixo meu agradecimento mais sincero. O envolvimento direto de mais de 650 participantes, desde os pré-testes até os experimentos realizados, assim como o apoio na divulgação científica e no contato com famílias e instituições infantis, foi imprescindível para a realização desta tese.

Foi um percurso bonito e difícil, mas extremamente instigante e gratificante. Por isso também que, como escreveu Clarice Lispector, “enquanto eu tiver perguntas e não houver resposta, continuarei a escrever”.

Viva à ciência — à minha mente curiosa e inquieta — e aos encontros inspiradores!

Todas as grandes coisas são difíceis e raras
(Spinoza)

RESUMO

Os processos cognitivos são, em essência, abstratos e emergem da interação entre fatores biológicos, psicológicos e ambientais. Esta tese investiga como esses processos se manifestam na aquisição da linguagem, do nível perceptual inicial à formação de representações conceituais. Partimos da hipótese de que o acesso a representações mentais é intrinsecamente abstrato, independentemente da natureza do conceito, e examinamos como isso se expressa na aquisição de palavras concretas e abstratas. Para selecionar os estímulos, realizamos dois pré-testes via Formulários Google: (i) uma readaptação do CDI do português brasileiro, respondida por 56 responsáveis de crianças entre 24 e 30 meses (compreensão e produção lexical), e (ii) a classificação de 445 adultos falantes nativos quanto à concretude/abstração das palavras. Em seguida, conduzimos um experimento-controle com rastreamento ocular on-line, por meio do *Lookit*, uma plataforma do *Massachusetts Institute of Technology*, com 53 adultos e um experimento infantil com 52 crianças (24–30 meses). Testamos: (i) diferenças no acesso e processamento de itens concretos versus abstratos; (ii) interesse relativo por cada tipo de item; e (iii) a acurácia do acesso lexical. Os resultados indicam que adultos e crianças acessam e interpretam palavras concretas e abstratas de modo distinto, sugerindo que a natureza conceitual dos itens modula a aquisição, o acesso e o processamento linguístico. Além de destacar o caráter transdisciplinar do estudo e a relevância para o contexto brasileiro, o trabalho introduz o uso inovador da plataforma *Lookit* e fornece parâmetros comparativos úteis para investigações com populações sob suspeita de transtornos do neurodesenvolvimento, subsidiando o desenvolvimento de instrumentos diagnósticos e intervenções precoces em contextos clínicos e educacionais.

Palavras-chave: aquisição de linguagem; nome concreto; nome abstrato; rastreamento ocular; acesso lexical.

ABSTRACT

Cognitive processes are fundamentally abstract and arise from the interplay of biological, psychological, and environmental factors. This dissertation examines how such processes unfold in language acquisition, from early perceptual stages to the formation of conceptual representations. We start from the hypothesis that access to mental representations is inherently abstract, regardless of conceptual type, and investigate how this is reflected in the acquisition of concrete versus abstract words. Stimulus selection involved two preregistration steps via Google Forms: (i) a Brazilian Portuguese CDI readaptation completed by 56 caregivers of children aged 24 to 30 months (lexical comprehension and production), and (ii) concreteness/abstractness ratings provided by 445 adult native speakers. We then conducted an online eye-tracking control experiment, via an MIT platform named Lookit, with 53 adults and a child experiment with 52 children (24–30 months). We tested (i) differences in access and processing of concrete versus abstract items, (ii) relative interest in each item type, and (iii) the accuracy of lexical access. Findings show that adults and children access and interpret concrete and abstract words differently, indicating that conceptual nature modulates acquisition, access, and linguistic processing. Beyond underscoring the study's transdisciplinary scope and its relevance to the Brazilian context, this work introduces the innovative use of Lookit and offers comparative parameters for research involving populations at risk for neurodevelopmental disorders, informing the design of diagnostic tools and early interventions across clinical and educational settings.

Keywords: language acquisition; concrete nouns; abstract nouns; eye tracking; lexical access.

LISTA DE FIGURAS

Figura 1 — <i>Attention grabber</i>	107
Figura 2 — Exemplo da primeira tela do <i>trial</i> de treinamento	107
Figura 3 — Exemplo da tela intermediária dos trials de treinamento	108
Figura 4 — Exemplo da segunda tela do <i>trial</i> de treinamento	108
Figura 5 — <i>Trial</i> de familiarização	109
Figura 6 — Cronologia do experimento	109
Figura 7 — Tela Inicial do ELAN — Experimento controle adulto	110
Figura 8 — Tela de apresentação do experimento infantil no <i>Lookit</i>	112
Figura 9 — Tela Inicial do ELAN — Experimento infantil	117

LISTA DE TABELAS

Tabela 1 — Resultados dos CDIs adaptados	94
Tabela 2 — Comparação dos itens lexicais selecionados para o teste psicolinguístico, extraídos da base do Lexporbr-Infantil de Estivalet (2017)	95
Tabela 3 — Distribuição dos participantes entre as listas do experimento adulto	101
Tabela 4 — Distribuição dos participantes entre as listas do experimento infantil	113

LISTA DE QUADROS

Quadro 1 — Quadro síntese dos estudos selecionados na revisão sistemática

71

LISTA DE GRÁFICOS

Gráfico 1 — Respostas ao questionário — <i>Quais palavras sua criança já aprendeu?</i>	85
Gráfico 2 — Respostas ao questionário — <i>Grau de escolaridade — responsável 01</i>	86
Gráfico 3 — Nível de escolaridade do outro responsável	87
Gráfico 4 — Respostas ao questionário — <i>Concreta ou abstrata: como você classifica essas palavras?</i>	89
Gráfico 5 — Nível de escolaridade dos respondentes ao questionário — <i>Concreta ou abstrata: como você classifica essa palavra?</i>	90
Gráfico 6 — Resultados do CDI administrado com 56 pais de crianças entre 24 e 30 meses de idade	92
Gráfico 7 — Categorização de palavras selecionadas em relação à concretude pelos adultos, falantes do PB	93
Gráfico 8 — Proporção média de olhares em cada condição (concreta e abstrata) — experimento adulto	118
Gráfico 9 — Diferenças temporais nos padrões de fixação ocular entre a condição concreta e a abstrata (<i>Cluster Permutation</i>) — experimento adulto	120
Gráfico 10 — Trajetória temporal da proporção de olhares por condição (concreta e abstrata) — experimento adulto	121
Gráfico 11 — Proporção média de olhares em cada condição (concreta e abstrata) — experimento infantil	123
Gráfico 12 — Diferenças temporais nos padrões de fixação ocular entre a condição concreta e a abstrata (<i>Cluster Permutation</i>) — experimento infantil	124
Gráfico 13 — Trajetória temporal da proporção de olhares por condição (concreta e abstrata) — experimento infantil	125

LISTA DE ABREVIATURAS E SIGLAS

AEC	Antes da Era Comum
ANOVA	Análise da Variância
APT	Teoria da Percepção de Ação (<i>Action Perception Theory</i>)
CBMM	<i>Center for Brains, Minds and Machines</i>
EFL	<i>English as a Foreign Language</i>
ERPs	Potenciais Relacionados a Eventos (<i>Event-Related Potentials</i>)
fMRI	Ressonância Magnética Funcional (<i>Functional Magnetic Resonance Imaging</i>)
HASP	<i>High Amplitude Sucking Paradigm</i>
IQR	Intervalo Interquartil
JSON	<i>JavaScript Object Notation</i>
LTA	Lobo Temporal Anterior
MB-CDI	<i>MacArthur-Bates Communicative Development Inventory</i>
MIT	<i>Massachusetts Institute of Technology</i>
PB	Português Brasileiro
PE	Português Europeu
PNAD	Pesquisa Nacional por Amostra de Domicílios
TCLE	Termo de Consentimento Livre e Esclarecido
TEA	Transtorno do Espectro Autista
TMS	Estimulação Magnética Transcraniana

SUMÁRIO

1	INTRODUÇÃO	19
2	ONTOLOGIA DOS CONCEITOS: UMA ANÁLISE HISTÓRICA SOBRE CONCRETUDE E ABSTRAÇÃO	30
3	A SENSIBILIDADE PRECOCE AOS ITENS FUNCIONAIS: EVIDÊNCIAS PSICOLINGUÍSTICAS	42
3.1	A AQUISIÇÃO LEXICAL DE ITENS CONCRETOS E ABSTRATOS	51
4	REVISÃO SISTEMÁTICA DA LITERATURA SOBRE A AQUISIÇÃO DE ITENS LEXICAIS CONCRETOS E ABSTRATOS	62
5	A AQUISIÇÃO DE NOMES CONCRETOS E ABSTRATOS POR BEBÊS BRASILEIROS: OS PRÉ-TESTES PSICOLINGUÍSTICOS	80
5.1	PRÉ-TESTE 01: <i>QUAIS PALAVRAS SUA CRIANÇA JÁ APRENDEU?</i>	82
5.1.2	PRÉ-TESTE 02: <i>CONCRETA OU ABSTRATA: COMO VOCÊ CLASSIFICA ESSAS PALAVRAS?</i>	87
5.1.3	CRUZAMENTO DAS RESPOSTAS DOS PRÉ-TESTES PARA A SELEÇÃO DOS ITENS LEXICAIS DO EXPERIMENTO	90
5.1.4	SELEÇÃO LEXICAL PARA OS ESTÍMULOS EXPERIMENTAIS A PARTIR DA ANÁLISE DOS DOIS FORMULÁRIOS	91
5.2	O EXPERIMENTO CONTROLE: <i>COMO ADULTOS BRASILEIROS PROCESSAM PALAVRAS CONCRETAS E ABSTRATAS DO PORTUGUÊS BRASILEIRO?</i>	97
5.2.1	ESTÍMULOS	99
5.2.2	PARTICIPANTES	100
	OBJETIVOS	102
	HIPÓTESES	102
	PREDIÇÃO	102
	MÉTODO	102
	MATERIAIS	104
	PROCEDIMENTOS	105

5.3	COMO CRIANÇAS BRASILEIRAS INTERPRETAM PALAVRAS CONCRETAS E ABSTRATAS DO PORTUGUÊS BRASILEIRO?	111
5.3.1	ESTÍMULOS	112
5.3.2	PARTICIPANTES	113
	OBJETIVOS	114
	HIPÓTESES	115
	PREDIÇÃO	115
	MÉTODO	116
	MATERIAIS	116
	PROCEDIMENTOS	116
6	RESULTADOS	118
7	DISCUSSÃO DOS RESULTADOS	130
8	CONSIDERAÇÕES FINAIS	135
8.1	ENCAMINHAMENTOS FUTUROS	136

1 INTRODUÇÃO

Um dos grandes desafios nos estudos sobre aquisição da linguagem é entender como bebês, mesmo expostos a um ambiente ruidoso, são capazes de adquirir sua língua materna, categorizando o mundo por meio de palavras e atribuindo significado a elas de forma tão rápida e instintiva. Ainda que estejam imersos em um ambiente altamente variável e repleto de estímulos concorrentes, bebês possuem uma notável capacidade de segmentar o fluxo contínuo de fala, identificar palavras e atribuir significados a elas.

Essas habilidades, que ocorrem de maneira aparentemente espontânea e sem instrução explícita, levantam questões fundamentais sobre os mecanismos cognitivos subjacentes à aquisição lexical. Esse processo decorre de fatores como a saliência perceptual do item e etapas do desenvolvimento neural, moldado por restrições cognitivas da espécie (Hensch, 2004; Werker; Hensch, 2015).

Desde a percepção das ondas sonoras, que se estruturam como padrões codificados em sinais eletroquímicos no cérebro, até sua estabilização na memória como representações linguísticas, todo o processo de aquisição linguística é inerentemente abstrato (Bedny; Saxe, 2012), independentemente da categoria do item lexical. Ainda que a representação semântica do conhecimento na mente/cérebro humano possa ser mapeada em termos neurofisiológicos por meio de exames de neuroimagem, investigar o processo cognitivo para determinar como esses conceitos são acessados e quando são adquiridos é ainda mais complexo (Brennan, 2022; Patterson; Nestor; Rogers, 2007; Willits *et al.*, 2013).

Visto que somente alguns itens podem ser imediatamente referenciados no mundo, como o bebê adquire conceitos ou sentimentos que não denotam objetos físicos? Como são adquiridas as palavras que exprimem reações fisiológicas como *medo*, *frio*, *dor* e outros itens vocabulares que expressam sentimentos e qualidades como *amor*, *lindo*, *saudável*? Como será que se dá a aquisição vocabular de tantas palavras que fazem parte do dia a dia dos bebês, mas não estabelecem uma correspondência direta entre palavra e objeto perceptível?

A distinção entre nomes mais concretos e abstratos é uma das discussões centrais e mais controversas nos estudos da semântica, isso porque nomes concretos se referem a entidades físicas que podem ser percebidas pelos sentidos, i.e., objetos que podemos ver, tocar ou ouvir, enquanto nomes abstratos codificam conceitos intangíveis, como as emoções (amor, tristeza, alegria, medo) e ideologias (democracia, liberdade), dentre outros (Brennan, 2022; Crutch *et al.*, 2009; Lupyan; Winter, 2018; Pulvermüller, 2013; Zdrazilova *et al.*, 2018;).

Qual seria, então, o *status* ontológico dos conceitos, que podem referenciar processos mentais, físicos, sociais ou imaginários? Conceitos como *bolo* ou *cadeira*, por exemplo, possuem um *status* ontológico concreto e objetivo, porque referenciam entidades extra-mentais, visualmente observáveis, passíveis de representação imediata e compartilháveis no mundo físico. Já palavras como *fome* ou *alegria* designam estados subjetivos, internos à experiência de um indivíduo, enquanto conceitos como *ideologia* ou *liberdade* envolvem construções sociais e abstratas.

Diante disso, qual seria o mecanismo que as crianças utilizam para aprender o sentido das palavras que elas escutam? Seria possível transpor essa lógica da materialidade versus imaterialidade para os conceitos? Seria mais fácil para um bebê mapear um objeto concreto como *mamadeira* e atribuir significado a ele, em comparação a uma reação fisiológica como *fome*? Haveria alguma contraparte física daquilo que é abstrato?

Todos esses conceitos podem ser acessados mentalmente independentemente dos objetos que referenciam, mas seus diferentes *status* ontológicos podem influenciar diretamente no modo como são adquiridos e compreendidos ao longo do desenvolvimento cognitivo (Margolis; Laurence, 1999). A dicotomia sobre a realidade mental de referentes físicos e não-físicos e a estabilização mental desses conceitos, durante a aquisição lexical, são algumas questões que motivaram esta pesquisa de doutorado.

À luz desse debate, no que tange propriamente à aquisição da linguagem, pesquisas apontam que bebês, já aos seis meses de idade, adquirindo a língua inglesa, compreendem muitas palavras de seus contextos sociais, como *mamãe*, *papai*, *mão*, *pé*, *boca*, *maçã*, *olhos*, *biscoito* (Bergelson, Swingley, 2012; Bergelson; Swingley, 2015; Tincoff; Jusczyk, 1999). Alguns meses depois, aos 18 e aos 24 meses, bebês neurotípicos já são capazes de juntar palavras que estabelecem relações composicionais sobre itens do mundo sem terem recebido nenhum tipo de ensinamento explícito (Booth; Waxman, 2002; Waxman; Booth, 2000; Willits et al., 2013; Wojcik; Saffran, 2013). Por essas evidências, seria plausível assumir que palavras concretas são as primeiras produções no vocabulário infantil. Enquanto palavras mais abstratas requerem mais modelos de suporte conceitual e linguístico (Crutch et al., 2009; Schwanenflugel, 1986) ou da cognição social (Borghi; Binkofski, 2014).

Interessantemente, essa diferença na tangibilidade do que os conceitos referenciam resulta em consequências cognitivas. Estudos em neurociência da linguagem (Amedi *et al.*, 2004; Bedny; Saxe, 2012) apontam que palavras concretas, mesmo sem referência dêitica eventual, são percebidas como itens mais fáceis de se visualizar, em função do nível de imageabilidade, e de se reter na memória, por causa da experiência direta (Bedny; Saxe, 2012;

Brennan, 2022; Paivio, 1990; Pulvermüller, 2013; Shebani, 2018; West et al., 2018). Em contrapartida, as palavras abstratas podem exigir operações mentais mais sofisticadas, as quais envolvem raciocínio inferencial e modelos de suporte conceituais¹ (*conceptual scaffolding*) para serem adquiridas (Borghi; Binkofski, 2014; Crutch et al., 2009; Veale; Keane, 1992).

Outros estudos (Brennan, 2022; Crutch et al., 2009; Lupyan; Winter, 2018; Pulvermüller, 2013) sugerem que o significado de alguns *conceitos concretos* emergem quando diversas áreas responsáveis pelo processamento de informações – sensoriais, imagéticas, táteis e motoras – são ativadas concomitantemente pelo cérebro, pois subjazem a ações e objetos diretamente percebidos pelos sentidos e por ações motoras.

A palavra *violão*, por exemplo, a princípio concreta, ativa os córtices sensoriais e motores, respectivamente o córtex sensorial primário no lobo parietal, referente ao processamento de informações sensoriais, como o som que este objeto produz, a cor e a representação da forma visual dele (as quais são quase sempre estáveis). O córtex motor primário no córtex frontal, por sua vez, é ativado em virtude da sensação tátil de pegar o violão e/ou da movimentação motora necessária para tocar suas cordas para que ele emita som (Brennan, 2022).

Por outro lado, palavras abstratas como *amor, dor, beleza, justiça, liberdade* ativam regiões associativas envolvidas na integração de experiências emocionais e sociais, mas também ativam áreas pré-frontais, as quais não têm relação com experiências perceptuais e motoras. Assim, o processamento de itens lexicais abstratos pode depender de redes neurais multimodais² distribuídas no cérebro (Brennan, 2022; Pulvermüller, 2013) e fazem parte inextricável de uma fisiologia sensorial sincrética.

Ainda que abordagens como a Cognição Corporificada³ (*Embodied Cognition*) e as Amodais⁴ tenham tentado explicar se o desenvolvimento conceitual e a aquisição vocabular

¹Para Veale e Keane (1992), *Conceptual Scaffolding* refere-se a estruturas conceituais provisórias que favorecem a construção progressiva do significado e funcionam como um suporte conceitual inicial para processos graduais de abstração semântica. Essa noção é interessante, principalmente ao pensarmos a aquisição de linguagem, pois sugere que os conceitos abstratos podem surgir a partir da elaboração de experiências e/ou representações mais concretas.

²Pulvermüller (2013, p. 458) esclarece que as áreas multimodais do córtex são: “regiões igualmente dedicadas ao processamento e à integração de informações provenientes de diferentes modalidades sensório-motoras.”

³Segundo Sakreida et al. (2013, p. 132, tradução da autora): na Cognição Corporificada, durante a compreensão da linguagem, ocorre uma simulação interna do conteúdo da palavra ou da frase. Assim, pressupõe-se que, durante o processo de simulação, seja recrutada a mesma rede neural sensório-motora ativada na interação com um objeto ou entidade, ou na execução da ação à qual a palavra se refere (Zwaan, 2004 *apud* Sakreida et al., 2013).

⁴Segundo Pulvermüller (2013, p. 458): o significado amodal (ou a semântica amodal) é definido em um espaço abstrato, por exemplo, em termos de traços semânticos abstratos (p. ex., SOLTEIRO = [+humano, +adulto,

são resultantes da experiência sensório-motora ou de algum outro domínio cognitivo inato, nenhuma dessas abordagens foi capaz de esclarecer como e quando as categorias conceituais com referências mais abstratas são adquiridas. Além disso, uma limitação da abordagem corporificada é que ela restringiria a aquisição à experiência sensório-motora.

Um estudo recente de Luchkin e Waxman (2025) traz evidências interessantes a essa limitação de abordagens corporificadas ao apontar que bebês típicos, falantes de língua inglesa, já aos 15 meses de vida são capazes de atribuir significado conceitual a novas pseudopalavras, mesmo na ausência de um referente visível. Em um experimento com rastreamento ocular, realizado na Plataforma *Lookit*, (Scott *et al.*, 2017; Scott; Schulz, 2017), disponibilizada pelo *Center for Brains, Minds and Machines* (CBMM) do Massachusetts Institute of Technology⁵, as autoras dividiram o estudo em duas fases:

Na primeira etapa, os bebês observavam três objetos familiares de uma mesma categoria semântica, por exemplo uma *maçã*, uma *laranja* e uma *banana*. O objetivo das autoras era ativar um efeito de *priming* semântico. Em seguida, uma pessoa apontava para algo escondido e dizia: “*Look, a modi!*” (pseudopalavra). Portanto, o bebê não via esse objeto, apenas ouvia essa pseudopalavra que poderia ser associada a este objeto oculto. Já na fase de teste, dois novos objetos eram apresentados na tela, sendo um da mesma categoria semântica da fase de familiarização, por exemplo uma *pitaya* e outro de uma categoria não relacionada, como um *puff*. Neste momento, o bebê ouvia novamente: “*Where is the modi?*” (“Cadê o modi?”). Os resultados apontam que os bebês de 15 meses olharam mais significativamente para o objeto da categoria congruente com a palavra nova aprendida, em comparação ao objeto de categoria não relacionada.

Esses dados são interessantes por apontarem para uma capacidade dos bebês, já aos 15 meses de inferir propriedades conceituais de um referente, generalizá-las e atribuí-las a uma categoria, o que evidencia um raciocínio representacional que ultrapassa a experiência sensório-motora imediata na atribuição do nome a um objeto. Isso aponta para uma habilidade precoce de bebês típicos de construir uma representação do significado de uma palavra, sem depender de estímulos perceptivos diretos como o objeto que designava aquela palavra.

Outras evidências sobre a limitação de abordagens corporificadas e sensório-motoras para explicar a aquisição lexical são os estudos com indivíduos cegos, os quais não seriam capazes de adquirir conceitos que nomeiam itens ou façam referência a experiências visuais,

+masculino, não casado]). Propõe-se que um sistema semântico amodal processe o significado de forma independente das informações provenientes dos sentidos.

⁵Para mais informações sobre, consultar: Research | The Center for Brains, Minds & Machines - CBMM @ MIT. Disponível em: <https://cbmm.mit.edu/research>. Acesso em: 15 out. 2025.

incluindo as experiências com cores (Landau; Gleitman, 1985). No entanto, há evidências de plasticidade ocorrendo para reorganizar os sistemas perceptuais em casos de privação sensorial, especialmente para os casos de privações congênitas. Por exemplo, dados de Estimulação Magnética Transcraniana (TMS) com cegos congênitos apontam para uma adaptação do lobo occipital (comumente associado à visão em indivíduos videntes) para atividades de processamento linguístico.

Em uma tarefa de produção verbal, na qual os participantes ouviam um substantivo e deveriam produzir oralmente um verbo, a aplicação de TMS durante a tarefa levou ao aumento de erros semânticos ou a atrasos na resposta em participantes cegos, efeito que não foi observado em participantes videntes. Esses resultados apontam para uma trans-modalidade sensorial e anatômica, uma vez que o córtex occipital se readaptou para processar estímulos auditivos ou táteis em tarefas linguísticas, confrontando a ideia de modalidade específica inata dos lobos cerebrais (Bedny; Saxe, 2012).

Outras evidências contundentes de sincretismo cognitivo provêm de estudos de neuroimagem, que utilizam Ressonância Magnética Funcional (fMRI) para mapear a atividade cerebral relacionada à linguagem em adultos cegos congênitos, revelando padrões de ativação cerebral distintos em comparação com um grupo controle pareado por idade. Eles ouviram frases, com uma estrutura sintática em dois níveis de dificuldade sintática.

Os resultados apontam que adultos cegos não apenas ativam áreas perisilvianas clássicas da linguagem do hemisfério esquerdo durante a compreensão da fala, como ocorreu com um grupo de adultos videntes, mas adicionalmente apresentam ativação nas estruturas homólogas do hemisfério direito e nos córtex occipital extraestriado e estriado. Tanto a atividade perisilviana quanto a occipital variaram em função da dificuldade sintática e do conteúdo semântico (Röder *et al.*, 2002).

Além disso, outros estudos qualitativos em neurociência da linguagem apontam que crianças cegas se utilizam também dos termos *olhar* e *ver* para se referir à percepção visual de pessoas videntes e à percepção tátil em si mesmas. Esses estudos indicam que cegos não só são capazes de adquirir e usar palavras que fazem referência a itens mais abstratos, como os verbos *cintilar*, *brilhar* ou *resplandecer*, mas também parecem ter uma estrutura conceitual similar a de videntes em relação à percepção de cores (Bedny *et al.*, 2011; Landau; Gleitman, 1985; Kim *et al.*, 2020).

Crianças cegas, por exemplo, foram capazes de atribuir cor a objetos concretos passíveis de terem cores, como *caminhão*, mas não a conceitos mais abstratos, como *ideias* (Landau; Gleitman, 1985). Elas também demonstraram uma compreensão acerca da

variabilidade de cor entre objetos de uma mesma categoria, por exemplo, as crianças compreendiam que é muito mais provável que carros tenham cores diferentes do que *bananas* (Bedny *et al.*, 2011).

Isso ocorre por meio da capacidade inferencial e da atuação de outras cognições como a linguagem, uma vez que a descrição de objetos e conceitos pode ser feita por outras pessoas. Essas evidências reforçam a proposição de que a tangibilidade transcognitiva impacta significativamente domínios distintos, uma vez que, para esses indivíduos especificamente a aquisição de conceitos pode ser representada mentalmente por meio de outros sentidos, como o tátil, o auditivo, o linguístico.

A partir dessas evidências, acreditamos que esses resultados possam trazer luz às nossas perspectivas em relação ao bebê, no período de aquisição de linguagem, considerando os conceitos de plasticidade e suas consequências no desenvolvimento cognitivo deles, sobretudo na primeira infância em razão dos períodos críticos do desenvolvimento (Hensch, 2004; Werker; Hensch, 2015). Isso porque a neuroplasticidade e o trânsito aberto entre diferentes interfaces e domínios cognitivos poderiam vir a ajudar os bebês, mesmo que de forma acessória, a processar o aprendizado de palavras sem referência direta no mundo. Todas essas descobertas nos falam em favor de uma efemeridade da divisão entre os conceitos de concretude e abstração.

Nesse viés, também é possível questionar se a aprendizagem inicial de palavras está só associada a experiências sensoriais. Será mesmo que os nomes concretos predominam nas primeiras fases do desenvolvimento lexical (Andrews *et al.*, 2008; Crutch *et al.*, 2009) em comparação aos abstratos (Johnston; Barry, 2006; Juhasz, 2005)?; ou será que a observação anedótica dos casos traria um viés das palavras mais facilmente observáveis? Certamente é mais fácil para os cuidadores perceberem que o bebê deles entendem e/ou tentam falar *bola* do que *frio*, mas isso não exclui a ideia de que os bebês consigam processar o significado de um item lexical mais abstrato como *frio* mesmo sem se referir a ele.

Um exemplo disso é como crianças, desde o primeiro ano de vida, adquirem propriedades e regras gramaticais abstratas, como os itens funcionais e as propriedades estatísticas, e as utilizam como pistas para identificar os itens lexicais na estrutura sintática da língua sem nenhum tipo de ensino formal ou esforço cognitivo.

Muitos estudos psicolinguísticos já forneceram evidências sobre a sensibilidade de crianças a esses itens funcionais (Saffran *et al.*, 1996; Shi *et al.*, 1999; Shi; Werker, 2001) e a outras propriedades, por exemplo as fonológicas e/ou distribucionais, prosódicas (De Carvalho *et al.*, 2019), estatísticas e probabilísticas (Shady, 1996) e morfofonológicas

(Corrêa, Name, 2003; Van Heugten; Shi, 2009; Van Heugten; Christophe, 2015) como forma de explicar como as crianças adquirem regras gramaticais abstratas, como as propriedades morfofonológicas e sintáticas das línguas às quais estão expostas.

Refletindo também sobre a aquisição de raízes verbais em eventos, há registros de que bebês, por volta dos treze meses, conhecem as exigências eventuais dos verbos, concernentes aos seus argumentos mesmo que não haja referência imediata a algo físico no mundo. Por exemplo, os verbos *abraçar* e *comer* abarcam conceitos mais complexos e específicos que não são propriamente coisas que existem no mundo, mas são mobilizações transitórias que se colocam no centro de eventos (Bergelson; Swingley, 2013). Alguns outros estudos investigam também como e se a capacidade preditiva de crianças, nos primeiros anos de vida, serve como um mecanismo para que elas identifiquem e aprendam itens lexicais semanticamente relacionados (Havron *et al.*, 2019; Rabagliati; Ferguson; Lew-Williams, 2019; Reuter *et al.*, 2023; Willits *et al.*, 2013).

Portanto, como explicar a capacidade de bebês de adquirir e manipular itens funcionais e identificar e manipular propriedades morfossintáticas, mesmo eles não tendo referentes concretos no mundo? Será que, em relação aos itens lexicais, haverá diferença no curso de aquisição em detrimento da natureza conceitual mais (ou menos) concreta desses elementos? Outra questão é: se qualquer domínio cognitivo de acesso a representações mentais é abstrato independentemente da sua natureza conceitual, será que a diferença entre itens funcionais, que dizem respeito à computação, por exemplo, os determinantes e as preposições em *A casa de boneca* e *A casa para boneca*, e os lexicais, i.e., aqueles que referenciam um objeto ou uma ação no mundo — como *boneca* e *casa* são adquiridos posteriormente a itens lexicais abstratos como *sono* e *fome*?

Para lidar com este tema, algumas questões metodológicas surgem, como **(i)** se o vocabulário dessas crianças, aferido por testes standardizados, realmente reflete maior produção e/ou maior produção e compreensão de palavras concretas do que de palavras abstratas ou **(ii)** será que as palavras abstratas não foram exploradas suficientemente de forma que ainda não houve evidências sobre a aquisição desse tipo de conceito? **(iii)** será que essa ausência de evidência recai no fato de que os inventários de palavras não possuem categorias semânticas de classes mais abstratas e as bases de dados sobre o vocabulário infantil foram respondidas com base na intuição dos responsáveis pelas crianças sobre aquilo que elas já compreendem ou compreendem e dizem? **(iv)** na fala direcionada às crianças, será que os nomes abstratos são adquiridos posteriormente devido às suas possíveis exigências cognitivas ou o contexto linguístico desempenha um papel fundamental em sua assimilação?

Ainda que alguns estudos psicolinguísticos tenham apontado para a sensibilidade precoce de bebês a itens funcionais (De Carvalho *et al.*, 2019; Shady, 1996; Shi *et al.*, 1999; Shi; Werker, 2001) e a capacidade de identificar ou de predizer itens lexicais concretos e que fazem referência a objetos das cenas diárias deles (Bergelson; Swingley, 2012, 2015; Havron *et al.*, 2019; Rabagliati; Ferguson; Lew-Williams, 2019; Reuter *et al.*, 2023; Tincoff; Jusczyk, 1999), a questão sobre como as crianças adquirem e quando interpretam itens lexicais abstratos, mas não os funcionais, ainda não foi amplamente explorada.

Por essa razão, um estudo empírico sobre como e quando as crianças começam a interpretar palavras abstratas, como *frio*, *sujo*, *alegria*, *fome* parece relevante, pois o tema não é recorrente na literatura psicolinguística, sobretudo com crianças brasileiras e, pode fornecer dados experimentais sobre as etapas da aquisição lexical. Além disso, discutimos aqui como esse processo cognitivo altamente complexo que compreende desde a percepção sensorial à formação de representações conceituais abstratas pelo cérebro, possibilita que sons, palavras e gestos ultrapassem sua materialidade física e se convertam em significados na mente humana.

Assim, os objetivos gerais deste estudo são verificar se crianças brasileiras, entre 24 e 30 meses de idade, adquirindo o PB, já interpretam palavras concretas e abstratas; compreender as diferenças cognitivas na aquisição e no processamento de itens lexicais concretos e abstratos; e aferir se há maior interesse atencional por itens lexicais concretos em comparação aos abstratos, em função da natureza conceitual deles.

Além disso, há também objetivos específicos, a saber: fornecer dados inéditos a partir da elaboração de inventários de palavras sobre a aquisição vocabular de crianças brasileiras e da publicação de resultados experimentais sobre as etapas iniciais da aquisição linguística, em um país no qual os parâmetros socioeconômicos são muito diferentes da maioria dos estudos que compõem a literatura em aquisição de linguagem, proveniente dos chamados países WEIRD⁶ - *Western, Educated, Industrialized, Rich and Democratic* (Singh *et al.*, 2023).

Objetivamos, também, verificar a possibilidade de utilizar o teste aqui desenvolvido como uma ferramenta linguística preditiva para transtornos do neurodesenvolvimento que acometem o processamento linguístico abstrato e democratizar a pesquisa em neurodesenvolvimento cognitivo no Brasil por meio da divulgação e da ampliação do *Lookit*, enquanto plataforma experimental.

A execução de experimentos sobre o desenvolvimento linguístico com crianças brasileiras é relevante para a área das ciências cognitivas por promover uma compreensão mais representativa dos fenômenos cognitivos sobre a aquisição e a categorização da

⁶A tradução da sigla WEIRD corresponde a: Ocidental, Educado, Industrializado, Rico e Democrático.

linguagem em uma questão ainda não explorada sobre a aquisição de nomes concretos e abstratos e com uma amostra populacional diversa da maioria dos estudos já publicados que se concentram majoritariamente no hemisfério norte (Kidd; Garcia, 2022).

De forma a suprir as questões acima apresentadas, foi desenvolvido um experimento no *Lookit*, uma plataforma desenvolvida pelo *Massachusetts Institute of Technology*⁷, através do paradigma da Fixação Preferencial do Olhar (*Looking While Listening* — LWL) com rastreamento ocular (Bánki *et al.*, 2022; Slim *et al.*, 2024). Este paradigma, além de ser consagrado na literatura para experimentos com bebês na faixa etária de interesse deste estudo, entre 24 e 30 meses de idade (Frota; Name 2017), também é adequado por permitir avaliar a compreensão de palavras em tempo real, capturando a rapidez e a precisão com que as crianças associam os estímulos auditivos aos visuais. Ademais, é possível captar a movimentação ocular delas de forma mais natural e instantânea, durante o acesso lexical.

Neste teste, realizado *online*, investigamos a aquisição de nomes concretos e abstratos por crianças brasileiras de 24 a 30 meses de idade. Nele, os bebês são apresentados a duas imagens, cada uma ocupando a metade da tela do monitor (ex., a imagem de um balão vs. a imagem de um fogão) e poderão ouvir um áudio pedindo a atenção para uma das duas palavras concretas (ex., “Olha! fogão/Olha! balão”). Alternativamente, as mesmas duas figuras (a imagem de um balão vs. a imagem de um fogão) são usadas para ilustrar palavras abstratas (ex., “Olha! Quente/Olha! Festa!”).

Essas quatro possibilidades (fogão, balão, quente, festa) estão distribuídas em um quadrado latino. A variável dependente é o tempo de reação do participante para olhar a imagem-alvo em cada condição (ex., olhar para o fogão quando ele escutar “fogão” em comparação à quando ele ouvir “quente”). As hipóteses são de que as crianças brasileiras de 24 a 30 meses já são capazes de reconhecer os itens com referentes concretos e abstratos; haverá diferença no acesso e no processamento desses itens lexicais, em função da natureza conceitual deles e que as crianças apresentarão maior engajamento atencional para itens concretos, em função de seu maior nível de imageabilidade. Se o resultado deste experimento for ao encontro das nossas hipóteses, será possível indicar que a natureza conceitual do item lexical influencia no acesso dessas palavras pelos bebês.

A questão de como definir os itens lexicais para o experimento se constituiu como um desafio, porque não só era preciso selecionar palavras para o teste experimental, como também decidir a classificação delas em termos de concretude e abstração. Por essas razões, a

⁷Para mais informações sobre, consultar: Research | The Center for Brains, Minds & Machines - CBMM @ MIT. Disponível em: <https://cbmm.mit.edu/research>. Acesso em: 15 out. 2025.

primeira etapa foi realizar uma adaptação da versão brasileira do *Children Development Inventory* (Teixeira, 2000) a fim de evitar que a escolha dos itens vocabulares se desse por critérios próprios da pesquisadora.

Nesta versão adaptada, foram incluídas algumas palavras já existentes, porém adicionamos muitas outras de natureza abstrata, privilegiando substantivos. Posteriormente, recrutamos responsáveis por crianças brasileiras, entre 24 e 30 meses, para que indicassem, assim como no teste original, quais palavras as crianças compreendiam e quais elas já compreendiam e diziam. A versão final deste formulário foi respondida por 56 responsáveis de bebês brasileiros entre 24 e 30 meses.

Em seguida, somente as palavras que obtiveram uma taxa de classificação no nível *compreende e diz*, acima de 50% foram introduzidas em um novo formulário para serem finalmente classificadas como concreta ou abstrata por 445 adultos brasileiros. A partir do cruzamento de dados dos dois formulários, foram selecionadas 12 palavras abstratas (*quente; festa; amor; alegria; medo; frio; fome; carinho; vazio; sujo; escuro; sono*) e 12 palavras concretas (*fogão; balão; beijo; menino; tigre; casaco; pizza; abraço; prato; tênis; buraco; pijama*). Depois dos materiais do experimento terem sido confeccionados, realizamos um pré-teste controle com 53 adultos, falantes nativos do português brasileiro, também na plataforma *Lookit*.

A partir da discussão e dos resultados encontrados, aventamos a possibilidade de uso deste teste como uma ferramenta preditiva linguística que poderá ser validada futuramente para auxiliar na identificação precoce de transtornos cognitivos, especificamente de crianças com suspeita de Transtorno do Espectro Autista (TEA), que frequentemente apresentam dificuldades significativas com conceitos abstratos e linguagem figurada (Philofsky *et al.*, 2007; Tenenbaum *et al.*, 2024).

A partir de evidências empíricas das etapas iniciais da aquisição lexical, para sondar a interface sintaxe-semântica de nomes abstratos e concretos, esperamos, com este estudo, fornecer dados sobre como itens lexicais de natureza conceitual diferentes passam a ser compreendidas como entidades dotadas de significado. Também esperamos revelar parte da arquitetura oculta do significado na mente de bebês e de adultos neurotípicos, enquanto são contrastados nomes concretos e abstratos semanticamente relacionados.

Para tanto, esta tese está organizada da seguinte forma: além da presente introdução, o segundo capítulo discute a ontologia dos conceitos, sob uma perspectiva histórica que vai desde a filosofia até os estudos em neurociência da linguagem. O objetivo é compreender a natureza dos conceitos: seriam eles entidades mentais ou produtos da percepção sensorial?

Qual é o substrato neuropsicológico desses termos arbitrários e socialmente convencionados? O terceiro capítulo conta com algumas evidências da psicolinguística, em relação à sensibilidade precoce de bebês a itens funcionais e lexicais, destacando que os itens funcionais, embora não possuam um referente concreto no mundo, são adquiridos precocemente e servem como suporte para que as crianças identifiquem, por meio da sintaxe, os itens lexicais, i.e., aqueles que, em sua maioria, fazem referência a elementos concretos e observáveis do mundo.

O quarto capítulo apresenta uma revisão sistemática da literatura sobre a aquisição de itens concretos e abstratos, utilizando o *Google Scholar* como banco de dados. No quinto capítulo, são descritos os pré-testes psicolinguísticos, elaborados pelo *Google Forms*, para selecionar os estímulos dos testes principais. No primeiro, 56 responsáveis por crianças entre 24 e 30 meses de idade responderam sobre quais palavras do PB as crianças desta faixa etária já compreendem ou já compreendem e dizem. Em seguida, todas as palavras que tiveram um percentual acima de 50% na taxa de resposta *compreende e diz* foram inseridas em um novo questionário e definidas, em termos de concretude/abstração, por 445 adultos falantes adultos do PB.

Além desses dois pré-testes, são apresentados aqui os resultados do teste controle aplicado em adultos e do experimento principal com crianças entre 24 e 30 meses de idade, usando uma nova plataforma de aplicação remota – *Lookit* – chancelada pelo *Massachusetts Institute of Technology*. A utilização desta metodologia permite avaliar o rastreamento ocular de forma manual com alta precisão na inferência do olhar do participante, porém sem o uso do *eye-tracker*.

No sexto capítulo, são reportados os resultados do experimento controle realizado com adultos e com crianças, entre 24 e 30 meses de idade, para apontar as diferenças na aquisição e no processamento de conceitos concretos e abstratos. O sétimo capítulo apresenta a discussão dos resultados, o oitavo reúne as considerações finais e os encaminhamentos futuros desta pesquisa.

2 ONTOLOGIA DOS CONCEITOS: UMA ANÁLISE HISTÓRICA SOBRE CONCRETUDE E ABSTRAÇÃO

Compreender como conceitos, enquanto formas abstratas de representação, podem ou não ter referência na realidade material, assim como identificar os processos neurocognitivos que sustentam os processos linguísticos são questões que remontam à filosofia antiga, com Platão e Descartes (Marcondes, 2010; Philippe, 2002), até os modelos de processamento da neurociência cognitiva contemporânea (Binder, 2009; Binder, 2011; Pulvermüller, 2013; Vigliocco, 2014). Neste capítulo será apresentada uma visão não exaustiva das tentativas de compreender o que está por trás da percepção da dicotomia concreto *vs.* abstrato ao longo dos séculos.

Ao tratar a questão acerca do significado das palavras, a filosofia antiga estabeleceu o problema do conhecimento e instaurou o chamado *projeto epistemológico* como objeto central da sua investigação. Nessa tentativa, foram contemplados temas sobre o significado e a natureza desses estados mentais que resultaram na formulação das noções de conceito, ideia e representação mental. Um dos principais nomes deste período é Platão, discípulo de Sócrates.

Em *Crátilo*, o diálogo pioneiro aos estudos da linguagem, Platão discorre sobre a problemática da convencionalidade dos signos e traz a proposição dos mobilistas e a dos sofistas para tentar respondê-la. Nesse diálogo, o naturalismo, representado por Crátilo, coloca que o signo compartilha a natureza daquilo que ele representa e há uma correspondência natural entre ele e o objeto por ele designado. Em contrapartida, o convencionalismo, interpretado por Hermógenes, propõe que as palavras são apenas convenções sociais e que não há uma relação intrínseca entre a palavra e a coisa por ela referenciada. Por isso, os signos não possibilitam o acesso direto à realidade (Marcondes, 2010; Philippe, 2002).

Platão expôs as limitações de ambas as abordagens, dado que, no naturalismo, a explicação de uma suposta relação intrínseca entre palavra e objeto se torna frágil, sobretudo diante da variabilidade das línguas, que contradiz a noção de uma correspondência direta entre os termos linguísticos e seus referentes. No convencionalismo, por sua vez, há a dificuldade de explicar a própria origem das convenções: de que modo seria possível instaurá-las antes do surgimento da linguagem, se a linguagem seria indispensável no estabelecimento dessa convenção palavra-objeto? Diante disso, Platão, em *Crátilo*, discorre sobre a “correção dos nomes” e sobre a necessidade de ir além deles. Essa constatação é pertinente à relação entre palavras (nomes) e a realidade correspondente, a qual foi

posteriormente tratada nos estudos linguísticos, sobretudo nos de aquisição de linguagem, como *o problema da referência*, tratado por Waxman (2006).

Essa questão é, particularmente, interessante ao pensarmos sobre a conceptualização dos conceitos, principalmente diante dessa dicotomia de palavras que representam/nomeiam itens concretos e outros mais abstratos. Se o *problema da referência* se impõe ao tratarmos unicamente sobre a natureza conceitual dessas palavras e o estabelecimento de suas representações mentais, torna-se ainda mais instigante pensarmos no processo de aquisição linguística, em que não só a língua, como também o mundo precisa ser categorizado pelo bebê.

Aristóteles também se ocupou do problema do significado, ao investigar a relação entre a mente e a realidade. Ele foi o precursor ao propor a relação entre linguagem e mente, influenciando outros pensadores como Descartes, Port-Royal e Locke, os quais serão abordados posteriormente. Para Aristóteles, em seu *Tratado da Interpretação (De Interpretatione)*, as palavras, enquanto signos, relacionam-se com a realidade apenas mediante a intervenção da mente. Assim, a entidade mental se constitui como um sinal natural que é captado pelos nossos sentidos e se estabelece na nossa mente.

Para Aristóteles, a noção de conceito era o de *afecções da alma*⁸ e essas seriam como mediadores entre as coisas do mundo e as palavras. Ao mesmo tempo em que os conceitos se constituem como uma entidade mental, eles é que dão acesso à realidade. Por isso, eles são representações mentais de caráter subjetivo e interior, que se referem a uma realidade objetiva e externa (Marcondes, 2010; Philippe, 2002).

De acordo com Aristóteles, o signo linguístico precisa estar vinculado a um conceito para ter significado. Porém, ele assume a concepção de que o signo é convencionalizado e a sua conexão só se dá pela percepção da mente com a realidade. A significação linguística, por sua vez, depende das representações mentais, dos sinais aos quais os conceitos estão vinculados.

A natureza dos conceitos, isto é, as *afecções da alma* são universais e resultam da experiência sensível e da abstração intelectual. Segundo ele, as representações mentais emergem mediante a generalização derivada da percepção concreta. Ainda que os signos linguísticos variem e sejam convencionalizados culturalmente, conforme apontado por Platão no *Crátilo*, a relação entre a palavra e conceito não é direta, mas é fruto dessa entidade mental. Há, portanto, uma concepção mais mentalista nesta abordagem (Marcondes, 2010; Philippe, 2002).

⁸Do latim: *Pathemata Psychai*.

Ockham (c.1285-1349) contrapõe-se ao inatismo de Platão e ao Empirismo de Aristóteles e se caracteriza como um empirista-nominalista. Em seu tratado *Summa totius logicae* (publicada postumamente em 1488), ele aborda uma concepção nominalista sobre a natureza do signo. Para ele, não há entidades abstratas e todo conhecimento parte da apreensão de indivíduos concretos. Por exemplo, o termo *cachorro* não se refere a nenhuma entidade, mas pode referenciar um único cachorro quando precedido por um pronome como *este*, ou um conjunto de cachorros. Sua ontologia se baseia na *Navalha de Ockham*, um princípio metodológico segundo o qual as entidades não devem ser multiplicadas além do necessário⁹, isto é, não se deve multiplicar além do necessário (De Boni, 1995; Marcondes, 2010).

Para *Ockham* existem apenas objetos singulares, i.e., só indivíduos concretos possuem realidade, sendo que gêneros e espécies não existem por si mesmos. Assim, na sua visão, só há indivíduos concretos e passíveis de serem referenciados, como *este girassol*, mas o termo *flor* é apenas uma palavra usada para agrupar indivíduos semelhantes e não corresponde a nada que exista independentemente. Elas são apenas nomes ou conceitos usados para organizar o pensamento e a linguagem, mas não têm realidade própria. Além disso, são signos convencionados e que, por meio de conceitos mentais, nos permitem designar objetos concretos.

Ainda que *Ockham* atribua aos conceitos mentais o significado das palavras, permanece o problema em relação à referência, uma vez que ele não explica como acessamos esses conceitos, nem como eles estabelecem o vínculo entre linguagem e realidade. Em relação à semântica, o filósofo parece ter dado continuidade ao *Tratado da Interpretação* aristotélico, pois não dissocia as palavras das *afecções da alma*¹⁰.

Segundo *Ockham*, o sentido da palavra só é possível por meio da associação a um conceito mental, o qual é uma entidade psíquica natural que faz a mediação entre linguagem e realidade. Por exemplo, ao falarmos a palavra *árvore*, não nos referimos imediatamente a todas as árvores concretas, existentes no mundo, mas sim a um conceito mental de *árvore*, i.e., uma representação formada pela mente, a partir da experiência com os indivíduos particulares. Por essa razão, o filósofo argumenta que é o próprio conceito que porta o significado, sendo ele um intermediário necessário entre o signo linguístico e os objetos a que ele se aplica. *Ockham* é, portanto, mais um conceitualista que nominalista, visto que ele supõe a existência de uma *lingua mentis* em que os conceitos mentais funcionam como signos

⁹Do latim *entia non sunt multiplicanda praeter necessitatem*.

¹⁰Do latim, *pathemata psychês*.

naturais, diferentemente dos termos falados ou escritos que são convencionados (De Boni, 1995; Marcondes, 2010).

Partindo da tese ockhamista de que o sentido é veiculado por conceitos mentais que mediam linguagem e mundo, a reflexão moderna passa a perguntar como esses conceitos se formam e em que medida podem fundamentar o conhecimento. Neste momento, surgem duas correntes que influenciaram fortemente o pensamento linguístico: o empirismo, que busca explicar se o conhecimento humano provém da experiência sensível, e o racionalismo, que defende que tudo advém da razão (Murphy, 2017).

Para os empiristas, a mente humana é como uma *tábula rasa*, i.e., como uma folha em branco (Kennedy, 2013). Por isso, o conhecimento é construído a partir da percepção do mundo, através dos sentidos e da experiência do indivíduo com o meio. Em oposição, os racionalistas argumentam em favor da noção de inatismo, i.e., de que a razão é o princípio do conhecimento e que a experiência sensorial, *per se*, pode ser distorcida. Essas questões serão apresentadas a fim de problematizar como ocorre a formação dos conceitos, sobretudo se a restringirmos à aquisição de palavras concretas e abstratas.

Na perspectiva empirista, representada por Locke (1632-1704) e Hume (1711–1776), a formação dos conceitos ocorre por meio da experiência sensível, i.e., a visão, a audição, o tato, o olfato, o paladar e de outros domínios sensoriais. Para Locke, são as ideias que fornecem o conteúdo cognitivo das palavras e que, por fim, constituem-se como signos convencionais das coisas. Isso só ocorre porque as ideias provém da experiência e da percepção sensível, passando por processos de generalização, abstração e reflexão. Essa postulação da relação das ideias com as coisas e das palavras com as ideias ficou conhecida como dupla conformidade das ideias, na obra *Ensaio Sobre o Entendimento Humano*, de Locke (Marcondes, 2010).

A *Teoria Clássica dos Conceitos*, elaborada por Locke, pressupõe que os conceitos têm uma estrutura interna composta por um conjunto de características definidoras, sendo alguns mais tangíveis e outros mais abstratos. Nesse viés, conceitos que designam objetos concretos, como *banana*, *copo* ou *brinquedo* foram tratados por Locke como *ideias simples*, uma vez que são diretamente percebidos e vinculados aos seus significados pelos nossos domínios sensoriais e, por isso, são mais facilmente observáveis. Por outro lado, palavras mais abstratas como *equilíbrio*, *justiça* ou *pensamento* seriam resultantes de generalizações de experiências particulares que são associadas (Ayers, 2001; França, Ferrari; Maia, 2016; Locke, 1847).

Nesta perspectiva da *Teoria Clássica dos Conceitos*, as crianças, adquirindo uma língua, não só precisam associar um som a um objeto, mas também precisam construir internamente um conjunto de propriedades que definem aquele conceito, principalmente para palavras concretas, como *bola* ou *gato*, as quais podem ser adquiridas mais facilmente do que palavras mais abstratas como *amizade*, *amor*, *dodói*. Uma criança, por exemplo, ao aprender o conceito de cachorro pode associar características como *animal*, *abana o rabo*, *peludo*, *lambe*, *tem quatro patas*, *late*. Já *bravo* pode ativar uma rede bem diferente como *malvado*, *forte*, *morder*, *medo*, *dodói*.

Segundo Locke, muitos dos conceitos lexicais são chamados de complexos por possuírem uma combinação de características definidoras. Porém, alguns outros conceitos, tratados como primitivos, requerem representações que envolvem natureza sensorial ou perceptiva. Embora a *Teoria Clássica dos Conceitos* de Locke remonte à ideia de conceitos definidos por propriedades essenciais, elaborada por Platão e Aristóteles no século IV AEC¹¹, ela é uma tentativa de explicar não só como os conceitos são formados, mas também como são organizados a partir de propriedades perceptivas e afetivas associadas a essas palavras.

Essa noção de ativação de rede foi reforçada, no século XXI, com estudos de neuroimagem, que indicam a ativação de outros domínios cognitivos, mais próximos do sistema límbico, no processamento de palavras abstratas. Esse mecanismo poderia, por exemplo, facilitar a aquisição da língua, sobretudo a de itens mais abstratos (Locke, 1847; Margolis; Laurence, 1999). De qualquer forma, ainda que a natureza dos conceitos pareça influenciar na forma como são adquiridos, tanto para Locke quanto para Hume, é a experiência sensorial que prediz a conceitualização.

Em oposição, filósofos como Descartes e Leibniz defendem que a razão é suficiente para a aquisição de conceitos. Na visão deles, conceitos que não apresentam materialidade física e/ou concreta no mundo como *paz*, *igualdade*, *leveza* não podem ser compreendidos apenas pela experiência perceptiva, portanto não podem ser reduzidos a impressões sensoriais. Na visão dos racionalistas, apenas mecanismos inatos e capacidades de raciocínio lógico seriam capazes de explicar a aquisição desse tipo de conhecimento.

Em *O Discurso do Método*, Descartes (1596-1650) desenvolve a lógica do pensamento aplicada ao método científico e discorre sobre questões epistemológicas voltadas à teoria do conhecimento científico, influenciando toda a ciência moderna. Para o referido filósofo, a

¹¹Era Comum (EC) e Antes da Era Comum (AEC) são notações cronológicas equivalentes a d.C. e a.C., respectivamente, adotadas como alternativa em textos científicos sem conotação religiosa. Ambas seguem o calendário gregoriano e não incluem um ano zero.

mente percebe o real por meio das ideias, sendo elas que representam as coisas. A linguagem, por sua vez, não só é a expressão do pensamento, como é também uma evidência da razão (Cottingham, 1999; Marcondes, 2010).

A Lógica de Port-Royal, elaborada depois de Descartes, na metade do século XVII, também influenciou o pensamento moderno com uma epistemologia de ordem psicológica. Segundo os pensadores de Port-Royal, as ideias, no sentido cartesiano, representam os objetos, e o significado das palavras tem prioridade da ideia sobre o termo. Assim, o sentido não está no som ou na grafia, mas na ideia que o signo evoca. Por isso, o significado de um termo é a ideia a ele associada. Nessa relação palavra-ideia-coisa, os signos são convencionais, porém sua inteligibilidade não é totalmente arbitrária, pois está relacionada à adequação das ideias aos objetos (Arnauld; Lancelot, 2001; Marcondes, 2010).

Nesse viés, a distinção entre concreto e abstrato diz respeito a como a palavra designa aquela ideia e/ou objeto. Por exemplo, *homem* é concreto, porque designa um sujeito humano. Enquanto, *humanidade* é um termo abstrato, pois nomeia essa forma em si e não designa indivíduos. Essa distinção é compatível com a “realidade psicológica” de *Port-Royal*, pois ancora a significação na relação palavra-ideia-coisa, e não apenas na convenção que a estabeleceu.

Depois da tradição mentalista de Port-Royal, o século XIX desloca o problema do significado para a psicologia experimental. Neste momento, Wundt (1832-1920) funda o laboratório de Leipzig e desenvolve a *Völkerpsychologie*, diferenciando a experiência imediata, a partir da sensação/percepção das funções superiores como linguagem e pensamento. Nesse viés, conceitos concretos estão mais relacionados à experiência imediata/direta, resultantes desses processos sensoriais e perceptivos. Enquanto a abstração é resultante de uma operação de apercepção, i.e., um processo complexo que envolve a percepção, a consciência, a significação e a interpretação de uma experiência sensorial¹², além de outros processos cognitivos complexos como a memória e a atenção (Wundt, 1912).

Embora tenha havido em Wundt, uma psicologização do significado e a inteligibilidade estivesse relacionada ao status psicológico dessa relação, a concepção de palavra e seu significado foi fortemente impactada pela publicação da obra póstuma de Saussure (1857-1913). Para ele, o significado não é exclusivamente mental, mas estrutural e relacional. Essa diferença no paradigma se baseia na diferença referencial entre o que para

¹²Etimologicamente, segundo o dicionário Michaelis, 2022, apercepção significa “ato ou efeito de perceber”. A diferença entre o processo de percepção e apercepção pode ser exemplificada com a percepção de um som, como o motor de um carro. A apercepção, por sua vez, é a consciência de que esse som indica que um carro está funcionando e/ou se aproximando.

Descartes e, na lógica de *Port-Royal*, era concebido como a relação palavra–ideia–coisa e aquilo que, em Saussure, passa a ser formulado como a relação entre significante e significado. No Curso de Linguística Geral, o significado, de forma geral, passa a ser estudado dentro da perspectiva de signos, os quais diferem entre si dentro de um sistema autônomo (Saussure, 1916).

A respeito disso, Marcondes (2010, p. 86) explica:

O signo é arbitrário e se define pela diferença em relação aos outros signos no interior do sistema, se estudar a língua como um sistema de signos. Essa proposta consiste em analisar a linguagem não como representação mental do real, nem como comunicação, mas sim primordialmente como um sistema dotado de uma organização interna a partir da qual o significado dos signos se constitui, tornando então possíveis a referência e a comunicação. O signo é arbitrário e se define pela diferença em relação aos outros signos no interior do sistema.

Para Saussure, o conceito se estabelece não pela relação da palavra com o objeto em si, mas por ser a união de um significado com o significante, i.e., a relação arbitrária de um conceito com a sua imagem acústica. Essa imagem não diz respeito ao som físico, mas a impressão psíquica dele. Por isso, é “material” só nesse sentido sensorial, em contraste com o conceito, que é algo mais abstrato.

A noção de concretude e abstração recai, portanto, em como o significado é compreendido nessas perspectivas. Em primeiro lugar, a oposição em relação ao significado diz respeito à transição dessa lógica cartesiana de ordem psicológica, em que o valor do signo decorre da natureza da ideia. Enquanto que, para Saussure, o signo linguístico se estabelece na relação com os outros elementos dentro deste sistema linguístico autônomo e estruturado (De Carvalho, 1976; Saussure, 1916).

Na década de 1930, sob influência do behaviorismo e do empirismo lógico, as hipóteses mentalistas perdem centralidade. O foco, então, passa a ser a descrição do que é observável na língua em uso e dos aspectos culturais que incidem sobre ela. A aquisição da linguagem recebe grande influência da psicologia behaviorista e, com a publicação de *Verbal Behavior* (1957), é tratada como aprendizagem por condicionamento operante, i.e., respostas verbais modeladas por reforço, controle de estímulos, discriminação e generalização. Nesse enquadramento, a diferença entre palavras concretas e abstratas descreve sobretudo variações no alcance dos estímulos que controlam a resposta verbal.

Fora do behaviorismo, Sapir (1884–1939) e Whorf (1897–1941) reformularam a hipótese Sapir–Whorf. De acordo com ela, a língua não só é apenas reflexo da sociedade, como também contribui para a constituição da cultura e da identidade da comunidade de fala.

Desse modo, o que é tomado como mais concreto ou mais abstrato pode variar entre línguas e culturas, conforme o sistema linguístico e as práticas culturais de cada sociedade (Boas, 2007; Marcondes, 2010; Sapir, 1980;).

Neste mesmo período, a publicação da obra *Syntactic Structures*, em 1957, foi uma revolução nos estudos linguísticos. A gramática gerativa de Chomsky recupera a tradição cartesiana de Port-Royal ao tratar a linguagem como faculdade da mente (Murphy, 2017) e retoma a concepção mentalista como forma de se compreender e se explicar a competência linguística. Esse momento representou uma mudança do enfoque empirista (influenciado por Hume e Locke) para o racionalista (inspirado em Descartes e Kant), deslocando o referencial teórico dos pesquisadores, para uma corrente denominada cognitivista em que:

a maior parte dos adeptos da nova corrente reconhece o substrato material dos processos mentais, que ocorrem no sistema nervoso central, daí uma preocupação para com os fundamentos biológicos da linguagem (Chomsky *apud* Selier-Cabral, 1991, p. 21).

Por essas razões, a compreensão daquilo que é concreto e abstrato muda de *status* ontológico. Essa noção deixa de ter um enfoque lógico-representacional, inspirado em universais lógicos, como na Escola de Port-Royal, para ser tratada como parte da arquitetura formal da língua. Em 1957, Chomsky discorre sobre a autonomia da sintaxe e, em 1965, fala sobre as propriedades lexicais, i.e., traços e restrições de seleção que condicionam as combinações sintáticas (Chomsky, 1965). A noção de concretude e abstração passa a ser tratada como propriedades do léxico mental que, por meio da seleção e da subcategorização, regulam a compatibilidade entre itens e a interpretabilidade das sentenças¹³.

Com a mente como centro do interesse de investigação e com o avanço da neurociência e das técnicas de imagem, a discussão sobre a representação dos conceitos e sobre os processos neurobiológicos que subjazem o acesso lexical e às suas representações mentais se torna um campo de interesse da linguística cognitiva e da neurociência da linguagem. Nesse cenário, mapear empiricamente a diferença entre palavras concretas e abstratas, a partir de evidências comportamentais e neurais, traz à luz a pergunta sobre a natureza da representação semântica: são os significados amodais e abstratos, codificados em um sistema simbólico central, ou dependem, ao menos em parte, de sistemas sensório-motores?

¹³Para mais detalhes sobre a noção de traços e de interpretabilidade na sintaxe, segundo a teoria de Chomsky, recomenda-se a leitura de Chomsky (1995).

Para a semântica abstrato-simbólica, o significado é proveniente de um processo mais abstrato e não depende dos nossos receptores sensoriais. Peacocke e Rey definem que conceitos são entidades puramente abstratas e não podem ser concebidos como representações mentais concretas (Margolis; Laurence, 1999). Nas palavras de Peacocke (1992, p. 3): “é possível que um mesmo conceito receba representações mentais diferentes em indivíduos distintos” e Rey (1994, p. 186) reforça a sua ideia e traz uma noção pertinente a respeito da dissociação entre representações mentais e conceitos:

Muitos filósofos sustentam a ideia de que esses tipos de representações internas não seriam mais idênticos aos conceitos do que as palavras em uma língua natural. Uma pessoa pode expressar o conceito CIDADE pela palavra 'city', outra pela palavra 'ville'; ainda outra talvez por uma imagem mental de avenidas movimentadas; mas, apesar disso, todas elas poderiam compartilhar o mesmo conceito CIDADE; uma poderia acreditar e outra duvidar de que cidades são lugares saudáveis para viver. Além disso, pessoas diferentes poderiam usar a mesma representação para expressar conceitos distintos: uma poderia usar uma imagem de Paris para expressar PARIS, outra para expressar FRANÇA.

Por isso, Rey enfatiza a dissociação entre a estrutura conceitual e as representações individuais e aponta que, ainda que os conceitos sejam acessados ou expressos por representações mentais, eles não se limitam a elas. Além disso, o autor cita o fato de que uma palavra como *cidade* pode ser *ville*, em francês, retomando também uma questão interessante sobre a arbitrariedade saussuriana do signo linguístico, já que várias formas sonoras podem designar variações dessas representações mentais. Logo, tanto as formas linguísticas quanto as representações mentais são arbitrárias.

Essa discussão sobre a dissociação entre a forma do conceito, seja ele entendido como representação mental ou como entidade abstrata, foi retomada por Jackendoff (1999). Ao defender que a linguagem é usada para expressar conceitos, o autor desenvolve a Semântica Conceitual, argumentando a favor da necessidade de uma teoria que integre tanto a linguagem quanto a mente. Além disso, ele questiona quais são os elementos conceituais primitivos subjacentes à cognição social e como os sistemas conceituais podem ser adquiridos (Jackendoff, 1999). Esse momento representa uma transição de paradigma para as abordagens corporificadas e para os estudos sobre as redes semânticas, principalmente pelo avanço nas técnicas de neuroimagem.

Nesse contexto, para a semântica corporificada (Barsalou, 2008; Bedny; Saxe, 2012; Glenberg et al., 2002; Glenberg et al., 2008; Fischer et al., 2008; Pulvermüller, 2013; Shebani, 2018; West *et al.*, 2018), a experiência sensório-motora de um indivíduo prediz o processamento lexical e ativa estruturas corticais distribuídas por várias vias no lobo frontal,

temporal e parietal, responsáveis pelo processamento linguístico de informações sensório-motoras (Brennan, 2022).

Palavras concretas, por exemplo, como verbos de ação como *correr*, *comer*, *tocar*, dentre outros recrutam maior ativação em córtices perceptuais e sensório-motores, sugerindo que sua representação está mais firmemente vinculada à experiência corporal. Não obstante, elas parecem ser acessadas e processadas mais rapidamente em comparação às abstratas, o que foi definido por Paivio (1990) como *efeito de concretude*.

O processamento de palavras de conteúdo mais abstrato, por sua vez, ativa redes neurais distribuídas que integram regiões referentes ao processamento emocional e semântico, incluindo o giro frontal inferior esquerdo e os lobos temporais anteriores (Binder, 2009; Pulvermüller, 2013; Vigliocco, 2014). Além disso, alguns conceitos que não possuem referência única como *amizade*, *justiça*, *liberdade*, *amor*, *responsabilidade* dependem de experiências sociais e individuais para serem adquiridos (Borghi; Binkofski, 2014; Borghi *et al.*, 2018; Pexman *et al.*, 2022; Snedeker; Gleitman, 2004).

Além da questão da cognição corporificada, evidências sobre os hubs semânticos indicam que regiões corticais interconectadas atuam como áreas de integração entre diferentes tipos de informações semânticas (Binder, 2011). Duas ênfases aparecem na literatura: **(i)** modelo *Hub-and-Spokes*, no qual um hub heteromodal, comumente localizado no lobo temporal anterior, integra atributos provenientes de sistemas sensório-motores; e **(ii)** modelos de circuitos distribuídos, nos quais múltiplas zonas de convergência apresentam preferências por determinadas categorias ou modalidades (Pulvermüller, 2013).

No primeiro modelo, *Hub-and-Spokes*¹⁴ (Patterson; Nestor; Rogers, 2007; Patterson, 2016), postula-se a existência de um hub heteromodal central, localizado no lobo temporal anterior, responsável por integrar informações provenientes de *spokes* sensório-motores distribuídos. No segundo, a representação semântica é distribuída em várias áreas neuroanatômicas, as quais apresentam preferências de domínio, i.e., contribuem diferencialmente para tipos específicos de processamento semântico, como cor, tamanho, cheiro, sabor ou tipo de objeto. Pulvermüller (2013) argumenta que várias regiões corticais funcionam como zonas de convergência para informações semânticas de naturezas distintas.

Nas palavras de Pulvermüller (2013, p. 460, tradução da autora):

¹⁴De acordo com Patterson (2016, p. 765), o modelo *Hub-and-Spoke* propõe que a memória semântica seja organizada por um sistema distribuído. Nesse sistema, regiões corticais específicas de modalidade (*spokes*) representam aspectos sensoriais, motores e linguísticos dos conceitos, enquanto um núcleo integrador amodal (*hub*) combina essas informações, abstrai das propriedades específicas de cada modalidade, i.e., não depende de como o conceito é percebido, executado ou nomeado e codifica a estrutura de similaridade semântica entre os conceitos.

Todos os possíveis hubs semânticos parecem apresentar, pelo menos em algum grau, especificidade em relação a categorias. Efeitos semânticos específicos de categoria também são observados em regiões que vão muito além dos candidatos a *hubs*, situando-se dentro e próximas de áreas sensoriais e motoras específicas de modalidade – ou, mais precisamente, preferenciais a determinada modalidade.

Evidências de patologias neurológicas graves de pacientes com algum tipo de *déficit* semântico (Brennan, 2022; Patterson, 2007) têm apontado que diversas regiões corticais interligadas no cérebro subjazem ao processamento semântico e lexical. Na demência semântica, caracterizada por atrofia bilateral (muitas vezes assimétrica) dos lobos temporais anteriores (LTA), há perda progressiva do conhecimento conceitual, tanto para itens que tenham mais atributos sensório-perceptuais, quanto a sua reversão (vantagem relativa para conceitos mais abstratos), dependendo do tipo de tarefa, do nível de especificidade e do grau de dependência de atributos sensório-perceptuais (Patterson; Nestor; Rogers, 2007). Esse tipo de comprometimento é compatível com a noção do efeito de concretude de Paivio (1991).

Em outros quadros de *déficits* semânticos, como nas afasias, há duplas dissociações no processamento de itens lexicais concretos e abstratos. Essas ocorrências sugerem que conceitos concretos que fazem referência a objetos mais facilmente identificados pelos sistemas sensoriais dependem mais de representações distribuídas em regiões sensoriais e temporo-occipitais. Em oposição, conceitos mais abstratos recrutam mais domínios linguístico-associativos e executivos, como o giro frontal inferior esquerdo e redes de integração de alto nível (Patterson; Nestor; Rogers, 2007; Schwanenflugel, 1986). Essa literatura tem apontado uma arquitetura neurocognitiva híbrida, em que há a ativação distribuída em *Hubs-and-Spokes* de domínios sensoriais e linguístico-associativos no processamento de itens lexicais concretos e abstratos.

Este capítulo teve como objetivo apresentar uma breve perspectiva histórica das primeiras formulações sobre o significado e a noção de concretude e abstração, problematizando a questão da referencialidade do significado, da lógica da materialidade, da percepção imediata de conceitos e de sua natureza psicofisiológica. Essa revisão aponta para uma mudança de paradigma, que parte de elaborações filosóficas de cunho especulativo e teórico em direção a uma investigação empírica ancorada em evidências comportamentais, voltada aos estudos linguísticos, como é o caso desta pesquisa.

Por fim, é interessante perceber como essa perspectiva, que vai da ontologia dos conceitos, entendidos como entidades mentais ou como convenções sociais, até as arquiteturas neurocognitivas, amodais ou híbridas, aponta para um caso-limite informativo a respeito da

dicotomia concreto *vs.* abstrato: os itens funcionais. Ao mesmo tempo em que exibem máxima abstração referencial porque não “apontam” para objetos no mundo, eles são densamente constrangidos por propriedades formais da língua (distribuição, prosódia, morfologia), figurando como candidatos ideais para testar a articulação entre representações amodais e circuitos sensório-motores.

À luz desta perspectiva, o próximo capítulo apresenta evidências psicolinguísticas de sensibilidade precoce aos itens funcionais. Veremos que sua alta previsibilidade e baixa carga imagética servem como pistas para a aquisição de padrões sintáticos e para o mapeamento ulterior entre forma e significado pelas crianças. É nesse cruzamento entre a ontologia dos conceitos e a aquisição da linguagem que a discussão subsequente se insere.

3 A SENSIBILIDADE PRECOCE AOS ITENS FUNCIONAIS: EVIDÊNCIAS PSICOLINGÜÍSTICAS

Há uma ampla literatura psicolinguística a respeito da sensibilidade precoce de bebês aos itens funcionais e de como esses elementos servem como pistas para identificar e adquirir a sintaxe da língua à qual estão expostos. Curiosamente, a aquisição desses itens funcionais pelas crianças não se constitui como um desafio, mesmo que esses elementos não possuam uma correspondência direta no mundo físico. Elementos funcionais, como artigos, preposições e pronomes, apesar de sua natureza abstrata e de serem adquiridos precocemente, desempenham um papel essencial na estruturação da linguagem, distinguindo-se dos itens lexicais por não apresentarem referência direta a objetos ou entidades do mundo.

A principal questão que emerge desse fenômeno e que será tratada neste capítulo é: como os bebês adquirem itens lexicais? Seriam os itens funcionais adquiridos primeiro? Será que esses itens podem ajudar os bebês a adquirir itens lexicais? Diferentemente dos nomes concretos, que podem ser ancorados em estímulos visuais e perceptuais do ambiente, os itens funcionais não possuem uma referência direta no mundo físico, mesmo assim, eles são adquiridos de maneira precoce e sem nenhuma dificuldade aparente, sugerindo que as crianças utilizam mecanismos sofisticados de extração de regularidades de padrões estatísticos da fala e usam esses dados no reconhecimento de padrões aos quais estão expostas (Saffran *et al.*, 1996; Saffran, 2001).

Evidências psicolinguísticas indicam que, antes do segundo ano de vida, os bebês já são sensíveis às diferenças fonológicas e distribucionais (Shady, 1996; Shi *et al.*, 1999; Shi; Werker, 2001), prosódicas e morfofonológicas (De Carvalho *et al.*, 2019), estatísticas e probabilísticas (Saffran *et al.*, 1996; Saffran, 2001) dos itens funcionais¹⁵ das línguas às quais estão expostos. Os bebês também identificam a frequência com que certos elementos ocorrem e suas posições nas sentenças, inferindo que esses itens possuem funções estruturais importantes. Além disso, a prosódia e as pistas fonológicas dos itens funcionais (De Carvalho *et al.*, 2019) podem facilitar a detecção e a categorização de palavras, mesmo na ausência de um referente concreto.

¹⁵De acordo com Shi (2013, p. 3, tradução da autora), “os morfemas funcionais constituem um conjunto pequeno de itens, cada um deles extremamente frequente. No que diz respeito às propriedades sonoras, em diferentes línguas, os elementos funcionais apresentam forma fonética reduzida na fala. Eles possuem menos sílabas por palavra e menos consoantes e vogais por sílaba do que os elementos lexicais, além de tenderem a uma prosódia fraca (por exemplo, menor intensidade, menor duração e menor proeminência de altura; Shi, 1996; Shi *et al.*, 1998)”.

Essa sensibilidade precoce em adquirir os itens funcionais tem implicações importantes para o estudo da aquisição de itens abstratos de forma geral. Como já falamos, enquanto os nomes concretos podem ser mais facilmente aprendidos por meio de uma associação direta entre palavra e objeto designado, os itens abstratos, por sua vez, deveriam exigir uma abordagem cognitiva mais complexa a qual pode envolver raciocínio inferencial e modelos de suporte conceituais (*conceptual scaffolding*) para serem acessadas (Crutch *et al.*, 2009).

A respeito disso, Shi *et al.* (1999) investigaram como recém-nascidos, entre 1 e 3 dias de vida, utilizam pistas acústicas para distinguir entre itens lexicais (como nomes e verbos) e funcionais (como artigos e preposições) e adquirir novas palavras. Participaram do estudo 77 recém-nascidos saudáveis, com idades entre 1 e 3 dias, dos quais 32 completaram o experimento. Entre os participantes, 17 tinham mães falantes nativas de inglês e 15 tinham mães que falavam outras línguas.

Os bebês foram submetidos a um experimento que utilizou o paradigma de sucção de alta amplitude (HASP — *High Amplitude Sucking Paradigm*). Nesse método, os bebês ouviam palavras enquanto sugavam um bico de chupeta conectado a um dispositivo que registrava a intensidade da sucção. A recuperação da taxa de sucção após a habituação foi usada como medida da capacidade de discriminação. Os estímulos incluíram itens lexicais (e.g., *toys, play*) e funcionais (e.g., *the, in*), extraídos de gravações de falantes nativos de inglês e regravadas para preservar suas características acústicas.

O *design* experimental foi feito em duas fases. Na fase de habituação, os bebês ouviam repetidamente uma lista de palavras de uma única categoria, como itens lexicais (*toys, play, chew*) ou itens funcionais (*the, in, your*), até que a taxa de sucção diminuísse, indicando familiarização. Na fase de teste, os bebês eram apresentados a uma nova lista de palavras. No grupo controle, a lista de teste continha palavras da mesma categoria apresentada na habituação (p. ex., itens lexicais seguidos por mais itens lexicais). No grupo experimental, no entanto, a lista de teste continha palavras de uma categoria diferente (e.g., itens lexicais seguidos por itens funcionais). Os pesquisadores previram que os bebês no grupo experimental apresentariam maior recuperação da taxa de sucção, refletindo sua capacidade de discriminar entre as duas categorias de palavras.

Os resultados de Shi *et al.* (1999) apontam que os bebês no grupo experimental mostraram uma recuperação significativa na taxa de sucção ao serem expostos a listas de uma categoria diferente, enquanto no grupo controle essa recuperação não foi significativa. Um segundo experimento foi realizado para controlar fatores como número de sílabas e proporção

de *types* e *tokens*, e os achados foram replicados. Esses resultados indicam que recém-nascidos possuem sensibilidade a pistas perceptuais que diferenciam itens lexicais e gramaticais com base nas características acústicas desses elementos. Esses achados são relevantes, pois reforçam a ideia de que as propriedades acústicas do *input* linguístico servem como pistas para que os bebês possam adquirir itens lexicais.

Shi e Werker (2001), em uma série de três experimentos investigaram se bebês aos seis meses de vida, expostos à língua inglesa, não só seriam capazes de discriminar categoricamente itens lexicais e funcionais, mas se utilizariam essas pistas como forma de adquirir a sintaxe da língua. Além disso, as autoras verificaram também se os bebês apresentariam uma predileção por algum desses itens.

No primeiro experimento, Shi e Werker (2001) utilizaram um paradigma de *habituation-dishabituation* para verificar se bebês de seis meses discriminam categoricamente elementos lexicais e funcionais. Para isso, foi adotado o paradigma de fixação visual. Participaram dessa etapa 32 bebês a termo, saudáveis, falantes nativos de língua inglesa, que haviam composto o grupo experimental do primeiro estudo com recém-nascidos (Shi, Morgan, 1999) e que foram convidados a participar novamente ao completarem seis meses de vida.

Os estímulos auditivos utilizados no teste foram os mesmos de Shi *et al.* (1999), ou seja, duas listas de elementos lexicais (por exemplo, as palavras *chew*, *hide*, *chair*) e duas listas de elementos funcionais (como *the*, *you*, *that*), selecionadas aleatoriamente a partir de uma gravação de fala espontânea entre mães e seus bebês. Cada lista continha entre 38 e 42 palavras. Na fase de habituação, metade dos bebês ouviu itens funcionais, enquanto a outra metade foi exposta a elementos lexicais. Na fase de teste, cada participante foi apresentado a duas novas listas de palavras: uma nova lista da mesma categoria ouvida durante a habituação (*same trial*) e outra pertencente à categoria oposta (*switch trial*). Todos os estímulos, desde a fase de habituação até as duas condições de teste, foram contrabalanceados.

A previsão das autoras era de que, se bebês de seis meses ainda fossem capazes de categorizar itens lexicais e funcionais com base nas propriedades fonológicas e morfossintáticas, eles apresentariam maior tempo de fixação na condição *switch trial* do que na condição *same trial*, durante a fase de teste. Não houve diferença significativa no tempo de fixação entre as condições durante a fase de habituação, ou seja, os bebês, independentemente de estarem na condição de itens lexicais ou funcionais, tiveram o mesmo tempo de exposição na fase de familiarização.

Uma ANOVA com design fatorial 2 (condição de habituação: categoria lexical vs. funcional) x 3 (*trial type*: final da habituação vs. *same trial* vs. *switch trial*) foi realizada para verificar o tempo de fixação durante o bloco final de habituação (a média das três últimas tentativas de habituação) e em cada uma das tentativas de teste. Essa análise apontou efeitos principais para ambas as condições de habituação e para o tipo de *trial* (*same* vs. *switch*), e de uma interação significativa entre condição de habituação e tipo de *trial*. Além disso, houve uma diferença significativa no tempo de fixação entre o *switch trial*, o *same trial* e o bloco final de habituação, para bebês que haviam sido habituados a itens funcionais.

Diferentemente do estudo de Shi e Morgan (1999), em que os bebês aumentaram a taxa de sucção independentemente da condição de habituação, os resultados desse primeiro experimento apontam que, aos seis meses, os bebês ouviram por mais tempo os itens lexicais após serem habituados a itens funcionais. Contudo, o inverso não ocorreu: os bebês que ouviram itens lexicais na fase de habituação não demonstraram predileção pelos itens funcionais. De acordo com Shi e Werker (2001), uma explicação plausível para essa preferência é que os itens lexicais são mais salientes que os itens funcionais, tanto na forma fonológica e acústica quanto no significado. Por isso, eles podem ser mais interessantes para os bebês que os itens funcionais.

Com o objetivo de confirmar o resultado de que bebês de seis meses, adquirindo a língua inglesa, têm predileção por itens lexicais em detrimento dos funcionais, Shi e Werker (2001) elaboraram um segundo experimento. Diferentemente do primeiro, os doze bebês participantes foram testados utilizando o paradigma de preferência de fixação central (*central-fixation preference procedure*), com duas listas de palavras em vez de quatro. As outras duas listas de cada tipo de item (lexical e funcional) foram combinadas. Assim, dez *trials* foram apresentados de forma aleatória, alternando cinco *trials* para cada tipo de item.

Para verificar se havia uma predileção por um tipo de item em detrimento do outro, Shi e Werker (2001) compararam o tempo de fixação do olhar de cada bebê para cada tipo de palavra (as funcionais e as lexicais). Os tempos médios de fixação para cada lista foram calculados individualmente. Uma ANOVA mista com *design* fatorial 2 x 2 foi conduzida, tendo como fatores a ordem de apresentação das listas (*between-subjects*: primeiro lexical vs. primeiro funcional) e a categoria de palavras (*within-subjects*: lexicais vs. funcionais). Os resultados indicaram um efeito significativo do tipo de palavra, com um tempo médio de fixação maior nos itens lexicais do que nos itens funcionais. Outra ANOVA, removendo o primeiro *trial*, confirmou o efeito principal para a categoria de palavra. Dessa forma, esses

dados corroboram os resultados do primeiro experimento quanto à preferência dos bebês pelos itens lexicais.

Os experimentos conduzidos por Shi e Werker (2001) apontam para a sensibilidade dos bebês aos itens funcionais, reforçando a ideia de que esses elementos servem como *andaimes* na segmentação do contínuo sonoro, na delimitação dos itens lexicais e, portanto, na capacidade dos bebês de adquirir a estrutura sintática da língua. Além disso, os resultados indicam que, entre os dois tipos de itens, os bebês aos seis meses de idade já demonstram uma preferência pelos lexicais em detrimento dos funcionais.

Outros pesquisadores se propuseram a explorar a sensibilidade de bebês às propriedades fonológicas e distribucionais, Shady (1996) conduziu um experimento com 24 bebês, adquirindo a língua inglesa, com 10,5 meses de idade, utilizando a técnica de Escuta Preferencial. O objetivo era verificar se os bebês diferenciavam itens funcionais reais de pseudoitens funcionais. Durante a fase de teste, foram apresentadas pequenas sentenças extraídas do livro infantil *The Curious Little Kitten* em duas versões: uma original, em que os itens funcionais (os verbos auxiliares *is* e *was*, determinantes *the* e *that*, e preposições *of* e *with*) eram mantidos, e outra modificada, na qual esses itens foram substituídos por pseudoitens. A análise revelou que os bebês escutaram por mais tempo as sentenças originais em comparação com a versão modificada. Esses dados sugerem que bebês de 10,5 meses possuem sensibilidade aos itens funcionais e demonstram preferência por estruturas nas quais esses elementos estão preservados.

Em um segundo experimento, Shady (1996) buscou investigar se essa sensibilidade se estendia também aos elementos lexicais ou se era específica aos itens funcionais. Para isso, 45% dos elementos lexicais de uma história foram substituídos por pseudoitens, enquanto os itens funcionais permaneceram intactos. Os resultados indicaram tempos médios de escuta muito próximos entre a versão original e a modificada, sem diferença estatística entre as condições. Assim, os achados apontam que bebês com 10,5 meses respondem seletivamente aos itens funcionais, não mostrando rejeição aos pseudoitens lexicais.

No terceiro experimento, Shady (1996) investigou se os bebês também eram sensíveis à posição estrutural dos itens funcionais nas sentenças. Foram comparadas uma versão não alterada e uma modificada, na qual a posição dos itens funcionais foi trocada. Por exemplo, a sentença original “*This man has bought two cakes*” foi modificada para “*Has man this bought two cakes*”. Os bebês de 10,5 meses não demonstraram preferência significativa entre as duas versões. Ao repetir o experimento com crianças de 13 meses, os resultados foram similares.

Contudo, bebês de 16 meses apresentaram uma resposta diferente: o tempo de escuta foi significativamente maior para a versão original do que para a modificada. Esses resultados indicam que a sensibilidade à posição estrutural dos itens funcionais parece emergir mais tarde no desenvolvimento. Enquanto bebês de 10,5 e 13 meses não demonstraram uma preferência direta, o desempenho de crianças de 16 meses foi estatisticamente significativo para as sentenças em que a estrutura sintática estava preservada.

Em 1996, Saffran *et al* conduziram um estudo pioneiro que investigou se bebês de 8 meses conseguem identificar fronteiras entre palavras em um fluxo contínuo de fala. A pesquisa revelou que, mesmo sem pistas acústicas explícitas, como pausas ou entonações, os bebês podem usar informações estatísticas para segmentar palavras. Para testar essa hipótese, os pesquisadores desenvolveram uma língua artificial composta por quatro pseudopalavras trissilábicas: *bi-da-ku*, *pa-do-ti*, *tu-pi-ro* e *go-la-bu*. Essas pseudopalavras foram combinadas e apresentadas em ordem aleatória, formando um fluxo de fala contínuo, por exemplo: *bidakupadotigolabu*. No fluxo, as probabilidades de transição entre sílabas dentro de palavras eram altas (1.0, como entre *bi* e *da* em *bi-da-ku*), enquanto as probabilidades entre fronteiras de palavras eram mais baixas (0.33, como entre *ku* e *pa* em *ku-pa-do*).

Para verificar a sensibilidade dos bebês a essas possíveis transições, Saffran *et al.* (1996) conduziram dois experimentos com 24 bebês em cada. No primeiro experimento, os estímulos de teste incluíam palavras familiares, ou seja, presentes no fluxo de fala, e pseudopalavras, combinações de sílabas que nunca haviam sido apresentadas juntas (por exemplo, *da-pi-ku*). Os resultados apontaram que os bebês passaram mais tempo escutando as pseudopalavras do que as palavras familiares, com uma diferença estatisticamente significativa. Esses dados indicam que os bebês reconheceram as palavras familiares do fluxo e foram capazes de distingui-las das pseudopalavras, demonstrando preferência por estímulos novos.

No segundo experimento, os estímulos incluíam palavras familiares e *part-words*, que eram sequências de sílabas cruzando fronteiras de palavras no fluxo (por exemplo, *ku-pa-do*). Apesar de as *part-words* incluírem sílabas familiares, elas eram menos prováveis estatisticamente devido à baixa transição entre fronteiras de palavras. Os resultados apontaram que os bebês passaram mais tempo escutando as *part-words* do que as palavras familiares, com diferença estatisticamente significativa. O estudo de Saffran *et al.* (1996) apontou que bebês de 8 meses são capazes de segmentar palavras em um fluxo contínuo de fala utilizando apenas pistas estatísticas.

Em um outro trabalho, Saffran (2001) investigou, a partir do *Head-Turn Preference Procedure*, se bebês ingleses de 8 meses utilizam pistas estatísticas para segmentar pseudopalavras em um fluxo contínuo de fala e se elas seriam tratadas como possíveis palavras da língua inglesa. Participaram do estudo 56 bebês monolíngues de 8 meses. Eles foram divididos aleatoriamente em dois grupos, sendo um exposto a frases em inglês e o outro a frases sem sentido. Na fase de familiarização, todos os bebês ouviram um fluxo contínuo de fala artificial, composto por palavras como *pabiku*, *golatu*, *daropi* e *tilado*. Essas pseudopalavras foram concatenadas sem pausas ou variações prosódicas, formando sequências uníssonas como *pabikugolatudaropitilado....* O fluxo foi estruturado para que as probabilidades de transição entre sílabas fossem altas dentro das palavras (1.0, como entre *pa* e *bi* em *pabiku*) e baixas entre palavras (0.33, como entre *ku* e *go* em *pabiku golatu*).

No primeiro experimento, as pseudopalavras aprendidas e as *part-words* (como *kupabi*) foram inseridas em sentenças em inglês, como *I like my golabu*. O objetivo era verificar se os bebês tratavam as pseudopalavras aprendidas como itens lexicais, já que elas possuíam propriedades morfofonológicas da língua inglesa, como o uso de itens funcionais, como *I like* e pistas prosódicas naturais. Os resultados apontaram que 21 dos 28 bebês demonstraram preferência significativa pelas palavras aprendidas, observando-as por mais tempo que as *part-words*. Isso sugere que os bebês trataram as pseudopalavras como possíveis palavras do inglês.

No segundo experimento, as mesmas pseudopalavras aprendidas e *part-words* foram inseridas em sentenças sem sentido, como *Zy fike ny golabu*. O objetivo era avaliar se os bebês poderiam identificar e discriminar as pseudopalavras na ausência de propriedades morfofonológicas familiares, como os itens funcionais ou pistas prosódicas. Nessa condição, os bebês não demonstraram preferência significativa entre os estímulos. Esses resultados indicam que, sem as outras pistas morfofisiológicas e prosódicas, apenas as informações estatísticas não são suficientes para que os bebês tratem a sentença como um estímulo linguístico.

O terceiro experimento replicou o primeiro, mas substituiu os *frames* de sentenças sem sentido por *frames* compostos por sequências com padrões prosódicos não linguísticos. Nesse caso, palavras segmentadas como *tiba* foram testadas em *frames* que apresentavam sequências de tons organizados segundo padrões regulares de duração e frequência, mas sem estrutura gramatical. Assim, a palavra *tiba* era apresentada em meio a padrões prosódicos artificiais, sem pistas morfosintáticas ou prosódicas do inglês. O objetivo era verificar se esses elementos poderiam influenciar no desempenho dos bebês. Novamente, foi encontrada uma

interação significativa entre o tipo de *frame* e os itens. Em *frames* compostos por sentenças em inglês, os bebês preferiram as pseudopalavras a *part-words*. No entanto, não houve diferença significativa em *frames* prosódicos não linguísticos, sugerindo que o aprendizado estatístico é sensível ao domínio em que é aplicado.

Esse resultado é interessante, porque sugere que o aprendizado estatístico, quando integrado a propriedades fonotáticas e morfossintáticas do inglês, foi tratado pelos bebês como possíveis itens lexicais. Por outro lado, nos *frames* prosódicos não linguísticos, que não tinham estrutura gramatical ou fonotática preservada, os bebês não mostraram diferença na preferência entre as pseudopalavras e *part-words*. Isso indica que, fora do domínio da linguagem (como em padrões sonoros abstratos), os resultados do aprendizado estatístico não parecem ser um mecanismo utilizado pelos bebês. Em síntese, os resultados de Saffran (2001) sugerem que propriedades estatísticas e probabilísticas constituem um mecanismo essencial que, nas etapas iniciais da aquisição linguística, é utilizado pelas crianças para identificar e para adquirir novos itens lexicais.

Ainda em relação às pistas perceptuais como forma de interpretar o significado e a categoria sintática de uma palavra, De Carvalho *et al.* (2019) investigaram, em dois experimentos, a sensibilidade de bebês franceses aos itens funcionais e às fronteiras prosódicas. Os resultados demonstraram que, aos 18 meses, os bebês não apenas utilizam elementos funcionais em uma sentença para determinar a categoria sintática de uma palavra, mas também recorrem à prosódia frasal, em combinação com os itens funcionais, para mapear a estrutura sintática e interpretar uma pseudopalavra como substantivo ou verbo, dependendo de sua posição na estrutura prosódica da sentença.

Dois experimentos foram conduzidos para avaliar se os bebês eram capazes de usar exclusivamente itens funcionais (Experimento 1) e uma combinação de itens funcionais e prosódia (Experimento 2) para diferenciar nomes de verbos. Participaram 48 bebês, adquirindo o francês, em uma tarefa com o paradigma de habituação-troca para familiarizá-los com vídeos de um pinguim realizando duas ações distintas, como girar e dar cambalhotas. Cada vídeo foi associado a áudios que introduziam as pseudopalavras ora como verbo ora como substantivo. Por exemplo, na condição de substantivo, os bebês ouviam frases como “*Oh regarde! C’est une bamoule!*” (“Olha! É uma bamoule!”), em que *bamoule* fazia referência ao pinguim.

Em contextos de verbo, as sentenças eram: “*Regarde! Elle doripe!*” (Olha! Ela está ‘doripando’), referindo-se à ação de dar cambalhotas. As combinações palavra-objeto e palavra-ação foram contrabalanceadas entre os participantes. Assim, durante o teste, na

condição em que havia troca de substantivo (*noun-switch*), o bebê ouvia uma frase como “*Oh regarde! C’est une bamoule!*” e assistia a um vídeo que mostrava o pinguim realizando uma ação errada, como girar quando deveria estar dando cambalhota. Já na condição em que o verbo era trocado (*verb-switch*), ele ouvia uma frase como “*Regarde! Elle doripe!*”, junto com um vídeo em que o pinguim realizava uma ação incompatível, como girar, ao invés de dar cambalhotas.

Os resultados indicaram que os bebês apresentaram maior tempo de fixação na condição de troca de verbo em comparação com a condição de troca de substantivo. Isso sugere que os bebês utilizaram itens funcionais para atribuir categorias sintáticas e restringir os significados das palavras novas. Uma ANOVA confirmou uma interação significativa entre condição e fase, sugerindo que bebês usaram os itens funcionais e a distribuição sintática para restringir o significado da pseudopalavra. Isso porque eles perceberam mais claramente a inconsistência quando a troca envolvia verbos (ações). Diante dessas evidências, o trabalho de De Carvalho *et al.* (2019) fornece evidências de que bebês franceses, já aos 18 meses de idade, não só são sensíveis aos itens funcionais, à prosódia frasal e à análise distributiva de itens lexicais, como também podem utilizar essas pistas no processo de aquisição de novas palavras.

Conforme visto na revisão acima, diversos são os estudos psicolinguísticos, realizados em diferentes línguas, que apontam que bebês, antes do primeiro ano de vida, já são sensíveis aos chamados itens funcionais (Cauvet *et al.*, 2013; Shi *et al.*, 1999) da língua, i.e., itens que cumprem apenas funções gramaticais e que não fazem, portanto, referência a algo físico no mundo. Nas palavras de Shi e Werker (2001, p. 70, tradução da autora): “A distinção binária entre itens lexicais e itens funcionais reflete essa importante diferença entre palavras que carregam significado e aquelas que contribuem principalmente para as relações estruturais.”.

No entanto, ainda que esses elementos pudessem ser mais difíceis de serem adquiridos por bebês, eles emergem no vocabulário infantil sem que eles apresentem nenhum tipo de dificuldade ou que sejam expostos a algum tipo de instrução explícita. Se esses itens fossem mais difíceis de serem adquiridos por bebês, uma vez que não possuem uma contraparte física no mundo real, seria factível propor, então, que a aquisição lexical estaria restrita aos itens lexicais concretos que fazem referência direta a um elemento físico no mundo e cuja referência poderia ser mais facilmente inferida pela criança, a partir de um contexto de cena.

Nessa perspectiva, a aquisição lexical não se reduz a uma associação direta entre sons/palavras e objetos no mundo, mas envolve a construção conceitual que torna esse mapeamento possível, mesmo quando não há referência material ou direta desse item no

mundo. Assim, a criança, ainda no primeiro ano de vida, precisa: “segmentar a palavra relevante no fluxo contínuo da fala, identificá-la como nome no contexto sintático em que se encontra; identificar a entidade relevante da atividade em curso e estabelecer uma correspondência palavra-mundo entre elas.” (Waxman, 2006, p. 133, tradução da autora).

Essa capacidade de estabelecer correspondências palavra-mundo, embora essencial, é apenas o início de um processo cognitivo complexo. Para lidar com palavras que não possuem uma representação física direta, como muitos itens funcionais e até mesmo palavras abstratas, os bebês precisam ultrapassar essa correspondência palavra-objeto para estabelecer correspondências palavra-conceito.

A criança, então, não só precisa extrair a palavra do contínuo sonoro, como também mapear o referente que ela designa. Pelo fato de a língua não se limitar a objetos concretos e materiais do mundo físico, o bebê precisa formar correspondências mais abstratas, como palavra-conceito. Esse processo envolve uma transição do nível perceptivo-sensorial para um mapeamento mais associativo e abstrato, no qual sons são conectados a conceitos que não possuem uma representação física direta.

Em síntese, se, por um lado, a história das ideias mostrou que a significação não se reduz à materialidade do referente, por outro, a evidência psicolinguística contemporânea sugere que a previsibilidade distributiva e a organização sintática, em que os itens funcionais desempenham papel central, estruturam o acesso lexical e a composição de sentido. Assim, itens como determinantes, preposições, clíticos e morfemas flexionais tornam-se uma janela empírica para observar como o sistema cognitivo mapeia pistas formais em representações conceituais, mesmo quando a “concretude” é mínima.

Essa transição teórica tem consequência metodológica direta: se a aquisição de conceitos concretos e abstratos depende, em graus distintos, de redes semânticas distribuídas e de pistas linguísticas, então a sensibilidade precoce a elementos de classe fechada deve emergir como um marcador robusto da arquitetura gramatical em formação. Em outras palavras, antes mesmo de consolidarem rótulos estáveis para entidades do mundo, bebês já podem explorar regularidades fonológicas, prosódicas e estatísticas associadas aos funcionais para inferir limites de palavras, categorias e esboços de estrutura sintática.

É precisamente nesse ponto que a discussão deixa o plano histórico-conceitual e muda para o experimental: investigar, com dados comportamentais e neurais, como e quando a cognição em desenvolvimento detecta e utiliza itens funcionais como “âncoras” de segmentação e como “moldes” composicionais, para o entendimento completo dos itens concretos e abstratos. Essa questão será abordada em 3.1.

3.1 A AQUISIÇÃO LEXICAL DE ITENS CONCRETOS E ABSTRATOS

Como bebês, diante de um ambiente altamente estimulante, em que todos os sons poderiam, a princípio, ser iguais ou transmitir a mesma mensagem, atribuem sentidos a esses estímulos de forma tão natural, rápida e eficiente? Atribuir sentido aos sons, mesmo utilizando os itens funcionais como pistas para chegar ao nível sintático-semântico, não seria um desafio, sobretudo se pensarmos na variabilidade de itens lexicais e dos significados que eles podem ter?

Será que existiriam, dentro dessa mesma classe conceitual de palavras, algumas mais complexas de estabelecer uma referência direta no mundo, como quando o bebê escuta algo como “o carrinho é muito leve”, em que a palavra “carrinho” poderia ser mais facilmente compreendido pela criança do que “leve”? Visto que o significado de um item lexical possui uma natureza psicológica, subjetiva e idiossincrática (pois cada pessoa possui uma representação mental única de um termo), eles também se constituem como representações mentais relativamente estáveis e objetivas do ponto de vista comunicativo, já que são comunicáveis e compreendidas socialmente.

Essa questão sobre o que é relativamente mais estável é essencial para compreender como conceitos são formados e torna-se ainda mais instigante ao se refletir sobre as etapas iniciais da aquisição da linguagem, uma vez que a criança, exposta a uma infinidade de estímulos, precisa caracterizar o mundo em suas formas linguísticas para ancorar o pensamento que se complexifica rapidamente.

Nesta perspectiva, Gleitman *et al.* (2005) e Gleitman e Trueswell (2020) debateram essas questões ao apontar que a dificuldade na aquisição lexical não está apenas na complexidade conceitual das palavras, mas na maneira como seus significados podem (ou não) ser inferidos a partir do contexto. A autora propõe a distinção entre *Easy Words*, como nomes concretos com referência direta e observável no mundo físico, e *Hard Words*, como conceitos abstratos ou estados mentais (pensar, saber, justo) sem contraparte física no mundo. Para Gleitman, a aquisição de nomes concretos pode ser mais fácil para as crianças, porque a associação palavra-objeto é direta, enquanto a de termos mais abstratos dependeria de outros mecanismos para a aquisição, como o *Bootstrapping* Sintático¹⁶ (Gleitman, 1990).

¹⁶A hipótese do *Bootstrapping* Sintático (Gleitman, 1990) prevê maior importância à informação sintática do que à informação de natureza conceitual na atribuição de sentido a um enunciado pela criança. Para mais informações a respeito, recomenda-se a leitura aqui referenciada.

Embora a discussão sobre quais propriedades os bebês são mais sensíveis e quais pistas — como itens funcionais, fronteiras prosódicas, propriedades estatísticas e probabilísticas — são utilizadas para a aquisição lexical ou sintática já tenha sido amplamente explorada na literatura psicolinguística, destaca-se aqui que a capacidade precoce de crianças em abstrair propriedades fônicas e distribucionais a partir dos itens funcionais não parece representar um obstáculo e, ao contrário, pode atuar como um guia na aquisição de itens lexicais e da estrutura sintática, mesmo diante de uma preferência inicial por itens lexicais concretos. Essa habilidade sugere que, apesar de ainda não dominarem um repertório lexical vasto de termos abstratos, os bebês já são capazes de utilizar essas pistas linguísticas pertencentes a categorias sintáticas distintas e atribuir significado a eles.

Em relação aos itens lexicais, propriamente, poucos estudos forneceram evidências sobre o processo de aquisição de itens concretos e abstratos. Um dos estudos pioneiros sobre o tópico foi o de Tincoff e Jusczyk (1999), que explorou a capacidade de associação específica e compreensão de itens lexicais por bebês americanos aos seis meses de idade. Utilizando o paradigma de Fixação Preferencial do Olhar Intermodal, do inglês *Intermodal Preferential-Looking Paradigm* (Name; Corrêa, 2018), 24 bebês foram expostos a vídeos nos quais apareciam seus próprios pais enquanto ouviam as palavras *mamãe* e *papai*.

O experimento foi dividido em três fases: (1) quatro tentativas de pré-teste silencioso, em que os bebês eram familiarizados com as localizações dos vídeos de cada pai em monitores distintos; (2) quatro tentativas de *baseline*, também sem áudio, para avaliar possíveis preferências visuais iniciais; e (3) oito tentativas de teste com estímulos auditivos (quatro apresentações da palavra *mamãe* e quatro da palavra *papai*, distribuídas aleatoriamente em dois blocos de teste). Os áudios foram gerados por um sintetizador de voz que simulava a voz de uma criança sem características de gênero, garantindo controle sobre possíveis influências externas. Cada tentativa teve 10 segundos de duração.

Os resultados do primeiro experimento indicaram que os bebês olharam por mais tempo para os vídeos dos pais nomeados no áudio, sugerindo que eles foram capazes de associar corretamente o som à imagem correspondente. O tempo de fixação visual ajustado pela linha de base foi analisado por meio de uma ANOVA de medidas repetidas. Observou-se que 18 dos 24 bebês fixaram o olhar por mais tempo nos vídeos dos pais nomeados em comparação aos vídeos não nomeados, com resultados estatisticamente significativos.

Para verificar se os bebês realmente compreendiam os termos *mamãe* e *papai* e os associavam especificamente aos próprios pais, foi realizado um segundo experimento. Outros 24 bebês, da mesma faixa etária, assistiram aos vídeos gravados no primeiro experimento (ou

seja, com os pais do primeiro grupo), enquanto ouviam as palavras *mamãe* e *papai*. Esse procedimento tinha o objetivo de investigar se, ao ouvirem essas palavras, os bebês generalizariam sua associação a outras mulheres e homens.

Os resultados do segundo experimento indicaram que os bebês não fixaram o olhar por mais tempo nos vídeos de pais não familiares nomeados. Uma ANOVA de medidas repetidas revelou que o efeito principal do tipo de vídeo não foi significativo. Ao comparar os resultados dos dois experimentos, observou-se que os bebês do Experimento 1 olharam por mais tempo para os vídeos nomeados em comparação aos bebês do Experimento 2. Esses resultados sugerem que, aos seis meses, os bebês compreendem e associam corretamente as palavras *mamãe* e *papai* aos seus próprios pais, sem generalizá-las para outras mulheres e homens. Assim, o estudo de Tincoff e Jusczyk (1999) demonstra que bebês já começam a formar associações entre padrões sonoros e seus respectivos significados aos seis meses de idade.

Em um estudo mais recente sobre a compreensão de itens concretos por bebês, Bergelson e Swingley (2012) investigaram, por meio de uma tarefa do tipo Fixação Preferencial do Olhar com rastreamento ocular, se bebês norte-americanos, também aos seis meses de idade, já compreendem palavras concretas como *maçã*, *bola* e *garrafa*. O experimento contou com 33 bebês, divididos em dois grupos etários: 6–7 meses e 8–9 meses. Todos os participantes foram expostos a palavras familiares do cotidiano e não receberam nenhum treinamento antes dos *trials* experimentais.

Os autores elaboraram dois tipos de condições experimentais: testes de cena e testes de imagens pareadas. Nos testes de cena, uma única imagem era apresentada no centro da tela, como uma mesa contendo quatro alimentos diferentes — *pão*, *banana*, *leite* e *maçã*. Durante a apresentação da imagem, os bebês ouviam uma sentença que designava um dos itens, como “Can you find the X?”, “Where’s the X?”, “Do you see the X?”, e “Look at the X!” (Você pode encontrar o X?”, “Onde está o X?”, “Você vê o X?” ou “Olhe para o X!”), em que X era a palavra-alvo (cada item foi associado a apenas um formato de frase). Cada bebê visualizou apenas uma imagem por vez. O objetivo desta condição era avaliar se os bebês conseguiam reconhecer palavras em contextos de cena mais complexos. Já nos testes de imagens pareadas, os bebês visualizaram simultaneamente duas imagens na tela: uma pertencente à categoria de alimentos e outra de partes do corpo.

As imagens, pareadas lado a lado, consistiam em uma *maçã* e uma *boca*, enquanto o bebê ouvia sentenças do tipo “Look at the apple” (Olhe para a maçã”) ou “Look at the mouth” (Olhe para a boca). O objetivo dessa condição era verificar se os bebês conseguiriam

identificar as duas imagens, as quais eram mutuamente exclusivas. Já nos *trials* de cena, os bebês viam uma única imagem chamada de “complexa”, pois continha vários objetos: um corpo, um rosto ou uma mesa em que estavam um dos quatro alimentos. Os bebês foram expostos a 48 *trials*, divididos em 32 *trials* de imagens pareadas e 16 *trials* de cena. Cada condição incluiu oito itens de alimentos e oito itens de partes do corpo. Todos os participantes foram expostos a ambos os tipos de *trial*, mas a ordem das apresentações foi pseudorandomizada, assim como a sequência das imagens e dos itens-alvo.

Como as duas condições avaliavam habilidades distintas, os resultados foram analisados separadamente devido às suas diferenças metodológicas. Nos *trials* de cena, o objetivo era avaliar a habilidade de generalização dos bebês em contextos visuais mais complexos, nos quais eles deveriam mapear a cena apresentada para identificar a palavra-alvo. Por outro lado, nos *trials* de imagens pareadas, seria mais fácil mapear se o bebê identificaria a palavra-alvo, já que só havia duas imagens, sendo uma o alvo e a outra que não tinha nenhuma relação com o estímulo auditivo.

Interessantemente, os resultados apontaram uma diferença entre os grupos etários dos bebês de 6–7 meses e os de 8–9 meses em relação à identificação das palavras concretas. Nos *trials* de imagem-pareadas, os bebês mais novos foram capazes de mapear a imagem-alvo, ainda que tenham ficado próximos ao nível da chance. Em contrapartida, o grupo mais velho apresentou um desempenho melhor, indicando uma maior capacidade no mapeamento do tipo palavra-objeto. Nos *trials* de cena, as médias do grupo de bebês de 6–7 meses indicaram valores referentes à fixação próximos ou abaixo de zero, sugerindo que as crianças, nessa faixa etária, não fixam tanto a imagem-alvo durante os *trials* de cena, similarmente ao que aconteceu na condição de imagens pareadas. Porém, o grupo de bebês de 8–9 meses teve um melhor desempenho na fixação da imagem-alvo nos *trials* de cena. Esses resultados indicam que, mesmo em contextos de cena mais desafiadores, os bebês mais velhos de 8–9 meses tiveram mais facilidade de reconhecer essas palavras em comparação aos de 6–7 meses.

Em suma, dos 33 bebês testados, 22 apresentaram uma proporção positiva de fixação na imagem-alvo, ou seja, os bebês foram capazes de identificar o objeto correto após a escuta do estímulo auditivo. A análise das médias, realizada a partir de um teste de *Wilcoxon*, revelou um efeito estatisticamente significativo. Esse dado é uma evidência favorável de que bebês entre 6 e 9 meses já são capazes de identificar palavras concretas, presentes em seus contextos diários, como as que fazem referência a itens alimentícios ou a partes do corpo. Ao analisar o desempenho em itens específicos (palavras individuais), também por meio de um teste de *Wilcoxon*, os bebês apresentaram índices acima do nível da chance em 12 dos 16 itens

experimentais. No entanto, conforme destacado pelos autores, o efeito de significância ficou próximo ao nível da chance.

Além dos bebês de 6 a 9 meses, Bergelson e Swingley (2012) testaram três grupos adicionais, compostos por crianças mais velhas, com o objetivo de investigar se o reconhecimento lexical após os 9 meses de idade, nos *trials* com imagens pareadas, era consistente com os resultados de estudos anteriores. O primeiro grupo foi composto por 30 crianças de 10 a 13 meses. O segundo grupo incluiu 7 crianças de 14 a 16 meses, enquanto o terceiro grupo consistiu em 13 crianças de 18 a 20 meses.

Analisando os quatro grupos etários: 6–9 meses, 10–13 meses, 14–16 meses e 18–20 meses, a fixação na imagem-alvo aumentou progressivamente após o início da palavra, indicando que os bebês foram capazes de associar a palavra escutada ao objeto correspondente. Interessantemente, foi observada uma correlação positiva entre o aumento da idade dos bebês e a fixação na imagem-alvo. As crianças de 18–20 meses apresentaram o maior aumento na fixação, enquanto as de 6–9 meses tiveram um desempenho pior, ainda que acima do nível de chance. Esses resultados sugerem que a habilidade de identificar palavras concretas é mais evidente no segundo ano de vida.

Em relação a itens lexicais mais abstratos, Bergelson e Swingley (2013) investigaram a capacidade de bebês entre 6 e 16 meses de idade de compreender palavras como *oi* (*hi*) e *acabou* (*all gone*), que apresentam maior dificuldade de representação em comparação a palavras mais concretas e imagináveis, como “maçã” (*apple*) e “boca” (*mouth*). As hipóteses do estudo eram de que bebês de 6 a 9 meses não conseguiriam associar palavras abstratas aos seus referentes, devido à necessidade de habilidades cognitivas e sociais mais avançadas para a aquisição desses itens.

A segunda hipótese consistia no fato de que bebês entre 10 e 12 meses começariam a mapear corretamente esses referentes, impulsionados por avanços na cognição social. Por fim, a terceira hipótese considerava que o desempenho na identificação de palavras abstratas melhoraria significativamente por volta dos 14 meses, em decorrência do amadurecimento geral das habilidades cognitivas e sociais nessa faixa etária.

O estudo contou com 98 participantes divididos em três grupos etários: 34 bebês de 6 a 9 meses; 46 bebês de 10 a 13 meses; 18 bebês de 14 a 16 meses. A partir de uma tarefa de Fixação Preferencial do Olhar, com rastreador ocular, os bebês assistiram a pares de vídeos lado a lado, organizados em 7 pares de palavras relacionadas, como *all-gone-hi* (acabou–oi); *eat-hug* (comer–abraçar); *dance-kiss* (dançar–beijar); *more-splash* (mais–respingo); *drink-smile* (beber–sorrir); *bye-uh-oh* (tchau–óh) e *sleeping-wet* (dormir–molhado). Cada

vídeo tinha 5 segundos de duração. Durante a exibição, os pais ouviam a palavra-alvo em fones de ouvido e a repetiam em sentenças como “Look! Eat, Eat!” (Olhe! Comer! Comer!). Os pares de palavras foram selecionados com base em sua frequência no vocabulário materno e em relatórios parentais sobre a compreensão infantil.

Não foram encontrados resultados significativos na correlação dos itens abstratos para os bebês de 6 a 9 meses. Por outro lado, para o grupo de bebês de 10 a 13 meses, foi encontrado valor de significância. O grupo de bebês de 14 a 16 meses foi, por sua vez, o que apontou maior capacidade de mapear os referentes corretos, a partir dos vídeos.

Interessantemente, ao contrastar esses resultados com o estudo anterior de Bergelson e Swingley (2012) sobre a aquisição de itens concretos, observa-se uma diferença no curso de aquisição entre itens lexicais concretos e abstratos. Enquanto bebês de 6 a 9 meses demonstraram maior habilidade em associar palavras concretas, como *banana* e *mão*, a seus respectivos referentes visuais, somente bebês de 10 a 13 meses foram capazes de mapear corretamente conceitos verbais abstratos, como *comer* ou *fazer carinho*.

Ainda que os autores atribuam o desenvolvimento progressivo da capacidade de bebês de 10 a 13 meses em compreender palavras abstratas ao amadurecimento da cognição social, que favorece a interpretação de conceitos mais complexos, tecemos duas críticas aos resultados. A primeira diz respeito às diferenças no desenho experimental entre os grupos etários. Isso porque, para o grupo de crianças de 6 a 9 meses, os estímulos apresentados eram em imagens e para o grupo mais velho, de 10–13 meses, as apresentações eram em vídeo. Essa alteração pode ter dificultado o desempenho dos bebês mais jovens, sugerindo que a dificuldade observada não reflete necessariamente uma incapacidade de reconhecer itens abstratos, mas sim uma limitação associada à tarefa aplicada.

A segunda concerne a diferença fonológica e morfológica dos estímulos lexicais, isso porque eles apresentavam uma variabilidade em relação ao número de sílabas, por exemplo os itens *all gone* (acabou) e *eat* (comer). Por essas razões, consideramos que os achados não se basearam em parâmetros robustos para comparar a capacidade das crianças de identificar itens lexicais de natureza conceitual distinta, visto que o tipo de tarefa experimental e as propriedades morfofonológicas desses itens eram bastante diferentes entre eles.

Um outro estudo relevante para o trabalho aqui desenvolvido investiga a capacidade preditiva e a identificação lexical por bebês de 12 a 24 meses. Nessa pesquisa, Reuter e Lew-Williams (2023) realizaram um estudo de rastreamento ocular que avaliou a habilidade de bebês na referida faixa etária de compreender nomes familiares e antecipar esses itens com base em palavras semanticamente relacionadas. Os bebês observaram duas imagens: um copo

e uma bola, enquanto ouviam frases correspondentes. Metade dos estímulos incluíam palavras e frases informativas que poderiam permitir que os bebês antecipassem um substantivo-alvo, por exemplo, “Let’s go play! Bring a toy! It’s fun to throw! Where’s the ball? Find the ball!” (Vamos brincar! Traga um brinquedo! É divertido jogar! Onde está a bola? Encontre a bola!), e a outra metade era composta por sentenças neutras, i.e., em que não havia nenhuma pista prévia para que o bebê pudesse antecipar a palavra-alvo, por exemplo “Look at that! There it is! Do you see it? Where’s the ball? Find the ball!” (Olha aquilo! Lá está! Você está vendo? Onde está a bola? Encontre a bola!).

Participaram do estudo, 67 bebês (sendo 37 meninos), falantes nativos do inglês, entre 11 e 25 meses de idade ($M = 18$ meses). As crianças foram divididas em quatro grupos etários, a saber: 12 meses ($n = 16$, $M = 12,6$ meses), 15 meses ($n = 16$, $M = 15,8$ meses), 18 meses ($n = 18$, $M = 18,7$ meses) e 24 meses ($n = 17$, $M = 24,6$ meses). O objetivo dos autores era avaliar se a capacidade de antecipação de itens lexicais conhecidos já estava presente no primeiro grupo etário ou se progrediria com o aumento da idade dos bebês.

Os estímulos auditivos do teste consistiam nas frases antecipatórias as quais tinham palavras informativas¹⁷, semanticamente relacionadas, e que os bebês podiam utilizar como pistas para antecipar um substantivo-alvo (por exemplo: eles ouviam uma sentença como “Let’s go eat. Oh, yum yum! Open your mouth! Where’s the cookie? Find the cookie!” (Vamos comer. Unh, nhammy, nhammy! Cadê o biscoito? Encontre o biscoito!) e frases neutras as quais não tinham nenhum componente em que o bebê pudesse antecipar o substantivo-alvo, por exemplo: “Look at that! There it is! Do you see it? Where’s the cookie? Find the cookie!” (Olhe para isso! Está ali! Você viu? Cadê o biscoito? Encontre o biscoito!”).

Os estímulos visuais eram imagens que representavam os quatro nomes-alvo: *bola*, *biscoito*, *copo* e *sapato*, com variação de duas imagens para cada um deles, por exemplo, uma bola azul e uma bola vermelha. Eles foram dispostos em pares, como *ball-cup* (bola-copo) e *cookie-shoe* (biscoito-sapato) e apareciam duas vezes durante o experimento (uma vez com uma frase neutra e outra com uma frase antecipatória). As condições foram distribuídas de forma intra-sujeitos (*within-subject*) e a posição do alvo foi aleatorizada, ora à direita ora à esquerda, garantindo que a posição do alvo e a condição (antecipatória/neutra) não se repetisse por mais de três *trials* consecutivos.

O modelo estatístico incluiu efeitos fixos para idade (contínua), condição (antecipatória/neutra) e tempo (intervalos de 100 ms, de 0 a 9000 ms a partir do início do

¹⁷Reuter e Lew-Williams (2023) reportam que a seleção das palavras experimentais foi feita a partir dos dados do *Wordbank*.

trial) e apontou efeitos significativos para idade, condição e tempo. Interessantemente, foram encontrados três efeitos principais: a fixação aumentou com o aumento da idade, os bebês fixaram o olhar por mais tempo nos *trials* com frases antecipatórias do que nas do tipo neutras e as fixações na imagem-alvo aumentaram à medida em que as frases eram apresentadas.

De forma geral, os resultados apontaram uma interação entre idade, condição e tempo. Isso não só significa que os bebês mais velhos foram melhores em inferir a condição correta, como também identificaram o alvo mais facilmente nas sentenças nas quais tinham pistas que pudessem antecipar a palavra-alvo, por exemplo, na frase “Let's go eat. Oh, yum yum! Open your mouth! Where's the cookie? Find the cookie!” (Vamos comer. Unh, nhammy, nhammy! Cadê o biscoito? Encontre o biscoito!)” em comparação às frases neutras “Look at that! There it is! Do you see it? Where's the cookie? Find the cookie!” (Olhe para isso! Está ali! Você vê? Cadê o biscoito? Encontre o biscoito!), em que o referente só era ouvido pela criança no final do trecho.

A fim de verificar se a habilidade de compreensão e antecipação se desenvolvem progressivamente com a idade, Reuter e Lew-Williams (2023) dividiram os bebês em quatro grupos etários (12, 15, 18 e 24 meses) e avaliaram a fixação com análises de permutação baseadas em *clusters*. Para o grupo de bebês de 12 meses não foram encontradas análises com *clusters* significativos. Em contrapartida, para o grupo de bebês de 15 meses foram observados dois *clusters* (na janela de 7100–8500 ms; e na de 8700–9000 ms). Também foi encontrado um *cluster* significativo para os bebês de 18 meses (no intervalo de 6700–9000 ms) e um outro para bebês de 24 meses (entre 6300–9000 ms).

Percebemos, a partir das análises, que o aumento gradual da capacidade de bebês de antecipar a imagem-alvo, a partir do estímulo sonoro, aumentou conforme a idade. A presença de *clusters* significativos apenas a partir dos 15 meses sugere que, nesse estágio, os bebês começam a demonstrar uma habilidade preditiva para identificar relações conceituais entre os estímulos auditivos e as imagens correspondentes.

Para ter resultados ainda mais robustos e indícios sobre essa habilidade preditiva das crianças, os autores optaram por calcular a proporção binarizada de fixações nos alvos durante os testes antecipatórios em intervalos de 100 ms (0–6000 ms a partir do início do teste). A hipótese deles era a de que, se os bebês utilizassem as palavras informativas (por exemplo, *comer*) para antecipar os nomes subsequentes (por exemplo, *biscoito*), a proporção de fixações no alvo seria maior que o nível da chance, após a apresentação da primeira palavra informativa e antes da apresentação do substantivo-alvo (1100–6000 ms a partir do início do

teste), indicando que os bebês mapeariam a imagem-alvo antecipadamente, antes que o substantivo fosse nomeado.

As análises das proporções binarizadas por grupos foram ao encontro das análises de permutação e também não indicaram *clusters* significativos para os bebês de 12 meses, mas identificaram dois *clusters* significativos para bebês de 15 meses (na janela de 2300–3100 ms; aos 4100–4500 ms; e aos 5000–5500 ms). Também foram identificados dois *clusters* significativos para bebês de 18 meses (aos 3200–4100 ms; e aos 4600–6000 ms) e um *cluster* significativo para bebês de 24 meses (aos 3700–6000 ms).

A partir desses resultados, é interessante observar que, aos 12 meses, as crianças não apresentaram diferenças claras entre as sentenças antecipatórias e neutras, com a proporção do olhar para o alvo permanecendo próxima ao nível da chance, durante quase todo o experimento. Esses dados sugerem que, nessa idade, ainda não há evidências consistentes da compreensão lexical ou da antecipação das palavras-alvo nas sentenças do tipo antecipatórias. Já no grupo de 15 meses, por sua vez, viu-se uma pequena diferença entre as sentenças do tipo antecipatórias e as neutras. No grupo de bebês de 18 meses, as crianças demonstraram uma distinção mais clara entre os dois tipos de sentenças e a proporção do olhar para o alvo nas sentenças antecipatórias aumentou, o que sugere a maior capacidade dessas crianças de antecipar o substantivo-alvo. Por fim, as crianças de 24 meses tiveram o melhor desempenho. Nesse grupo, durante as sentenças antecipatórias, a proporção do olhar para o alvo aumentou significativamente após o início da primeira palavra informativa, evidenciando uma maior habilidade de antecipação lexical.

Esses resultados são interessantes, uma vez que vão ao encontro dos resenhados no início deste capítulo a respeito da capacidade dos bebês de identificar padrões e regularidades morfossintáticas, por exemplo, ao ouvirem o verbo *comer* e o mapearem a um substantivo, semanticamente relacionado, como a imagem do *biscoito* e não a do *sapato*. No entanto, os bebês se mostram sensíveis a essas co-ocorrências só a partir dos 15 meses de vida, sugerindo que eles podem fazer uso dessas regularidades para identificar mais rapidamente referentes familiares.

Mais interessante ainda é que essa capacidade antecipatória foi especialmente evidente no grupo de bebês de 24 meses. Por isso, esses achados fornecem evidências robustas de que, aos dois anos de idade, ao contrário do observado em bebês de 12, 15 e 18 meses, as crianças já são capazes de compreender palavras informativas e utilizá-las para antecipar o substantivo-alvo, mesmo antes de escutarem o estímulo auditivo que referencia a imagem.

Este capítulo objetivou apresentar algumas pesquisas encontradas que exploraram a aquisição lexical de itens concretos e abstratos por bebês. Destacamos, ao longo do capítulo que, quanto maior era a acurácia de resposta em relação à identificação dos itens lexicais de natureza mais abstrata, como verbos ou como a capacidade preditiva de alguns itens, maior era a idade das crianças (Bergelson; Swingley, 2012, 2013; Reuter; Lew-Williams, 2023).

Em particular, os resultados apresentados por Reuter e Lew-Williams (2023) são relevantes para a presente pesquisa por trazerem a comparação da habilidade preditiva e antecipatória entre bebês de grupos etários diferentes e apresentar evidências significativas de que bebês, aos 24 meses de vida, em um experimento com rastreador ocular, já conseguem abstrair pistas semanticamente relacionadas e mapear o item lexical correto. Ainda que todos os estudos resenhados contribuam empiricamente para explicar a aquisição lexical de itens concretos e abstratos e adotem a mesma metodologia desta tese, os resultados de Reuter e Lew-Williams (2023) justificam a faixa etária selecionada para o teste aqui desenvolvido.

Embora alguns trabalhos tenham se dedicado a responder quais itens funcionais são adquiridos precocemente por bebês, e quais mecanismos cognitivos podem subjazer essa capacidade de apreender itens tão abstratos na sintaxe, a questão sobre a aquisição de itens lexicais mais abstratos ainda carece de evidências na literatura psicolinguística.

A respeito do processamento de itens mais abstratos, a neurociência da linguagem já forneceu evidências sobre como conceitos linguísticos são processados e como estão distribuídos em regiões diferentes no cérebro, o que é interessante ao se pensar que a representação conceitual, como no caso dos itens lexicais abstratos, pode ser acessada/ativada por meio de outras cognições sensoriais, como a visual, a tátil, a auditiva (Amedi *et al.*, 2004; Bedny, Saxe, 2012).

O próximo capítulo complementarará a discussão sobre achados experimentais com dados de uma revisão sistemática e apresentará estudos sobre o acesso lexical de conceitos e a natureza semântica mais abstrata que depende de ativações de redes semânticas distribuídas no córtex cerebral, por exemplo, algumas áreas sensório-motoras.

4 REVISÃO SISTEMÁTICA DA LITERATURA SOBRE A AQUISIÇÃO DE ITENS LEXICAIS CONCRETOS E ABSTRATOS

A aquisição de linguagem é uma área de pesquisa heterogênea que não só traz evidências experimentais a respeito do desenvolvimento linguístico de uma população específica, como também pode oferecer dados para outras áreas, como a neurologia, a psicologia, a fonoaudiologia, a ciência de dados, dentre outras. Embora as pesquisas em AL tenham um caráter transdisciplinar, a temática referente à aquisição de itens concretos e abstratos por bebês parece ser ainda incipiente na literatura linguística ou de outras áreas das ciências biológicas, sobretudo quando a população de interesse se restringe a bebês brasileiros.

Diante da escassez de dados sobre a aquisição de nomes concretos e abstratos, especialmente por bebês brasileiros, foi realizada uma revisão sistemática da literatura. O objetivo não foi apenas levantar rigorosamente estudos já publicados para compor a bibliografia desta tese, mas também contribuir para a literatura linguística e suas áreas de interface, aprofundando a compreensão dos processos neurocognitivos que subjazem essa aquisição.

Para a realização da revisão sistemática da literatura sobre a aquisição de itens lexicais concretos e abstratos, utilizamos o Google Acadêmico como mecanismo de busca. Essa ferramenta atua como um indexador/metabuscador que organiza e disponibiliza produções científicas e acadêmicas, indicando os *links* nos quais os arquivos estão hospedados. Diferentemente de bases de dados como a *Scopus*, *PubMed* ou *Web of Science*, o Google Acadêmico não adota critérios rigorosos de seleção e padronização dos conteúdos indexados.

Portanto, para a elaboração da revisão sistemática, as palavras-chave inseridas para realizar a primeira etapa, a qual consiste na seleção de todos os artigos da base, foram: *semantic features*; *concrete words*; *abstract words*; *language acquisition*, inserindo o período de seleção de 1990 até o ano de 2023. Os trabalhos foram organizados de acordo com a temática investigada nesta tese e, também, em relação aos seus aspectos metodológicos. Foram excluídos estudos que fossem de aquisição de línguas de sinais, de línguas adicionais (L2 ou mais) ou, ainda, de línguas artificiais. Também foram descartados aqueles que não tivessem metodologia experimental ou tratassem especificamente de valência emocional no processamento lexical.

Como resultado, obtivemos uma base de 376 artigos, a partir da inserção das quatro palavras-chave, transcritas acima. Desse total, foram excluídos 99 artigos que tinham

“bilinguismo” no título; 18 que continham “L2”; 8 com a palavra “distúrbios” no título e 9 com a palavra “afasia”. Outros critérios para a exclusão foram 4 artigos repetidos na base, 1 que não tinha título e nem *link* e 13 que não tinham sido encontrados para *download* naquela etapa. Portanto, 152 artigos foram descartados nessa fase.

Embora alguns estudos parecessem contemplar a temática e tenham aparecido como resultado na busca, após a leitura criteriosa do resumo de cada um dos artigos remanescentes da base de dados gerada, ainda foram excluídos 162 trabalhos, cujas justificativas para a eliminação desses da revisão foram detalhadas na própria base da revisão sistemática, nas respectivas colunas *Exclusion criteria - Full text* ou *General Comments/Highlights*¹⁸. Nesta terceira etapa da elaboração da revisão sistemática, restaram, então, 62 artigos para leitura na íntegra e para avaliação. Desse número, foram descartados os artigos abaixo pelas razões que serão descritas a seguir.

Alguns apresentam uma perspectiva teórica, como em *Semantic Representation* de Vigliocco e Vinson (2007), em que os referidos autores revisam as perspectivas neuropsicológicas sobre a representação do significado, discutindo como o significado lexical é representado na mente/cérebro dos falantes. De maneira semelhante, Vega *et al.* (2008), no capítulo, *Reflecting on the debate* argumenta a favor de um quadro simbólico e traz evidências experimentais, especialmente de efeito de compatibilidade de sentenças de ação (ACE), que vão ao encontro de abordagens corporificadas. Ademais, os referidos autores discorrem como as estruturas simbólicas e incorporadas podem explicar o processamento referencial e o significado abstrato. Não apenas, são apresentados dados acerca de como as áreas perceptivas e motoras do cérebro são recrutadas no processamento do significado, bem como o papel das áreas tradicionalmente relacionadas à linguagem e de outras regiões de processamento de ordem superior do cérebro. Por fim, ainda discorrem sobre como a inteligência artificial e a robótica podem contribuir para esclarecer algumas questões do processamento linguístico e semântico.

Outros estudos desse tipo consideram a abordagem de fenômenos neurocognitivos, como a representação conceitual, em abordagens de semântica corporificada (Kiefer; Pulvermüller, 2012, 2013; Fernandino *et al.*, 2015; Reilly *et al.*, 2016), em abordagens amodais (Reilly *et al.*, 2016) ou em modelos teóricos mistos, como a Groundability (Thill; Twomey, 2016). Em “*Semantic embodiment, disembodiment or misembodiment? In search of*

¹⁸A lista completa dos artigos excluídos em razão do título pode ser consultada no excel: Systematic review - -
Olivia Fernandes Bogo.xlsx. Disponível em:
<https://docs.google.com/spreadsheets/d/1k-X6oQRhwyTItpDsLu3kcWYs0kMIYe3n/edit#gid=1227332640>.

meaning in modules and neuron circuits”, Pulvermüller (2013) propõe uma abordagem integrativa às teorias corporificadas e às não-corporificadas. Trata-se da Teoria da Percepção de Ação (do inglês APT) a qual postula que tanto as áreas sensório-motoras quanto as multimodais do cérebro contribuem diferencialmente para tipos específicos de significados e conceitos.

Em “*The case of CAUSE: neurobiological mechanisms for grounding an abstract concept*”, Pulvermüller (2018) assume a causalidade como um exemplo concreto e postula que conceitos abstratos podem ser aprendidos e fundamentados na interação da vida real. O referido autor propõe mecanismos neurocognitivos para explicar alguns casos típicos de causalidade. Em *Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition*, Sakreida *et al.* (2013) desenvolveram um experimento com fMRI, porém com universitários alemães, para verificar os correlatos neurais de palavras concretas e abstratas em contextos de ação. Os achados dos referidos autores sugerem que o processamento de palavras concretas recruta preferencialmente o sistema sensório-motor, enquanto o processamento de palavras abstratas recruta preferencialmente o sistema linguístico.

Ainda nesse panorama teórico, foram encontrados os trabalhos de Haeri (2017) sobre a representação conceitual no sistema cognitivo humano e as diferenças entre conceitos concretos e abstratos e o trabalho de Borghesani e Piazza (2017), *The neuro-cognitive representations of symbols: the case of concrete words*, sobre as representações neurocognitivas dos símbolos: o caso das palavras concretas.

Em *How does meaning come to mind? Four broad principles of semantic processing*, Pexman (2020) apresenta quatro princípios acerca do processamento semântico à luz da semântica lexical. Para tanto, discorre sobre a questão da arbitrariedade e da iconicidade do signo linguístico, sobre a riqueza semântica na representação conceitual, sobre as propriedades de conceitos abstratos, os quais se constituem como um desafio para as teorias semânticas, e como eles são adquiridos e representados. Por fim, discorre, também, a respeito da influência da experiência para o processamento semântico.

No capítulo *Meaning in acquisition: Semantic structure, lexical organization, and crosslinguistic variation*, Li (2010) discute como a criança descobre o significado das palavras em um sistema lexical e quais os mecanismos que propiciam essa aquisição. Para discorrer sobre isso, o referido autor se atém a três questões: a aquisição do tempo verbal e do aspecto, a aquisição de *cryptotipos* e o desenvolvimento da estrutura lexical.

Já Borghi *et al.* (2018) em *Varieties of abstract concepts: development, use and representation in the brain*, realiza um relato teórico acerca dos conceitos abstratos e de como classificá-los. Ademais, discorre sobre como outros processos, como a sociabilidade, além de processos neurobiológicos são importantes para o desenvolvimento de conceitos abstratos. Não só, a referida autora apresenta ferramentas metodológicas para investigar esse tipo de palavra.

Similarmente, Pexman *et al.* (2022) em *Social Semantics: The Organisation and Grounding of Abstract Concepts* sugerem que o significado social é uma característica fundamental na representação de alguns conceitos. Além disso, os autores fornecem uma revisão bibliográfica a favor dessa proposta para os estudos psicolinguísticos e neurolinguísticos e apresentam teorias contemporâneas sobre a representação semântica e a cognição social tanto no viés neurocognitivo quanto no comportamental.

Ainda em relação à categorização de itens lexicais, Bruhl (2014), em *The role of situations and thematic reorganization in the conceptual processing of abstract concepts*, avaliou os princípios que fundamentam a estrutura conceitual, a organização e a representação de conceitos abstratos dentro do sistema cognitivo. Para isso, realizou experimentos comportamentais, utilizando tarefas de categorização e de julgamento de similaridade a fim de verificar o papel da informação situacional e da organização temática no processamento de conceitos abstratos.

Também, acerca da definição de itens lexicais, Gandía (2016), em sua tese *On Word Definition in Children and Adults: Effects of Word Category and Level of Abstraction*, desenvolveu uma sequência de experimentos para testar a definição morfológica lexical e seus níveis de abstração. Nesse viés, no livro *Longitudinal Developments in Vocabulary Knowledge and Lexical Organization*, Dóczy e Kormos (2016) tratam do desenvolvimento longitudinal do conhecimento de vocabulário e da organização lexical em ambientes de ensino de Inglês como língua estrangeira (do inglês EFL). Ademais, os autores correlacionam os estudos de desenvolvimento vocabular às teorias de aprendizagem de uma L2.

Huang e Federmeier (2015), por sua vez, em *Imaginative Language: What Event-Related Potentials have Revealed about the Nature and Source of Concreteness Effects* também revisam trabalhos os quais utilizaram ERPs - Potenciais Relacionados a Eventos, alguns associados à apresentação visual de meio campo, para explicar a natureza e a localidade dos efeitos de concretude. De acordo com os referidos autores, os resultados apontam para diversos mecanismos subjacentes às diferenças de processamento

comportamental para palavras concretas e abstratas. Além disso, indicam um papel predominante do hemisfério direito na associação de palavras com imagens sensoriais.

Outros estudos excluídos tratam da questão de traços semânticos (Muraki; Sidhu; Pexman, 2020) e seus valores em itens lexicais abstratos, a partir de tarefas de decisão lexical, com adultos (Recchia; Jones, 2012). Já Siew (2020), em *Feature distinctiveness effects in language acquisition and lexical processing: Insights from megastudies*, analisou as propriedades estruturais das palavras, inseridas em um espaço de características semânticas, na aquisição e no processamento lexical, a partir do levantamento de megaestudos e de base de dados, como o *English Lexicon Project* e o *Calgary Semantic Decision Project*.

Rabovsky (2013), em sua dissertação *Semantic Richness Effects in Visual Word Processing: Combining Event-Related Brain Potentials and Connectionist Network Modeling*, verificou a influência da riqueza semântica no processamento lexical e semântico, a partir da execução de experimentos com ERPs para verificar a influência da riqueza semântica no processamento lexical e semântico, a partir da simulação de modelo de rede de cognição semântica.

Foram mapeados, também, muitos artigos sobre valência emocional e tipo de palavras, como *The Emotional Embodiment of Words in a second language: An eye-movement Study*, de Sheikh e Titone (2015); *Investigating Emotion-label and Emotion-laden Words in a Semantic Satiation Paradigm*, de Hildebrandt (2020); *Interaction of Lexical-Semantic and Imagery Representations*, de Zinovieff (2000); *Age-Dependent Positivity-Bias in Children's Processing of Emotion Terms*, de Bahn et al. (2017); e *Affective neurolinguistics: towards a framework for reconciling language and emotion*, de Hinojosa, Moreno e Ferré (2019). Ferré et al. (2015), em *The role of emotionality in the acquisition of new concrete and abstract words*, investigou se a valência emocional e a concretude são efeitos facilitadores para as etapas iniciais da aquisição vocabular, porém o experimento foi feito com adultos, adquirindo uma L2.

Sobre a relação de áreas do cérebro relacionadas ao processamento semântico, Dreyer e Pulvermüller (2018), no trabalho *Abstract semantics in the motor system? – An event-related fMRI study on passive reading of semantic word categories carrying abstract emotional and mental meaning* conduziram um estudo investigando a possível correlação da ativação de áreas motoras e da semântica abstrata.

Nesse mesmo viés, Borghi et al. (2018), em *Abstract concepts, language and sociality: from acquisition to inner speech* também sugerem que tanto a experiência social quanto a linguística desempenham um papel na aquisição e no processamento de conceitos

abstratos. De forma similar, Bowers (2008) em sua tese *Hemispheric difference in semantic priming for early and late age of acquisition words* investiga as diferenças hemisféricas cerebrais em tarefas de *priming* semântico na aquisição precoce e tardia de palavras.

Alguns autores, apesar de utilizarem metodologia experimental, dedicaram-se a investigar a diferença no processamento de nomes e verbos, como Shapiro e Caramazza (2003, 2006, respectivamente) em *The representation of grammatical categories in the brain* e *Cortical signatures of noun and verb production*. Semelhantemente, Tsai *et al.* (2009), utilizando ERPs, testou o processamento dessas duas classes de palavras, por falantes adultos, falantes do chinês a fim de verificar questões referentes à imageabilidade.

Em *Noun and verb in the mind: An interdisciplinary approach*, Błaszczak e Klimek-Jankowska (2015) discorrem sobre os conceitos abstratos e acerca de alguns modelos teóricos que visam a compreender como esses são adquiridos, seja pela defesa de uma representação semântica fortemente corporificada em relação aos conceitos sensório-motores ou pela visão de que o significado é *fundamentado* através de sistemas cognitivos, incluindo aqueles envolvidos no processamento da linguagem e da emoção. Contudo, Pexman *et al.* (2015) argumentam em defesa de que o significado social é essencial na representação de alguns conceitos. Para isso, fornecem dados psicolinguísticos e de neuroimagem, discutem a definição de “sociabilidade” e realizam um panorama sobre teorias da representação semântica e da cognição social.

Na tese *Neural correlates of word learning and meaning acquisition*, Mestres-Missé (2007) apresenta uma série de teorias sobre a aquisição de palavras e sobre a representação conceitual, além de realizar sete experimentos para investigar os correlatos neurais envolvidos nesse processo. No entanto, tanto os experimentos com eletroencefalografia, quanto os com ressonância magnética funcional foram realizados com adultos, em paradigmas de aprendizagem de palavras e/ou significados.

Já no trabalho *Functional Neuroanatomy of Contextual Acquisition of Concrete and Abstract Words*, Mestres-Missé *et al.* (2008) simularam a aprendizagem de itens lexicais concretos e abstratos, com adultos, em um experimento que fornecia novas palavras relacionadas a contextos concretos e abstratos, enquanto a atividade cerebral dos participantes era mapeada por ressonância magnética funcional. Segundo os autores, a aprendizagem de palavras concretas novas ativou, mais significativamente, o Giro Fusiforme Anterior Ventral, uma área relacionada à imageabilidade, a qual já foi relacionada ao processamento de palavras concretas.

Similarmente, em *Mapping concrete and abstract meanings to new words using verbal contexts*, Mestres-Missé *et al.* (2014) investigaram os efeitos do contexto semântico e da concretude das palavras no mapeamento de significados para novas palavras a fim de discutir diversas abordagens que têm sido utilizadas para tentar explicar *o efeito de concretude*. Em uma série de três experimentos, a partir de um paradigma de aprendizagem de palavras novas, os participantes deveriam descobrir o significado de uma palavra, a partir de um contexto verbal específico. De acordo com os referidos autores, o efeito de concretude, sobretudo para as palavras concretas, observado na aprendizagem, é resultante da diferente organização da informação conceitual abstrata e concreta na memória semântica.

Além da questão da perspectiva da semântica corporificada, Fernandino *et al.* (2015) em *Concept Representation Reflects Multimodal Abstraction: A Framework for Embodied Semantics* apontam como áreas sensoriais e motoras desempenham um papel significativo na representação conceitual. Esses autores também indicam haver um sistema hierárquico de representações sensório-motoras abstratas que incorporam uma divisão importante entre interação e percepção de objetos. No estudo complementar, *Predicting brain activation patterns associated with individual lexical concepts based on five sensory-motor attributes*, Fernandino *et al.* (2015) pesquisam se um modelo de codificação baseado em cinco atributos semânticos — som, cor, movimento visual, forma e manipulação, diretamente relacionados à experiência sensório-motora, poderia eliciar a ativação cerebral de conceitos lexicais individuais concretos ou abstratos.

Antonucci e Alt (2011), em *A lifespan perspective on semantic processing of concrete concepts: does a sensory/motor model have the potential to bridge the gap?*, exploram, também, a possibilidade de que o conhecimento semântico sobre conceitos concretos pode ser construído com base em processos sensoriais/motores e em evidências comportamentais, neuropsicológicas e de neuroimagem.

Houve um trabalho cujo experimento era de natureza comportamental, trata-se de Zdrzilova (2017), na tese *Understanding the Meanings of Abstract Concepts: Keeping it Grounded and Real*, indo ao encontro, na análise dos resultados, de abordagens mais corporificadas no que tange ao processamento de conceitos abstratos. A fim de investigar os efeitos da imageabilidade na produção linguística, Ribu (2012), em *An Image Is Worth a Thousand Sounds? On Imageability and Phonological Neighborhood Density Effects in Speech Processing* aplicou seus experimentos e contrastou seus dados com população neurotípica e neurodivergente.

Foram também encontrados trabalhos sobre modelagem neural, como o de Radford (1997), o de Howell (2004), mencionado anteriormente, o de Henningsen-Schomers e Pulvermüller (2022) e o de Mikeleiz, Konstenius e Ezquerro (2011). Além desses que fornecem modelos neurais com abordagens teóricas diversas, Howell *et al.* (2005), no estudo *A model of grounded language acquisition: Sensorimotor features improve lexical and grammatical learning* argumenta que tanto conceitos pré-linguísticos quanto os da semântica corporificada dão base para a aquisição linguística. No entanto, o resultado do experimento de simulação, com adultos, apontou que modelos de redes neurais que considerem aspectos sensório-motores têm desempenho melhor no processamento natural de palavras, em relação à predição de palavras.

Nesse viés de articular representações sensório-motoras e modelos de aprendizagem neural, foi encontrada também a tese *Sensorimotor Representations of Meaning in Early Language Acquisition* de Howell (2004), com uma série de experimentos realizados com adultos. Já Rabovsky e McRae (2014), em *Simulating the N400¹⁹ ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning*, discute que ainda não são claros os mecanismos computacionais do N400. Por isso, investigaram, a partir de um modelo conexionista, a covariância na amplitude do N400, a partir de erros da rede semântica.

Mais recentemente, em *Modelling concrete and abstract concepts using brain-constrained deep neural networks*, Henningsen-Schomers e Pulvermüller (2022) desenvolveram uma rede neural que reproduz a função da área cortical responsável pelo processamento sensório-motor, conceitual e linguístico com o intuito de investigar os possíveis mecanismos biológicos subjacentes à formação de categorias conceituais e semânticas. De acordo com esses autores, as redes foram programadas para aprender padrões neurais que representam objetos específicos e ações para “fundamentar” semanticamente conceitos concretos e abstratos.

Sobre potenciais relacionados a eventos linguísticos, *Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential* de Kutas e Federmeier (2011) foi excluído por se tratar de uma revisão da literatura acerca do papel do N400 em relação aos ERPs linguísticos. Já o capítulo *Event-Related Potential (ERP) Research in Neurolinguistics: Part I* de Segalowitz e Chevalier (1998) aponta as vantagens e

¹⁹Entende-se por N400 no ERP: “um pico negativo de aproximadamente aos 400ms (N400), após o instante de estimulação – [o N400] passou a ser robustamente relacionado a incongruências semânticas de sentenças” (França, 2002, p. 130).

as desvantagens no uso de Potenciais Relacionados a Eventos para as pesquisas em neurociência da linguagem. Também foram revisados, no referido trabalho, paradigmas específicos de ERPs abordando questões de acesso lexical, de distinções de classes de palavras (classe aberta vs. classe fechada e substantivo vs. verbo) e de acesso fonológico.

Van Petten e Luka (2006), em *Neural localization of semantic context effects in electromagnetic and hemodynamic studies*, realizam um panorama sobre a utilização de técnicas eletromagnéticas e hemodinâmicas para os estudos linguísticos, indicando como o um ERP, aos 400ms (N400), é um indicador de facilidade ou de dificuldade de recuperação do conhecimento conceitual de uma palavra. Os autores supracitados discorrem sobre como evidências de pacientes com danos cerebrais, a partir de dados de magnetoencefalografia, apontam para uma extensa área do lobo temporal esquerdo como a maior região do efeito de contexto semântico, eliciando um N400.

Os estudos de ressonância magnética funcional relacionada a eventos, nos quais houve manipulação de contexto semântico, apontaram maior atividade hemodinâmica no Giro Temporal Superior Esquerdo quando os estímulos tinham incongruência semântica. Sendo esses achados compatíveis com os resultados eletrofisiológicos. Em menor grau, outra área comumente identificada nos estudos de fMRI foi o giro frontal inferior esquerdo, que não parece influenciar de modo significativo no efeito do N400.

Embora a base de dados desta revisão sistemática contasse, inicialmente, com 376 artigos para análise, é pertinente ressaltar que a busca pelas palavras-chave selecionadas não identificou trabalhos sobre aquisição de linguagem, e menos ainda trabalhos referentes à aquisição em PB.

Os poucos artigos (Bergelson; Swingley, 2012; Bergelson; Swingley, 2013; Gleitman et al, 2005; Gleitman; Trueswell, 2020; Reuter; Lew-Williams, 2023) que tratavam sobre a aquisição de linguagem, mais especificamente de itens lexicais concretos ou abstratos, com a faixa etária compatível com a do experimento realizado neste estudo e com paradigmas experimentais da psicolinguística foram encontrados pelo Google Acadêmico por busca da própria autora da tese. No quadro 1, abaixo, estão sintetizados os 9 artigos selecionados por meio da revisão sistemática e os que tratam, especificamente, de AL.

Quadro 1 — Quadro síntese dos estudos selecionados na revisão sistemática

Título	Autores e ano	Língua do Estudo	Participantes	Técnica Experimental	Palavras-Chave	Resultado
Are pumpkins better than heaven? An ERP investigation of order effects in the concrete-word advantage	Tolentino, Tokowicz (2009)	Inglês Americano	Universitários	EEG	Concreteness; ERPs; Stimulus order; N400; Semantic context; Feature activation;	As respostas do ERP aos tipos de palavras foram mais positivas do que as respostas a pseudopalavras, indicando que os participantes eram sensíveis à lexicalidade, aos 400Ms. As respostas às palavras concretas foram mais negativas que as respostas às palavras abstratas quando um bloco de palavras abstratas precedeu um bloco de palavras concretas e quando os dois tipos de palavras foram misturados em um único bloco; quando palavras concretas precederam palavras abstratas, não houve diferença nos ERPs.

Título	Autores e ano	Língua do Estudo	Participantes	Técnica Experimental	Palavras-Chave	Resultado
Modulation of the semantic system by word imageability	Sabsevitz <i>et al.</i> (2005)	Inglês Americano	Adultos (Idade média 22.8)	fMRI	Semantic system; Word imageability; Sensory-motor representation; fMRI; Abstract noun	Os resultados apontam os efeitos da imageabilidade das palavras na ativação neural, demonstrando que um sistema de memória semântica bilateral responde mais fortemente a palavras concretas do que abstratas. Os achados indicam uma forte modulação do córtex temporal ventral pela imageabilidade da palavra, conforme previsto pela teoria sensório-motora do processamento conceitual. Os dados indicam uma dissociação entre esta rede organizada bilateralmente, sensível a conceitos com atributos sensório-motores e um sistema lateralizado à esquerda, especializado no processamento de palavras abstratas.

Título	Autores e ano	Língua do Estudo	Participantes	Técnica Experimental	Palavras-Chave	Resultado
Contextual Acquisition of Concrete and Abstract Words: Behavioural and Electrophysiological Evidence	Mkrtychian <i>et al.</i> (2021)	Russo	Adultos (Idade média: 23.4 anos)	EEG	Abstract semantics; Concrete semantics; Word acquisition; Language learning; Event-related potentials	A aquisição de conceitos abstratos e concretos novos é compartilhada por mecanismos neurais, que exibem padrões diferenciais, desde o início da dinâmica do ERP até ~210 ms, indicando diferentes velocidades de ativação, distribuição e lateralização hemisférica em circuitos de memória. A distribuição da fonte de ativação para palavras concretas e abstratas apresentadas visualmente indica o papel de ambos os hemisférios cerebrais na aquisição de palavras em geral, e dos circuitos temporo-parietais nas proximidades da área de Wernicke e seu homólogo no hemisfério direito. As diferenças entre o processamento de novos conceitos concretos e abstratos são evidentes no nível comportamental. Os participantes tiveram um melhor desempenho nas tarefas que avaliaram a aquisição dos significados das palavras concretas.

Título	Autores e ano	Língua do Estudo	Participantes	Técnica Experimental	Palavras-Chave	Resultado
Investigating early semantic priming: Event-related potentials as a window into the organization of word meaning in children at age 24 months	Syversen (2007)	Norueguês	Bebês - 24m	EEG	Não há	Os resultados apontam que as palavras-alvo antecedidas por uma palavra semanticamente não relacionada provocaram um efeito semelhante ao N400, em comparação às palavras-alvo antecedidas por uma palavra semanticamente relacionada. O efeito de priming semântico, refletido pelo N400, pode ser uma ferramenta funcional para investigar o significado lexical, permitindo também examinar a memória semântica em um estágio muito inicial da aquisição da linguagem.
Hard Words - Language Learning and Development	Gleitman et al, (2020)	Inglês	Adultos universitários	<i>Human Simulation Paradigm</i>	Word learning; argument Structure; syntactic Bootstrapping; Conceptual development; Language acquisition	Os resultados indicam que substantivos concretos são muito mais facilmente identificados do que verbos. Verbos abstratos (ex.: think, know) raramente são inferidos apenas pelo contexto visual, mas são melhor compreendidos quando pistas sintáticas e distribucionais são combinadas com o contexto. Por isso, os autores argumentam, a partir de evidências de um paradigma de aprendizagem de palavras, a favor do Bootstrapping sintático. Para eles, as crianças combinam informações do mundo (palavra-referente) e pistas

Título	Autores e ano	Língua do Estudo	Participantes	Técnica Experimental	Palavras-Chave	Resultado
						linguísticas (como a estrutura sintática) para aprender o significado das palavras, inclusive das mais abstratas.
Easy words: reference resolution in a malevolent referent world	Gleitman et al., (2005)	Inglês	Revisão teórica e integrativa de outros estudos com participantes adultos, bebês de 14–18 meses e crianças de 2, 3 e 4 anos	<i>Human Simulation Paradigm</i> ; observação de interações pais-filhos em vídeo; análises experimentais e teóricas sobre aquisição lexical	Psycholinguistics; Language development; Word learning; Reference	Os resultados apontam que a aprendizagem do vocabulário inicial ocorre em duas etapas: primeiro, por associação direta entre palavra e mundo (palavras fáceis); depois, pelo uso de pistas sintáticas (syntactic bootstrapping) para aprender palavras difíceis. Esse processo depende de momentos em que há referência (gems) em meio a muitos contextos ambíguos (junk). O processo de aquisição é, então, caracterizado pela atenção conjunta e pela precisão temporal entre a fala e a presença do referente. Os autores propõem o modelo propose-but-verify, no qual a criança gera e testa hipóteses de significado, descartando as incorretas.

Título	Autores e ano	Língua do Estudo	Participantes	Técnica Experimental	Palavras-Chave	Resultado
At 6-9 months, human infants know the meanings of many common nouns.	Bergelson; Swingley (2012)	Inglês	33 bebês entre 6 e 9 meses de idade	Paradigma do Mundo Visual com rastreamento ocular.	Word learning; Cognitive development; Infant cognition	Entre 6 e 9 meses, os bebês mapearam o objeto nomeado, indicando que compreendiam o significado de diversos substantivos comuns, como partes do corpo e objetos distais. Os resultados sugerem que a compreensão lexical começa muito antes da produção de palavras e não se limita aos nomes familiares mais tarde no primeiro ano de vida.
The acquisition of abstract words by young infants	Bergelson & Swingley (2013)	Inglês	98 bebês de 6 a 16 meses (divididos em 3 grupos: 6-9 meses, 10-13 meses, 14-16 meses)	Paradigma do Mundo Visual com rastreamento ocular.	Language acquisition; Word learning; Cognitive development; Infancy; Psycholinguistics	Bebês de 6-9 meses não reconheceram palavras abstratas. Entre 10-13 meses, os bebês começaram a identificar corretamente palavras abstratas. Aos 14-16 meses, os bebês apresentaram uma melhora robusta na compreensão, semelhante ao desempenho com palavras concretas.

Título	Autores e ano	Língua do Estudo	Participantes	Técnica Experimental	Palavras-Chave	Resultado
Infants' lexical comprehension and lexical anticipation abilities are closely linked in early language development	Erel et al. (2023)	Inglês	67 bebês entre 12 a 24 meses (divididos em 4 grupos: 12 meses (16 bebês); 15 meses (16 bebês); 18 (18 meses bebês); 24 meses (17 bebês)	Paradigma do Mundo Visual com rastreamento ocular.	Não há	Aos 12 meses de idade, os bebês ainda não identificavam de forma consistente a palavra-alvo, e não apresentavam diferença significativa entre as condições antecipatória e neutra. Na condição antecipatória, a frase permitia prever o alvo antes de ele aparecer (ex.: “The boy eats the...”, enquanto a imagem da comida já estava na tela). Na condição neutra, a frase não oferecia pistas antecipatórias (ex.: “Look at the...”). Aos 15 meses, os bebês começaram a apresentar uma tendência de antecipação, mas sem significância. Aos 18 meses, os bebês já demonstravam uma antecipação lexical robusta e confiável, olhando para o alvo antes mesmo de ele aparecer. Só aos 24 meses, os bebês anteciparam o referente com diferença estatisticamente significativa. Os resultados apontam maior capacidade de identificação do alvo à medida que a idade dos bebês aumenta.

Embora este levantamento bibliográfico conte com 376 artigos na sua base, há abordagens teóricas heterogêneas na investigação de fenômenos correlatos que perpassam quadros teóricos mais formalistas, tratando a representação conceitual em traços semânticos; experimentos neurocognitivos, fornecendo evidências neurobiológicas dos processos cognitivos; e modelos neurocomputacionais, reproduzindo os mecanismos neurobiológicos relativos ao processamento motor, cognitivo e conceitual.

No entanto, sobre o processamento ou a compreensão de itens lexicais, especialmente nomes concretos e abstratos, poucos artigos exploraram essas questões. A perspectiva é ainda mais escassa quando se trata de estudos sobre a aquisição linguística, haja visto que apenas um artigo foi mapeado e nenhum com crianças brasileiras.

Durante a defesa de qualificação desta pesquisa, uma avaliadora trouxe uma questão relevante a respeito da produção científica brasileira e, mais especificamente, da indexação dos estudos disponíveis no Google Acadêmico. A busca realizada com descritores em inglês (*semantic features; concrete words; abstract words; language acquisition*) não retornou estudos sobre a aquisição da linguagem em português brasileiro (PB).

Essa ausência, contudo, não indica falta de pesquisas na área, uma vez que diversos laboratórios brasileiros investigam e publicam, de forma sistemática, sobre a aquisição de linguagem por bebês e crianças brasileiras. Entre eles, destacamos o Laboratório de Acesso Sintático (ACESIN Lab — UFRJ), ao qual a autora desta tese está vinculada, o Laboratório de Línguas de Sinais e Oraís em Psicolinguística e Aquisição (SOPA Lab — UFRJ), o Laboratório de Psicolinguística e Aquisição da Linguagem (LAPAL — PUC-Rio) e o Núcleo de Estudos em Aquisição da Linguagem e Psicolinguística (NEALP — UFJF), além de outros grupos de pesquisa distribuídos pelo país.

O fato dessas publicações não terem sido recuperadas a partir dos descritores em inglês sugere uma limitação de indexação e, conseqüentemente, de visibilidade internacional, o que pode ter gerado uma subnotificação de publicações sobre aquisição de linguagem com crianças brasileiras. Esse cenário revela que a ausência de estudos brasileiros nos resultados de busca não decorre de uma inexistência de pesquisas, mas sim de um possível viés na forma como o Google Acadêmico indexa e prioriza publicações, sobretudo as internacionais.

Embora seja uma ferramenta amplamente utilizada, o mecanismo de busca parece indexar artigos publicados em inglês e ou oriundos de países norte-centrais, reforçando um recorte da literatura majoritariamente proveniente de contextos WEIRD (Singh et al., 2023). Isso é problemático, como mencionado na introdução desta pesquisa, porque limita a

compreensão de fenômenos cognitivos de uma forma mais heterogênea, a partir de dados de uma amostra populacional diversa.

Portanto, a presente tese traz, diante desse cenário, evidências empíricas inéditas à literatura em neurociência da linguagem, ancoradas em uma discussão teórica extensa sobre a conceitualização e a representação de domínios cognitivos, em particular de conceitos linguísticos de natureza concreta e abstrata, nas etapas iniciais da aquisição linguística, por crianças brasileiras. É oportuno, também, relatar que todo o planejamento e a execução do experimento se constituiu como um desafio, visto que não foram encontrados parâmetros nem resultados prévios na literatura que pudessem servir de base para a elaboração do *design* experimental.

Percebemos, por fim, como os mecanismos neurocognitivos ainda precisam ser explorados, a fim de se compreender de maneira mais aprofundada o seu funcionamento, responsável por consolidar as funções cognitivas e possibilitar a integração entre linguagem, percepção, memória e atenção. Além disso, é necessário investigar como o conhecimento é adquirido, armazenado e processado nas etapas iniciais da aquisição linguística.

Essa questão será aprofundada nos próximos capítulos que abordam a metodologia e as partes experimentais desta tese.

5 A AQUISIÇÃO DE NOMES CONCRETOS E ABSTRATOS POR BEBÊS BRASILEIROS: OS PRÉ-TESTES PSICOLINGÜÍSTICOS

Apesar do mapeamento detalhado nos capítulos anteriores, não foi encontrado nenhum estudo de natureza experimental ou de revisão teórica sobre a questão da concretude/abstração em PB e sobre aquisição de linguagem. Por isso, esta tese preenche essa lacuna ao prover dados experimentais acerca das etapas iniciais da aquisição linguística de itens concretos e abstratos.

Em relação à seleção de palavras para o experimento, dentro do escopo de interesse desta pesquisa em itens lexicais de natureza concreta e abstrata, foram consultados alguns inventários de referência internacional para orientar a escolha dos itens, incluindo o *MacArthur-Bates Communicative Development Inventory* (MB — CDI), o CDI para o Português Europeu — Forma reduzida: Nível I (Viana *et al.*, 2018); o *Wordbank*²⁰, o *LexporBR-Infantil*²¹ (Estivalet *et al.*, 2017) e a adaptação do CDI para o português brasileiro²², elaborada pela professora Elizabeth Reis Teixeira da Universidade Federal da Bahia. A justificativa para a utilização desses inventários recai no fato de que esses são materiais de referência na área de AL, uma vez que suas bases de dados são compostas por itens lexicais já compreendidos ou compreendidos e produzidos por bebês.

Em um primeiro momento, realizamos um levantamento de 20 palavras no banco de dados do *Wordbank*. É oportuno mencionar que essa base permite ao pesquisador verificar a produção de itens lexicais específicos por crianças em determinadas faixas etárias (ambos os dados podem ser selecionados de acordo com o seu interesse de pesquisa). Esses 20 itens lexicais foram selecionados, a partir de palavras que já haviam sido utilizadas em experimentos psicolinguísticos, conduzidos com bebês, que investigaram a aquisição lexical de itens concretos e abstratos (Bergelson; Swingley, 2012, 2013), desde que essas palavras fossem dissílabas no PB.

No entanto, durante a busca por esses itens lexicais alguns problemas foram encontrados, como o fato de que não há registro da produção de palavras no PB no *Wordbank*.

²⁰O *Wordbank* é um banco de dados aberto sobre vocabulário infantil. É uma base com dados de colaboradores de todo o mundo. Ele possui dados do MacArthur-Bates Communicative Development Inventory (MB-CDI). Para mais detalhes, acesse: <http://wordbank.stanford.edu/>.

²¹Para acessar o LexporBR-infantil, acesse: <https://www.lexicodoportugues.com/infantil/>.

²²A versão do CDI para o PB foi elaborada pela professora Elizabeth Reis Teixeira da Universidade Federal da Bahia. A decisão de readaptar o CDI do português brasileiro foi comunicada à professora Elizabeth Teixeira para que esta pesquisa fosse cadastrada na base de projetos que utilizam a versão autorizada do CDI brasileiro e, então, a autorização foi concedida.

Existe apenas o banco de dados do português europeu (doravante PE), o que não permitiria uma seleção de estímulos adequada para o experimento sobre aquisição de linguagem com crianças brasileiras. Ademais, os 20 itens lexicais selecionados foram verificados em relação à taxa de produção por crianças de 24 meses. Isso significa que seriam escolhidas apenas as palavras cuja taxa de produção estivesse entre 50% e 100%, tanto no PE quanto no inglês e no francês, línguas nas quais os outros experimentos psicolinguísticos em aquisição de linguagem, reportados na literatura, foram realizados.

Ainda assim, como mencionado anteriormente, não se julgou metodologicamente adequado utilizar parâmetros da taxa de produção vocabular pelas crianças de línguas que não fossem o PB. Outro problema de ordem metodológica em relação à presente pesquisa era o fato de que nenhum dos inventários de palavras consultados, como o CDI para o Português Europeu, o *Wordbank*, o LexporBR-Infantil e a adaptação do CDI para o português brasileiro possuíam classificações referentes ao nível de concretude/abstração dos itens lexicais²³.

Visto que um dos objetivos da presente pesquisa é verificar as diferenças cognitivas na aquisição de itens concretos e abstratos por crianças brasileiras, não seria pertinente utilizar os inventários mencionados acima para selecionar os itens lexicais para o experimento aqui desenvolvido. Mesmo os dados do LexporBR, referentes às taxas de compreensão e de produção das crianças, ainda seriam insuficientes para correlacionar os dados de interesse, ou seja, a taxa de produção e a definição de concretude.

Diante desses desafios, criamos uma nova versão do CDI para o PB, com o objetivo de inserir itens²⁴ de natureza abstrata e elaborar um novo inventário de palavras, i.e., uma nova base de dados para que a seleção lexical fosse realizada. Para a elaboração desse novo inventário de palavras, utilizamos a base da lista de palavras do CDI do português brasileiro, o qual é a versão autorizada do *MacArthur-Bates Communicative Development Inventories* adaptada para o PB. Esse formulário é dividido em três partes, sendo a primeira a lista de palavras, a segunda sobre orações e gramática e a terceira sobre informações gerais.

Dado o interesse da presente pesquisa na aquisição lexical de nomes, algumas outras categorias de palavras foram excluídas, a saber: sons de coisas e de animais; rotina diária e fórmulas sociais; palavras relacionadas ao tempo; quantificadores e locativos; palavras de

²³A definição do que é um item concreto e abstrato é uma discussão bastante heterogênea na literatura linguística e até de outras áreas. Neste momento, não se pretende aprofundar em nenhuma abordagem. Apenas descrever quais foram os critérios e as motivações para selecionar as palavras para este experimento.

²⁴No que tange à versão do CDI brasileiro, este apresentava itens funcionais, as quais, por vezes, são caracterizadas como itens abstratos. No entanto, visto que a presente pesquisa se destina a investigar a aquisição de nomes, i.e., as chamadas palavras de conteúdo, não foram encontrados nomes abstratos nesse inventário. Por essa razão, optamos por elaborar uma nova adaptação do CDI para o português brasileiro.

ação; verbos auxiliares; perguntas; artigos; preposições; pronomes e conectores. Além disso, a parte II (Orações e Gramática) do CDI do PB foi excluída e alguns outros critérios foram alterados, como a disposição das palavras ao longo do formulário.

A adaptação do CDI para o PB possuía itens lexicais da categoria dos substantivos e dos adjetivos. Não obstante, foram inseridas seções referentes a palavras abstratas, separadas por categorias semânticas, assim como todos os modelos de CDI (não só o do PB). Também, foram inseridas, a partir de critérios pessoais da pesquisadora responsável, seções referentes a palavras abstratas, separadas por categorias semânticas, assim como todos os modelos de CDI (não só o do PB). Desse modo, as seções de palavras abstratas introduzidas foram: *Emoções e Sentimentos* e *Seres fantásticos e Personagens*. Além da categoria *Pessoas e Atributos/Características* que já existiam no CDI do PB, mas foram alteradas para a inserção de novas palavras.

Embora cada questionário tenha tido, respectivamente, 101 respostas válidas de responsáveis por crianças entre 24 e 30 meses de vida e por 288 adultos que categorizam as palavras em relação à concretude/abstração, percebemos, ao cruzar os dados de ambos os questionários, que havia muitos adjetivos e poucos substantivos. Portanto, ambos os questionários foram refeitos²⁵ e os dados serão apresentados a seguir.

5.1 PRÉ-TESTE 01: *QUAIS PALAVRAS SUA CRIANÇA JÁ APRENDEU?*

Este formulário de palavras, intitulado *Quais palavras sua criança já aprendeu?*, foi readaptado após constatarmos que os dados coletados no formulário *Sua Criança Sabe essas Palavras?* eram insuficientes para a seleção dos itens lexicais do teste psicolinguístico. Por essa razão, mantivemos a categorização das palavras pelos responsáveis entre “compreende” e “compreende e diz”, conforme o modelo do CDI. No entanto, todos os adjetivos que apresentavam taxa igual ou acima de 50%, no primeiro questionário e que haviam sido classificados como abstratos, foram alterados para substantivos.

Visto que uma primeira versão deste questionário já havia sido respondida por responsáveis de crianças entre 24 e 30 meses de vida, outro critério adotado para a nova versão foi retirar as palavras cujos percentuais de produção, i.e., cujas taxas de produção (classe *Compreende e diz*), no questionário *Sua Criança Sabe essas Palavras?*, estavam

²⁵A lista de palavras referente aos dois primeiros questionários que foram descartados pode ser consultada no apêndice A.

abaixo de 49%²⁶.

Para além desse critério metodológico, alguns responsáveis se queixaram, na primeira versão, da extensão do formulário. Já que alguns itens lexicais apresentavam um percentual de produção muito abaixo do nível de interesse para a seleção de palavras para o experimento, adotamos esse critério de exclusão para tentar diminuir o tamanho do questionário e reduzir o cansaço dos participantes.

O formulário *Quais palavras sua criança já aprendeu?* foi, também, desenvolvido por meio da plataforma do *Google Forms*, seguindo a forma da versão autorizada e adaptada do CDI brasileiro, com as alterações descritas acima. A divulgação foi realizada por meio digital, sobretudo por redes sociais (*whatsapp*, *instagram*) e *e-mail*. Este questionário foi segmentado em quatro seções. Na primeira, tinha a identificação da pesquisadora responsável e de sua orientadora, uma descrição resumida da pesquisa e a explicação de que esta era uma segunda adaptação do formulário de palavras. Foi explicado que mesmo que o participante já tivesse contribuído com a primeira versão, ele poderia realizar esta segunda.

Na seção dois, havia o Termo de Consentimento Livre e Esclarecido e o responsável deveria indicar se estava (ou não) de acordo em contribuir com a pesquisa. Na terceira, tinham as perguntas sobre a criança e, na quarta, perguntas sobre os responsáveis. Na quarta e última seção, as listas de palavras eram apresentadas em seções semânticas, como no questionário 01 (e no CDI), para que os responsáveis marcassem se a criança compreendia ou se já compreendia e dizia a palavra apresentada.

Em relação a divisão do questionário, foram mantidas as 14 seções dentre as categorias semânticas, já presentes no CDI, eram elas: *animais*; *brinquedos*; *roupas e acessórios*; *partes do corpo* (ou itens relacionados); *emoções/sentimentos*; *comidas e bebidas*; *lugares fora da casa* (ou *profissionais*); *seres fantásticos e personagens*; *objetos externos*; *objetos em geral*; *pessoas ou relações*; *atributos/características*; *veículos (reais e de brinquedo)* e *móveis e aposentos*.

Conforme descrito anteriormente, o número de palavras foi reduzido, em virtude da taxa de produção no formulário 01. As palavras *tambor*, *amendoim*, *máquina de lavar*, *colar*, *saia*, *bruxa* que haviam tido um percentual de 49% na taxa de produção do questionário 01 foram inseridas nesta nova versão, visto que estavam próximas ao nível da chance. Foram, ainda, inseridos alguns substantivos, a saber: *limpeza*, *sujeira*, *amiguinho*, *amiguinha*, *cafuné*,

²⁶A lista completa de palavras cujos percentuais de produção pelas crianças ficaram abaixo de 50% e que, portanto, foram excluídas para o questionário, pode ser consultada por meio da planilha, disponível em: https://docs.google.com/spreadsheets/d/1JdCthp2ar_E2aYf_6c4kHSi1yDWC1zEJ0AE1IPDBi28/edit?usp=sharing.

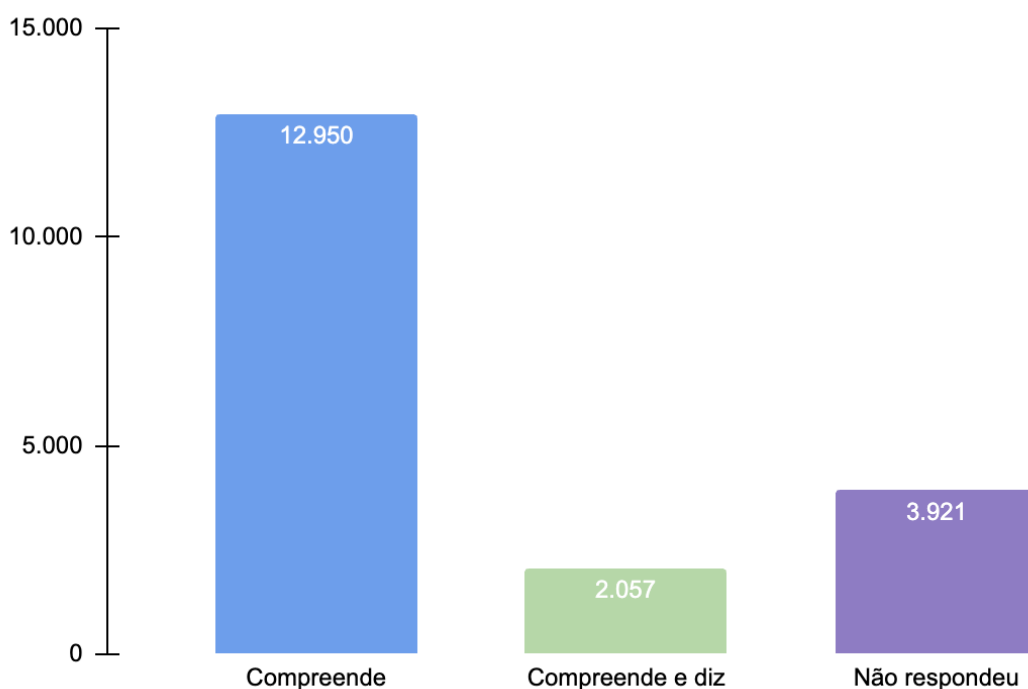
injeção, agulha, espera, riqueza, lábio, boca, machucado, vômito, sangue e enjoo à critério da pesquisadora responsável e de sua orientadora. Este formulário contou, portanto, com 338 palavras. Na última parte, havia um agradecimento e um pedido para que o participante compartilhasse a pesquisa com outros responsáveis por crianças na faixa etária de interesse desta pesquisa.

O formulário denominado: *Quais palavras sua criança já aprendeu?*, teve 85 respostas. Assim como nos formulários anteriores, a primeira etapa consistiu na análise dos dados para descartar os participantes que não estivessem dentro dos critérios de seleção. Diante disso, foram eliminados os dados de 29 crianças pelos seguintes motivos: idade superior a 30 meses (14); inferior a 24 meses (11); bilíngues (2); e ausência de informações sobre problemas congênitos (2). Isto posto, foram contabilizados na amostra final os dados de 56 crianças com idades entre 24 e 30 meses.

Como o responsável poderia selecionar, concomitantemente, os campos *Compreende* e *Compreende e diz*, para um mesmo item lexical, foi feita a correção automática dos campos em que ambas as opções tivessem sido inseridas, i.e., em que constavam tanto a opção *Compreende*, quanto a *Compreende e diz* para só *Compreende e diz*. Posteriormente, os dados finais referente as 56 crianças, foram compilados e analisados no Excel.

No que tange aos resultados, visto que os responsáveis poderiam selecionar a opção: *Compreende e diz* ou *Compreende* ou, então, deixar em branco (se a criança ainda não falasse aquela palavra) para cada item lexical, obtivemos, respectivamente, 12.950 respostas, seguidas de 2.057 e 3.921 em uma base de dados vocabular com 18.928 palavras no total (equivalente a 338 itens lexicais respondidos por 56 responsáveis por crianças entre 24 e 30 meses de vida).

Os dados supramencionados podem ser checados no gráfico 1, abaixo:

Gráfico 1 — Respostas ao questionário — *Quais palavras sua criança já aprendeu?*

Fonte: elaboração própria (2024).

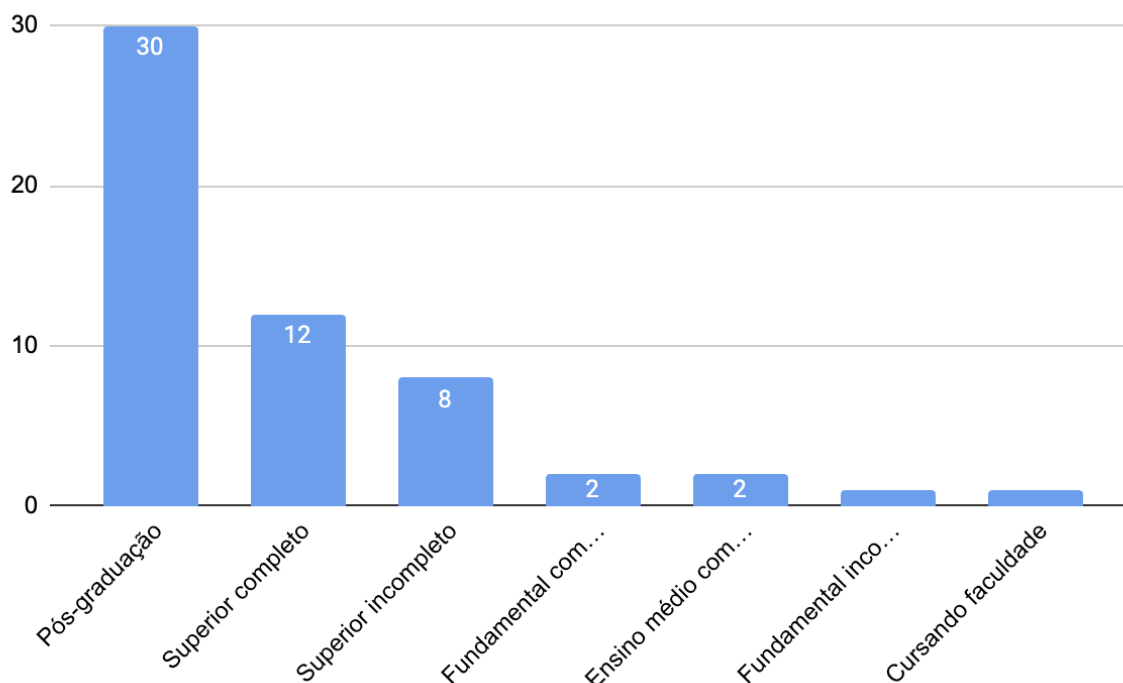
As respostas eram referentes a 56 crianças entre 24 e 30 meses de idade (idade média: 27m), falantes nativas do PB, monolíngues e sem histórico de doenças congênitas, sendo 29 crianças do sexo feminino e 27 do sexo masculino. Do total, 50 crianças eram nascidas no Sudeste, 3 eram nascidas no Sul e 3 eram nascidas no Nordeste. Em relação ao local onde a criança morava no momento do preenchimento do formulário, não houve nenhuma resposta diferente em relação àquela do local de nascimento.

No que tange ao número de irmãos, das 56 crianças, 27 não tinham nenhum irmão, 23 tinham um irmão, 05 tinham dois irmãos e 01 tinha três ou mais irmãos. Dos 23 participantes que tinham só um irmão, 02 desses eram gêmeos, 16 deles tinham irmãos mais velhos, 01 responsável reportou estar grávida no momento do preenchimento do questionário e um não soube informar a idade do irmão, pois era pela outra parte da família e apenas 03 tinham irmãos mais novos. Já das crianças que tinham dois irmãos, 02 eram gêmeos e tinham um irmão mais novo, e 02 tinham irmãos mais velhos. A única criança que tinha 03 irmãos, possuía 02 irmãos mais velhos e 01 mais novo.

Em relação ao nível educacional dos responsáveis que estavam preenchendo o formulário, 30 reportaram ter pós-graduação, 12 com ensino superior completo, 08 com ensino superior incompleto, 02 com ensino médio completo, 02 com ensino fundamental

completo, 01 com ensino fundamental incompleto. Todos eles eram brasileiros natos. Esses dados podem ser conferidos no gráfico 2, abaixo:

Gráfico 2 — Respostas ao questionário — *Grau de escolaridade — responsável 01*

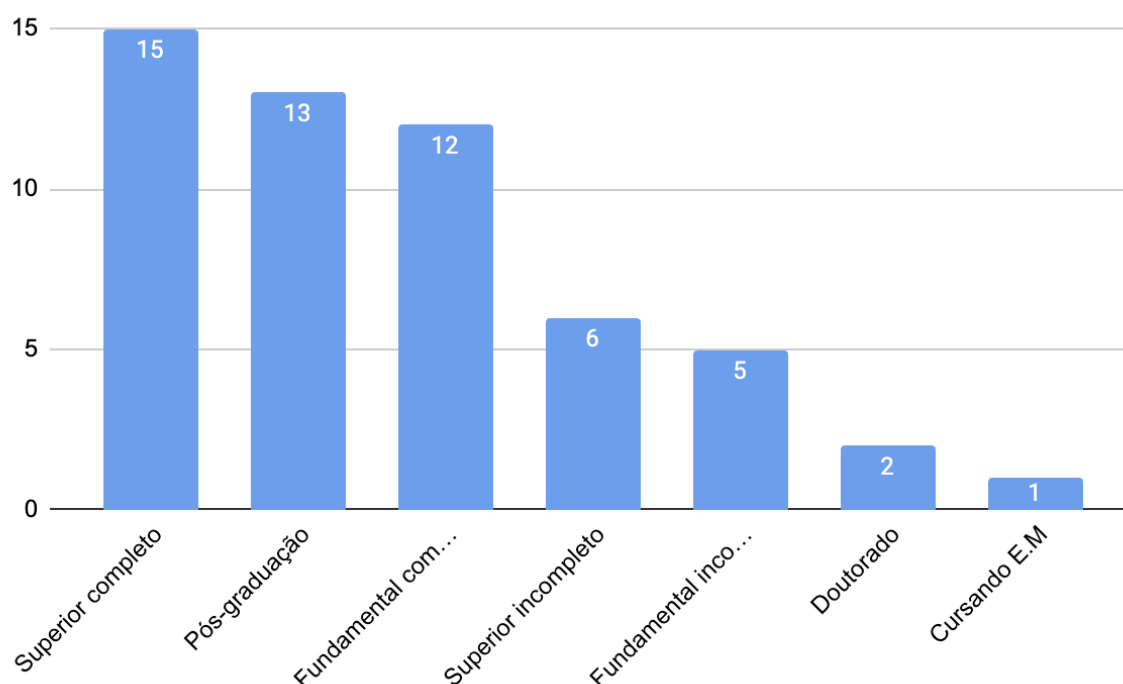


Fonte: elaboração própria (2024).

Também foi perguntado se mais alguém era responsável pela criança, se dividia tempo com ela, qual era o seu grau de parentesco e qual era o nível educacional desse outro cuidador. Do total de 56 participantes, 53 indicaram que havia um outro responsável dos quais: 15 possuíam pós-graduação, 15 tinham graduação completa, 06 tinham ensino superior incompleto, 01 estava cursando o ensino médio, 12 tinham ensino fundamental completo, 05 tinham ensino fundamental incompleto. Todos eram brasileiros natos.

Essa relação do nível educacional desse outro cuidador pode ser verificada no gráfico 3:

Gráfico 3 — Nível de escolaridade do outro responsável



Fonte: elaboração própria (2024).

É importante mencionar que esses dados referentes ao perfil socioeconômico e demográfico dos responsáveis se constituem como parâmetros importantes para pesquisas em AL, uma vez que estudos em neurociência cognitiva (Cuartas *et al.*, 2024; Fernald *et al.*, 2013; Miller, 2019; Pace *et al.*, 2017; Sperry; Sperry, 2019) têm apontado que quanto maior o nível socioeconômico dos cuidadores, maior é o impacto no desenvolvimento vocabular e em outros aspectos do desenvolvimento cognitivo, como as habilidades socioemocionais, funções executivas (Rathore *et al.*, 2023; Romeo *et al.*, 2018).

5.1.2 PRÉ-TESTE 02: *CONCRETA OU ABSTRATA: COMO VOCÊ CLASSIFICA ESSAS PALAVRAS?*

Como essa discussão sobre o que determina um item lexical ser concreto ou abstrato na literatura linguística é bastante heterogênea, adotar uma visão teórica para selecionar as palavras poderia não ser um critério adequado ao propósito do experimento. Por isso, optamos por elaborar um outro questionário para a classificação de itens lexicais em concretos e abstratos por falantes adultos, nativos do PB.

Elaboramos, então, pelo *Google Forms*, o pré-teste *Concreta ou abstrata: como você*

classifica essas palavras? e a sua divulgação foi realizada por meio digital, sobretudo por redes sociais (*WhatsApp, Instagram*) e *e-mail*. Ele foi dividido em três seções: a primeira continha a apresentação da pesquisa, a explicação de que era um segundo questionário para adultos e a identificação da autora e de sua orientadora; a segunda incluía o Termo de Consentimento Livre e Esclarecido e a lista de palavras; e a terceira apresentava um agradecimento. As palavras foram organizadas em blocos semânticos, assim como no CDI. No total, havia 338 itens lexicais.

Nesta versão, o participante deveria classificar as palavras entre as opções *concreta*, *abstrata* ou *não sei definir*. Não era possível selecionar dois campos duplamente, i.e., tanto o campo *concreto* quanto o *abstrato*, nem mesmo deixar o campo em branco. Portanto, a resposta era sempre obrigatória. Como descrito na análise de resultados, quando o participante deixava muitos campos em branco, ele era descartado, pois foi considerado que ele havia abandonado o questionário.

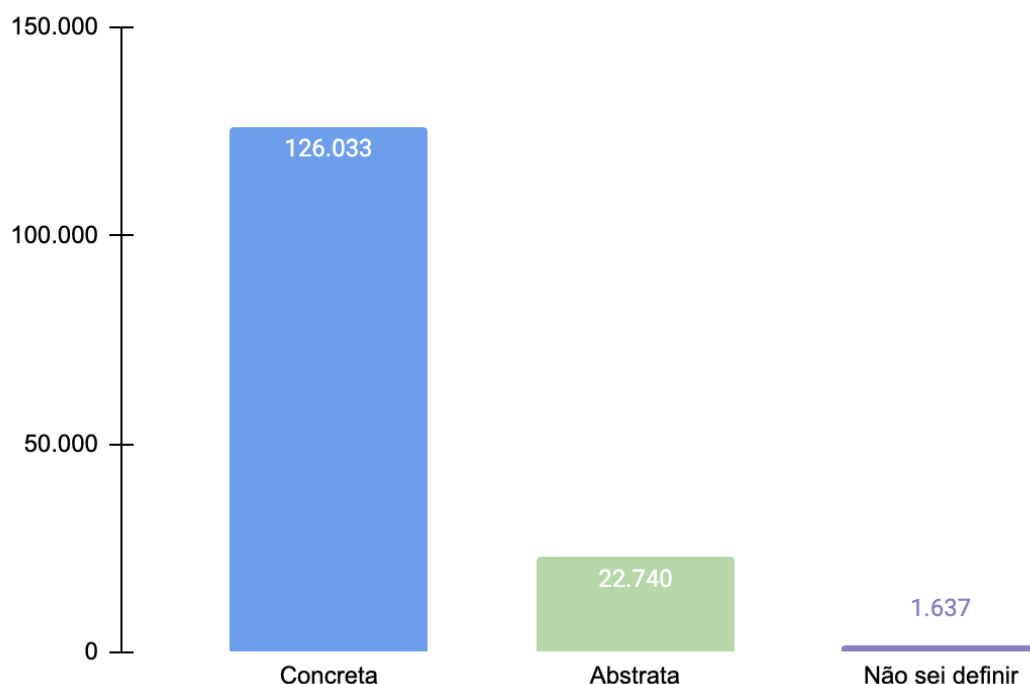
O formulário *Concreta ou abstrata: como você classifica essas palavras?* foi respondido por 472 participantes. Foi realizada, como primeira etapa, uma filtragem dos dados para a análise das respostas. Desses 472 participantes, foram descartados os dados de 25 pessoas que declararam terem sido alfabetizadas em outra língua que não o PB e 02 que não eram brasileiras natas.

Após a exclusão dos participantes que não se encaixavam no perfil, os dados relativos às respostas foram analisados. Conforme reportado anteriormente, os participantes poderiam inserir as opções *concreta*, *abstrata* ou *não sei definir* para cada uma das palavras apresentadas. Todas as listas de palavras foram marcadas como campos obrigatórios para os participantes a fim de evitar que o teste fosse enviado com muitos campos em branco. Além disso, não foi permitido inserir dois campos simultaneamente. Por fim, os dados de cada palavra por participantes foram compilados com o objetivo de verificar as respostas de cada participante para cada palavra. Para tanto, os dados, também, foram submetidos para análise no Excel.

As respostas poderiam ser inseridas em três níveis: *concreta*, *abstrata* e *não sei definir*, resultando, respectivamente, em 126.033 respostas, seguido de 22.740 respostas e 1.637. Esta base de dados contou com 150.410 palavras, provenientes de 338 itens lexicais categorizados por 445 brasileiros natos.

Esses dados estão presentes no gráfico 4, abaixo.

Gráfico 4 — Respostas ao questionário — *Concreta ou abstrata: como você classifica essas palavras?*



Fonte: elaboração própria (2024).

A respeito do perfil socioeconômico dos participantes, todos os 445 participantes eram brasileiros nativos e falantes do PB, sendo 138 homens, 302 mulheres e 03 pessoas se autodeclaradas como *outro*. Os participantes pertenciam aos seguintes grupos etários: 18–24 anos (73 participantes), 25–30 anos (82 participantes), 31–35 anos (82 participantes), 36–40 anos (45 participantes), 41–45 (52 participantes), 46–50 (32 participantes), 51–55 (23 participantes), 56–60 (26 participantes), 61–65 (13 participantes), 66–70 (11 participantes) e acima de 70 (6 participantes).

Os participantes eram naturais de regiões diversas do Brasil, sendo 330 do Sudeste, 43 do Nordeste, 41 do Sul, 15 do Centro-Oeste e 16 do Norte. No que tange ao local onde moravam, no momento do preenchimento do formulário, 312 residiam no Sudeste, 40 no Sul, 39 no Nordeste, 18 no Centro-Oeste e 17 no Norte do país. Não obstante, 20 participantes declararam que não moravam no Brasil.

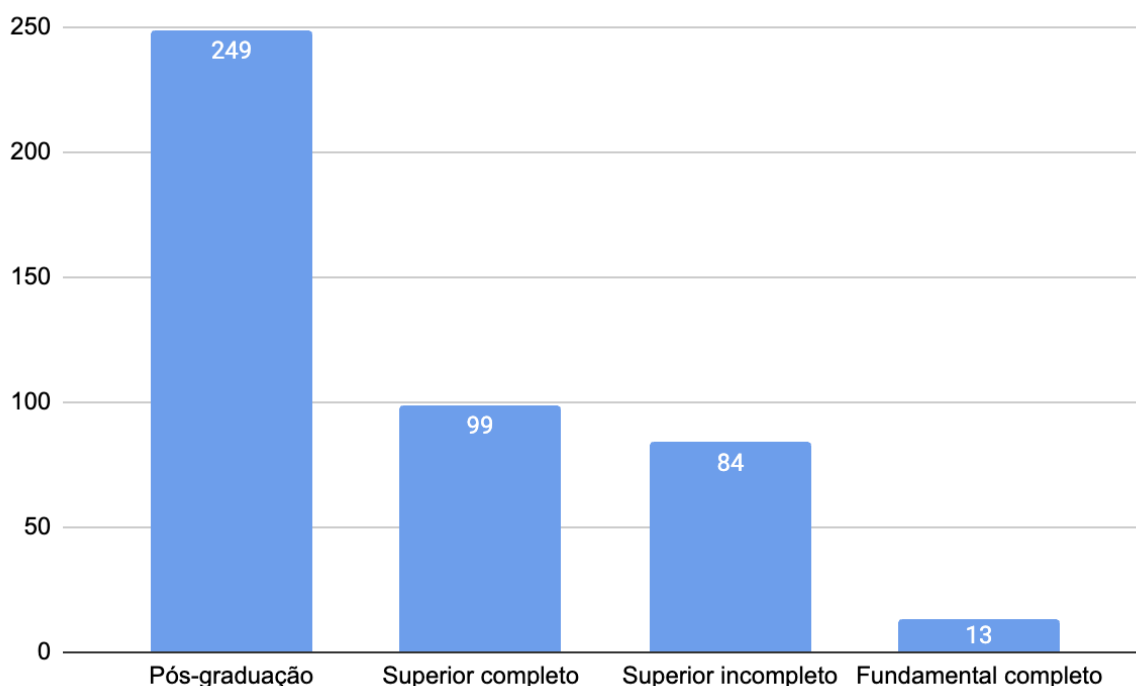
Sobre falar um outro idioma, 281 indicaram que falavam outra(s) língua(s) estrangeira(s), além do PB, dos quais 194 indicaram falar apenas uma língua adicional, 77 declararam falar duas línguas estrangeiras, 31 indicaram falar três idiomas adicionais e 06 disseram falar quatro ou mais línguas.

No que tange ao nível educacional dos participantes, 259 indicaram possuir

pós-graduação, 99 nível superior completo, 84 ensino superior incompleto, seguido de 56 com graduação incompleta, 13 com ensino fundamental completo.

Essas informações estão apresentadas no gráfico 5 abaixo.

Gráfico 5 — Nível de escolaridade dos respondentes ao questionário — *Concreta ou abstrata: como você classifica essa palavra?*



Fonte: elaboração própria (2024).

Na próxima seção, será descrito como os itens lexicais foram selecionados para o experimento psicolinguístico.

5.1.3 CRUZAMENTO DAS RESPOSTAS DOS PRÉ-TESTES PARA A SELEÇÃO DOS ITENS LEXICAIS DO EXPERIMENTO

As respostas desses formulários foram cruzadas, considerando a taxa de produção das crianças, a partir da classe *Compreende e Diz* da variável resposta no Questionário *Quais palavras sua criança já aprendeu?* e a classificação dos itens lexicais em relação à concretude, no Questionário: *Concreta ou abstrata: como você classifica essas palavras?*, pelos adultos.

Em relação ao formulário infantil, foram selecionados os itens vocabulares que tivessem valores a partir de 50%, portanto acima do nível da chance, no que tange à classe

Compreende e diz, correspondente à produção oral da criança. Além do cruzamento dessas informações (i.e., a taxa de produção pelas crianças e a categorização do item lexical pelos adultos). Dos 338 itens lexicais que compunham a lista vocabular dos formulários, já considerando a taxa de produção acima de 50%, ficaram 288 palavras no total (vide apêndice B).

É pertinente destacar que esse critério de análise da proporção das taxas de produção dos itens vocabulares pelas crianças é o mesmo adotado na base de dados do Wordbank, mencionada anteriormente. Portanto, trata-se de uma prática amplamente difundida nos estudos em aquisição de linguagem para a seleção de estímulos experimentais. É importante salientar que, com a substituição de adjetivos por substantivos e a inserção de alguns outros por critério da pesquisadora, a escassez de dados para a seleção dos estímulos experimentais foi parcialmente corrigida. Essa questão será melhor explicada na seção subsequente.

5.1.4 SELEÇÃO LEXICAL PARA OS ESTÍMULOS EXPERIMENTAIS A PARTIR DA ANÁLISE DOS DOIS FORMULÁRIOS

Para a seleção lexical dos itens que compõem o experimento desta tese, foram consideradas, após a análise dessas 288 palavras, os índices da classe *Compreende e Diz* do Questionário — *Quais palavras sua criança já aprendeu?*²⁷ e correlacionadas com a definição de concretude e abstração do Questionário — *Essa palavra é concreta ou abstrata?*, respondido por falantes nativos do PB. Foram selecionados, então, 23 itens lexicais, dos quais 12 foram categorizados por falantes do PB como concretos (*fogão; balão; pijama; pizza; beijo; casaco; tigre; prato; tênis; abraço; buraco; menino*) e os outros 11 como abstratos (*quente; festa; fome; amor; frio; medo; vazio; sujo; carinho; escuro; alegria*).

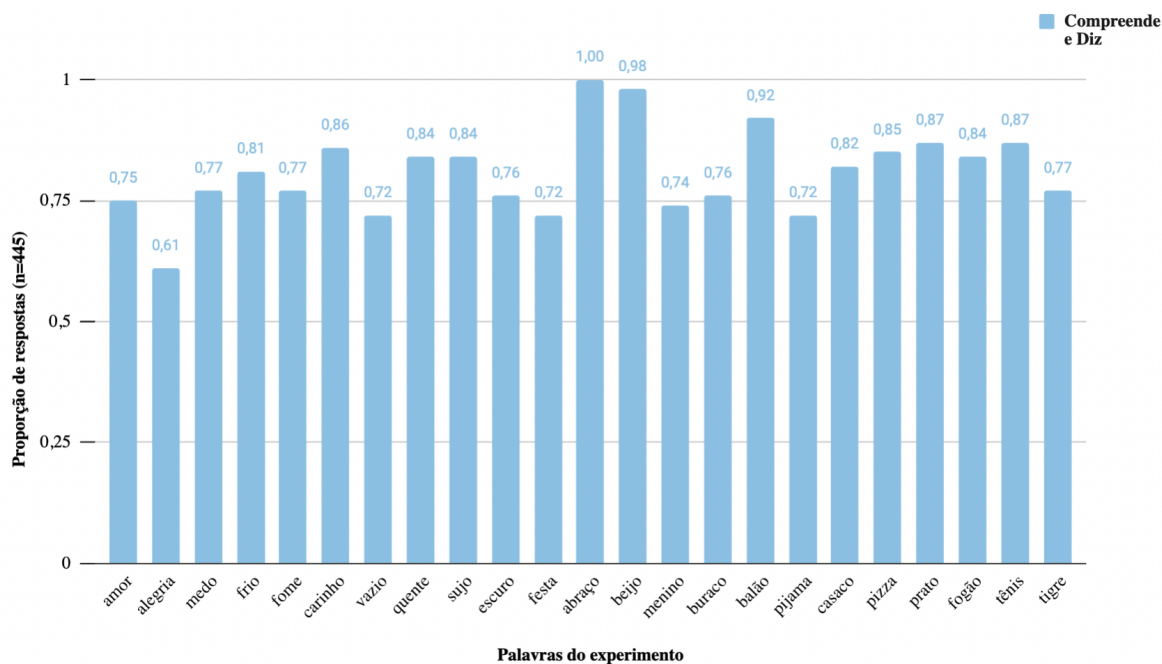
O item lexical *sono*, apesar de não ter sido selecionado a partir do formulário realizado pela autora desta tese, foi inserido para compor as palavras do experimento, considerando o inventário do Lexporbr-Infantil (Estivalet *et al.*, 2017). O Lexporbr-Infantil também foi utilizado para realizar uma análise de alguns parâmetros psicométricos, como a taxa de

²⁷Assim como nos CDI's, a taxa de produção dos itens lexicais foi calculada dividindo-se o total de respostas para a opção "já compreende e já diz" daquele item lexical pelo número de participantes totais (56) no Questionário - *Quais palavras sua criança já aprendeu?*.

frequência, a escala *Zipf*²⁸ e a imageabilidade dos itens selecionados para o experimento desta tese.

O gráfico 6 abaixo apresenta a taxa de produção de cada um desses itens lexicais.

Gráfico 6 — Resultados do CDI administrado com 56 pais de crianças entre 24 e 30 meses de idade

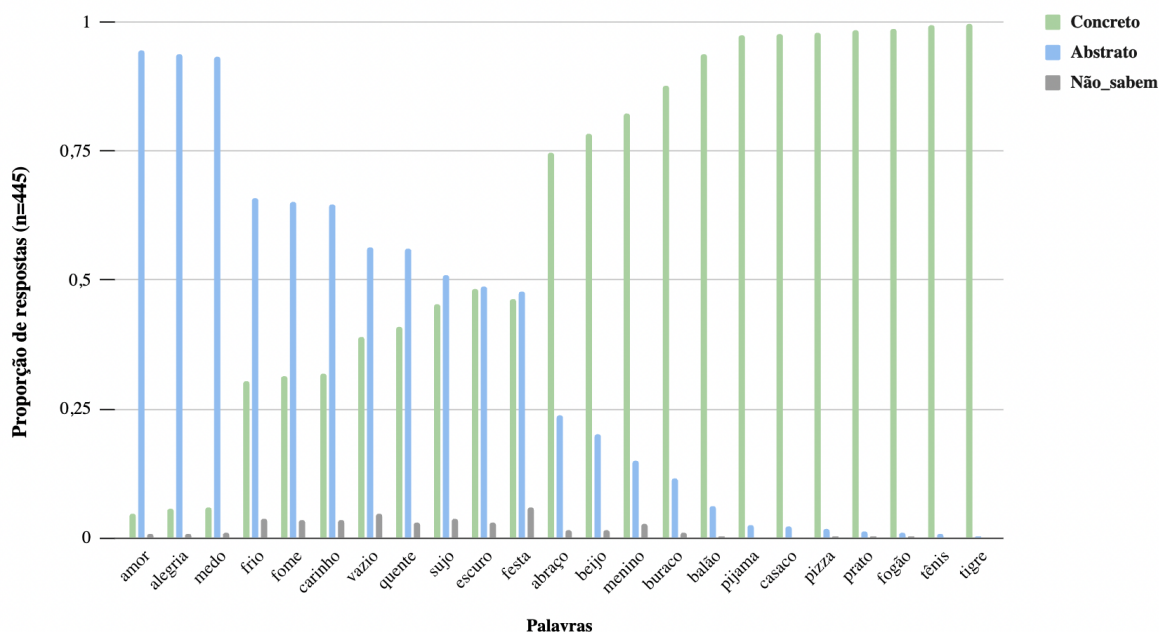


Fonte: elaboração própria (2024).

A seguir, está a definição de concretude e abstração do Questionário — *Essa palavra é concreta ou abstrata?*, realizada por 445 falantes nativos do PB, das palavras experimentais que compõem o experimento desta tese.

²⁸Conforme explicado por Van Heuven et al (2014), a escala Zipf mede a frequência lexical em valores de 1 a 7, sendo que palavras com $\text{Zipf} \leq 3$ são de baixa frequência e palavras com $\text{Zipf} \geq 4$ são de alta frequência. O cálculo é feito por meio de $\log_{10}$ (frequência por bilhão de palavras) ou $\log_{10}$ (frequência por milhão) + 3. Assim, um valor de Zipf 1 corresponde a 1 ocorrência a cada 100 milhões de palavras, Zipf 2 a 1 em 10 milhões, Zipf 3 a 1 em 1 milhão, e assim sucessivamente. A escala recebeu esse nome em homenagem ao linguista norte-americano George Zipf (1902–1950), que analisou, de forma pioneira, as regularidades da distribuição da frequência das palavras e formulou a chamada lei de Zipf (1949).

Gráfico 7 — Categorização de palavras selecionadas em relação à concretude pelos adultos, falantes do PB



Fonte: elaboração própria (2024).

É importante destacar que, das 24 palavras utilizadas no experimento, 19 são nomes (*amor; alegria; medo; carinho; festa; abraço; beijo; menino; buraco; balão; pijama; casaco; pizza; prato; fogão; tênis; tigre; sono; fome*) e 5 são adjetivos (*frio; vazio; quente; sujo; escuro*) com possibilidade de uso como nome.

Esses cinco ficaram na lista por serem palavras de alta frequência no vocabulário infantil e por possuírem versatilidade categorial para desempenhar diferentes funções sintáticas e semânticas em uma frase, como mostram os exemplos abaixo:

- (I) O vazio da geladeira me diz que é hora de ir ao supermercado.
- (II) Tenho pavor do escuro.
- (III) O sujo aqui no casaco está difícil de tirar.
- (IV) O leite estava tão quente que queimou a boca do bebê.

Na tabela 1 estão compiladas as informações dos CDIs adaptados por nós, exibindo todas as etapas de elaboração da lista final de palavras que compuseram o nosso teste. A tabela também inclui os resultados dos questionários sobre definição de concretude/abstração, aplicados a adultos brasileiros, inclusive dos dois primeiros questionários descartados. As colunas indicam: *lexema*, a forma lexical básica (lema) da palavra; *nb_sílabas*, o número de sílabas; e *nb_letras*, o número de letras.

Tabela 1 — Resultados dos CDIs adaptados

Lexema	Taxa de produção pela criança (Nível Compreende e Diz) - quest 01	Concretude (resposta dos adultos) - quest 02	Taxa de produção ²⁹ pela criança (Nível Compreende e Diz) - quest 03	Concretude (resposta dos adultos) - quest 04	nb_silabas	nb_letras
fogão	0,69	concreto	0,84	concreto	2	5
quente	0,85	abstrato	0,84	abstrato	2	6
balão	0,88	concreto	0,92	concreto	2	5
festa	0,76	concreto	0,72	abstrato	2	5
pizza	0,75	concreto	0,85	concreto	2	5
fome	0,69	abstrato	0,77	abstrato	2	4
pijama	0,73	concreto	0,72	concreto	3	6
sono	-	-	-	-	2	4
casaco	0,76	concreto	0,82	concreto	3	6
frio	0,77	abstrato	0,81	abstrato	2	4
beijo	-	-	0,98	concreto	2	5
amor	-	-	0,75	abstrato	2	4
tigre	0,76	concreto	0,77	concreto	2	5
medo	0,74	abstrato	0,77	abstrato	2	4
prato	0,79	concreto	0,87	concreto	2	5
vazio	0,55	abstrato	0,58	abstrato	3	5
tênis	0,76	concreto	0,87	concreto	2	5
sujo	0,84	abstrato	0,84	abstrato	2	4
abraço	-	-	1,00	concreto	3	6
carinho	-	-	0,86	abstrato	3	7
buraco	0,70	concreto	0,76	concreto	3	6
escuro	0,77	abstrato	0,76	abstrato	3	6
menino	0,79	concreto	0,74	concreto	3	6
alegria	-	-	0,61	abstrato	4	7

Fonte: elaboração própria (2024).

²⁹ Como mencionado na seção 5.1.4, lembramos que a taxa de produção desses itens lexicais foi realizada seguindo os parâmetros do CDIs. Para tanto, dividiu-se o total de respostas para a opção "já compreende e já diz" daquele item lexical pelo número de participantes totais (56) no Questionário — *Quais palavras sua criança já aprendeu?*.

Na tabela 2, inserimos as propriedades lexicais e estatísticas, apresentadas na base LexporBR-Infantil (Estivalet et al., 2017), das palavras selecionadas para o experimento psicolinguístico.

Tabela 2 — Comparação dos itens lexicais selecionados para o teste psicolinguístico, extraídos da base do Lexporbr-Infantil de Estivalet (2017)

lexema	freq_lexema	freqM_lexema	freqZS_lexema	nb_silabas	nb_letras	dl_orto
fogão	853	6,66	3,8237	2	5	7
quente	17665	138	5,1399	2	6	6
balão	1551	12,12	4,0834	2	5	23
festa	42979	335,76	5,526	2	5	21
pizza	7644	59,72	4,7761	2	5	2
fome	14449	112,88	5,0526	2	4	18
pijama	1954	15,26	4,1837	3	6	1
sono	4424	34,56	4,5386	2	4	28
casaco	6372	49,78	4,697	3	6	10
frio	11850	92,57	4,9665	2	4	17
beijo	9334	72,92	4,8628	2	5	8
amor	65996	515,57	5,7123	2	4	21
tigre	2090	16,33	4,2129	2	5	3
medo	38898	303,88	5,4827	2	4	24
prato	4696	36,69	4,5645	2	5	16
vazio	4321	33,76	4,5284	3	5	8
tênis	3575	27,93	4,446	2	5	14
sujo	4615	36,05	4,5569	2	4	23
abraço	3705	28,94	4,4616	3	6	13
carinho	1794	14,02	4,1466	3	7	14
buraco	8429	65,85	4,8185	3	6	5
escuro	5700	44,53	4,6486	3	6	7
menino	14711	114,92	5,0604	3	6	6
alegria	4459	34,83	4,542	4	7	5

Fonte: adaptado de Estivalet *et al.* (2017).

As medidas incluídas na tabela são: lexema, que indica a forma lexical básica (lema) da palavra; freq_lexema, a frequência absoluta desse lexema no corpus; freqM_lexema, a frequência normalizada por milhão de ocorrências; e freqZS_lexema, a frequência transformada em escala Zipf, permitindo comparações independentes do tamanho do corpus. As colunas nb_sílabas e nb_letras indicam, respectivamente, o número de sílabas e de letras de cada palavra, enquanto dl_orto representa a distância ortográfica média (OLD20), ou seja, a similaridade da palavra em relação a outras do léxico.

Foram selecionadas palavras que tivessem percentuais de produção próximos nas duas categorias de palavras (as concretas e as abstratas) a fim de evitar, no experimento, um efeito de frequência entre itens com maiores taxas de produção em comparação aos demais. Por exemplo, a palavra ténis apresentou taxa de produção de 0,87%, conforme respostas dos responsáveis ao formulário *Quais palavras sua criança já aprendeu?*, enquanto o item lexical abstrato sujo teve taxa de 0,84%. Essas palavras foram classificadas, respectivamente, como concreta e abstrata pelos adultos, no questionário *Concreta ou abstrata: como você classifica essas palavras?*.

O objetivo com essa produção era evitar um efeito principal de frequência (em virtude de uma taxa de produção muito alta de determinado item lexical em detrimento de outro, dado que palavras com uma taxa de produção maior tendem a ser as mais frequentes no *input* que a criança escuta). A respeito disso, Brysbaert *et al.* (2011, p. 415) esclarece:

A frequência de palavras é uma das variáveis mais importantes nas pesquisas em psicolinguística, estabelecendo que palavras que são lidas e ouvidas mais frequentemente são reconhecidas mais rapidamente do que palavras menos frequentes.

Por essa razão, os pares de palavras (concreta e abstrata) que compõem os *trials* experimentais foram rigorosamente controlados como forma de se evitar um efeito de frequência nos resultados experimentais. Outro fator importante é que, entre os itens lexicais selecionados, todos apresentam uma taxa na escala Zipf entre 3 e 5, sendo apenas a palavra fogão com 3,8, enquanto as demais variam entre 4 e 5. Isso indica que todos os itens vocabulares selecionados são altamente frequentes.

Na seção subsequente, será descrita a plataforma utilizada e o experimento elaborado com vistas a fornecer evidências experimentais do teste controle *Como adultos brasileiros processam palavras concretas e abstratas do Português brasileiro?*

5.2 O EXPERIMENTO CONTROLE: *COMO ADULTOS BRASILEIROS PROCESSAM PALAVRAS CONCRETAS E ABSTRATAS DO PORTUGUÊS BRASILEIRO?*

O experimento *Como adultos brasileiros processam palavras concretas e abstratas?* foi elaborado no *Lookit* (Scott *et al.*, 2017; Scott; Schulz, 2017), uma plataforma online desenvolvida pelo Massachusetts Institute of Technology (MIT) em que é possível desenvolver experimentos cognitivos. Como toda plataforma experimental, apresenta vantagens e limitações em seu uso, as quais serão descritas abaixo.

Os testes elaborados nesta plataforma são hospedados no *Children Helping Science*, um domínio no qual as famílias cadastradas podem participar, de forma remota, de pesquisas científicas, desenvolvidas por diferentes instituições internacionais que estejam associadas ao MIT. Portanto, para elaborar um experimento nesta plataforma é preciso que a instituição ao qual o pesquisador está associado assine um termo institucional³⁰ para estabelecer o vínculo entre a universidade e o MIT. Essa etapa será melhor detalhada na parte de procedimentos.

Em relação à elaboração do experimento, a personalização de cada protocolo experimental é determinada para o programa em um tipo de formato de texto especial chamado *JavaScript Object Notation* — JSON, que será lido pelo executor de experimentos do *Lookit*. Essa descrição em JSON é o que permite a estruturação de como o teste vai ser lido e executado pela plataforma.

Caso o pesquisador não tenha familiaridade com JSON ou com linguagem de programação, como *Java Script* — que pode ser utilizado para definir alguns parâmetros, a plataforma disponibiliza um tutorial³¹ para instruir os interessados a elaborarem o seu estudo na plataforma. Por essa razão, a execução de experimentos no *Lookit*, para quem não tem conhecimento em JSON e/ou em JavaScript pode ser, a princípio, desafiadora.

Outro ponto negativo é que a análise dos resultados é feita manualmente. Em experimentos com rastreamento ocular, por exemplo, o pesquisador precisa codificar os movimentos oculares dos participantes, o que torna o processo mais trabalhoso se comparado ao uso do rastreador ocular convencional, o qual já fornece os dados referentes ao rastreamento automaticamente. O detalhamento de como a análise foi feita, qual foi o *time point* utilizado para computar a posição do olhar será descrito na seção 5.2.2, na parte referente aos *procedimentos*.

³⁰Para mais informações sobre como filiar o seu laboratório ao *Lookit*, acesse: <https://lookit.readthedocs.io/en/master/researchers-start-here.html>.

³¹Para acessar o tutorial do *Lookit*, basta acessar o link: <https://lookit.readthedocs.io/projects/frameplayer/en/latest/>.

Apesar dessas características, a utilização desta plataforma fornece vantagens metodológicas significativas no âmbito científico por permitir, por exemplo, que responsáveis participem de pesquisas diretamente de suas casas, sem a necessidade de ir até um laboratório, preservando ao mesmo tempo a confiabilidade científica dos dados coletados nos experimentos de forma segura no próprio *Lookit*. Para isso, é necessário que os participantes façam um cadastro com as informações pessoais, como nome, faixa etária, nome da criança que participará do experimento, a data de nascimento dela, se tem irmãos, se moram em área urbana ou rural, se tem contato com línguas estrangeiras e se há alguma suspeita de doenças psiquiátricas, cognitivas ou sensoriais.

Posteriormente, eles podem acessar estudos de diferentes laboratórios diretamente pelo navegador, sem a necessidade de agendamento ou de videoconferência (a menos que o pesquisador opte por desenvolver uma pesquisa em formato de entrevista em tempo real). Portanto, como mencionado anteriormente, a execução dos testes ocorre de forma remota, utilizando o próprio computador pessoal da família da criança, desde que este seja equipado com uma câmera. Durante essas atividades, a reação dos participantes são registradas pela própria câmera do computador e, ao final do experimento, enviadas ao servidor do CHS, onde ficam disponíveis para análise posterior pelos pesquisadores responsáveis por aquele estudo.

Esse aspecto é positivo, sobretudo diante de todos os desafios metodológicos envolvidos nas pesquisas em AL, já que possibilita que os participantes sejam avaliados em um ambiente domiciliar, trazendo dados mais naturalísticos do que os coletados em um laboratório (Oakes, 2017).

É relevante destacar que estudos que investigaram a viabilidade de pesquisas sobre o desenvolvimento infantil conduzidas de forma *online* compararam o desempenho de crianças pequenas em tarefas realizadas virtualmente e em laboratório. Essas investigações apontaram a eficácia na implementação dos protocolos de pesquisa, a segurança na coleta de dados, a confiabilidade na codificação das medidas dependentes e a capacidade dos pais de conduzir os experimentos de forma autônoma (Kaduk *et al.*, 2023; Scott *et al.*, 2017; Scott; Schulz, 2017b; Slim *et al.*, 2024).

Outra questão de impacto social é que, quanto mais pesquisadores brasileiros realizarem experimentos no *Lookit*, maior será o banco de dados de famílias cadastradas que podem ser convidadas a participar de outras pesquisas. Como fomos os primeiros pesquisadores do Brasil a utilizar essa plataforma, os participantes do nosso estudo são pioneiros na formação dessa base.

Isso representa um aspecto inovador nas pesquisas em aquisição de linguagem, porque a maioria dos estudos em desenvolvimento cognitivo tem se baseado, predominantemente, em dados provenientes dos chamados países *WEIRD* – *Western, Educated, Industrialized, Rich, Democratic* (Singh *et al.*, 2023). Essa predominância limita a capacidade de generalização das pesquisas em desenvolvimento cognitivo, uma vez que os participantes desses países não só representam uma pequena parte da população global (aproximadamente 12%), como também apresentam resultados mais elevados em tarefas cognitivas complexas, como a linguagem, quando comparados aos participantes de países *WILD*³² — *Worldwide, In situ, Local, Diverse* (Alves *et al.*, 2022), como o Brasil.

Essa falta de diversidade nas amostras muitas vezes desconsidera fatores importantes que influenciam o desenvolvimento cognitivo, social e emocional de crianças (Peyre *et al.*, 2017; Singh *et al.*, 2023), como o nível socioeconômico, o acesso à educação, a identidade racial, o contexto sociopolítico e religioso e a origem linguística dos participantes. Essa exclusão de perspectivas de populações menos representadas (Henrich *et al.*, 2010; Singh *et al.*, 2023) pode resultar em uma visão restrita da diversidade das experiências humanas.

Isso tudo pode, como mencionado brevemente na introdução desta tese, limitar a compreensão dos fatores que influenciam o desenvolvimento da linguagem na infância e a interpretação dos dados, pois as conclusões acabam refletindo de forma mais significativa as experiências e as prioridades de grupos economicamente e socialmente privilegiados, que não só ocupam posições de destaque no campo da pesquisa, como também apresentam parâmetros socioeconômicos mais elevados do que os de países em desenvolvimento, como o Brasil.

Nesse contexto, o uso do *Lookit* surge como uma estratégia para incorporar novas ferramentas metodológicas, capaz de ampliar e diversificar as amostras nos estudos em neurociência cognitiva. Além disso, essa plataforma possibilita a aplicação de experimentos de forma mais simplificada, superando limitações logísticas de pesquisas que utilizam o rastreamento ocular como técnica experimental.

5.2.1 ESTÍMULOS

O teste é composto por 24 palavras-alvo (12 abstratas e 12 concretas), utilizando seis pares de imagens por participante (12 imagens ao todo - [consultar apêndice C]) que ilustram cada palavra utilizada no experimento. Os estímulos são apresentados por meio de oito listas

³² A tradução da sigla *WILD* corresponde a: Mundial, In situ, Local, Diversificado.

diferentes (com, no mínimo, seis participantes para cada lista [consultar apêndice D]), de forma que cada um escute apenas uma palavra associada a cada par de imagens, dispostas em posições distintas (ora à direita ora à esquerda) para evitar qualquer viés/preferência em relação à posição das imagens.

Optamos, também, por utilizar as mesmas imagens tanto para os estímulos concretos quanto para os abstratos, no mesmo tipo de tarefa, a fim de evitar possíveis diferenças no desempenho dos participantes. Isso pode ter ocorrido, por exemplo, nos experimentos de Bergelson e Swingley (2012, 2013), nos quais os estímulos eram apresentados ora como imagens, ora como vídeos, ou ainda em função de preferências visuais das crianças (Oakes, 2017) decorrentes de tarefas diferentes, que não necessariamente refletiram uma maior dificuldade no processamento dos estímulos por bebês.

Cada participante ouve apenas seis *trials* experimentais. A condição *concreta* e a *abstrata* é manipulada *within-subjects* de forma que, para cada sujeito, metade dos itens (n=3) aparece na condição concreta e a outra metade (n=3) na condição abstrata, totalizando seis *trials* experimentais, como mencionado anteriormente. A sequência dos itens é pseudoaleatória.

As sentenças experimentais consistem em frases que designam ora a condição concreta como “Olha! Fogão! Viu! Fogão!” ou “Olha! Balão! Viu! Balão!”, ora a abstrata como: “Olha! Quente! Viu! Quente” ou “Olha! Festa! Viu! Festa!”, no caso do *trial* balão-fogão. No *trial* prato-tigre, por exemplo, a criança poderá ouvir, a depender da lista que foi randomizada pelo programa, uma frase que faz referência à condição concreta, como “Olha! Prato! Viu! Prato!” ou “Olha! Tigre! Viu! Tigre!”, ou à abstrata, como “Olha! Vazio! Viu! Vazio!” ou “Olha! Medo! Viu! Medo!”. A lista completa das sentenças pode ser consultada no apêndice E.

Os *onsets* e as durações das palavras que compõem o teste foram analisadas (ver apêndice F), a partir do aplicativo *Praat* (versão 6.3). Esses parâmetros são interessantes para saber em qual momento exato do experimento o participante ouviu cada uma das palavras-alvo.

5.2.2 PARTICIPANTES

Participaram 88 adultos, brasileiros natos e residentes na região Sudeste do Brasil, entre 18 e 64 anos de idade. Destes, 35 participantes foram excluídos da amostra final pelas

seguintes razões: erro técnico na plataforma (4); distração (22); dificuldade na análise dos dados por causa de uso de óculos e/ou cílios postigos pelo participante (3); não ter ao menos dois *trials* válidos em cada condição - concreta ou abstrata - (6). A amostra final foi, portanto, composta por 53 indivíduos, com idades entre 18 e 64 anos ($M = 412,71$ meses; $DP = 120,33$), dos quais 39 eram mulheres e 14 homens.

Os participantes preencheram um formulário elaborado com base na Pesquisa Nacional por Amostra de Domicílios — PNAD (IBGE, 2024). Esse questionário foi adaptado para coletar, além de dados socioeconômicos, informações sobre aspectos culturais e fatores relacionados ao estilo de vida, como a exposição à leitura e à prática de atividades físicas. Essas variáveis têm sido apontadas como preditores importantes para o desenvolvimento neurocognitivo pleno (Hoff, 2003; Hutton *et al.*, 2020; Jirout *et al.*, 2019; Piccolo *et al.*, 2016a; Sarsour *et al.*, 2011; Singh *et al.*, 2023; Ursache *et al.*, 2016; Zeng *et al.*, 2017).

A respeito do nível de escolaridade desses 53 participantes, 24 possuíam pós-graduação completa, 16 graduação completa, 06 ensino superior incompleto, 01 ensino superior em andamento, 06 ensino médio completo. Sobre falar línguas estrangeiras, 35 declararam falar um (ou mais) de um idioma. 35 pessoas declararam gostar de ler, 14 marcaram “mais ou menos” e 3 disseram que não gostavam. Sobre a quantidade de livros que a pessoa possuía em seu domicílio, 10 pessoas declararam que havia mais de 200, 8 tinham de 101 a 200 livros, 18 de 26 a 100, 10 pessoas de 11 a 25 livros e 7 pessoas de 1 a 10.

Os participantes foram distribuídos da seguinte forma entre uma das oito listas em que os estímulos foram randomizados, conforme se vê na tabela 3:

Tabela 3 — Distribuição dos participantes entre as listas do experimento adulto

Número da lista	Total de participantes
01	7
02	6
03	6
04	7
05	6
06	6
07	7
08	8

Fonte: Elaboração própria (2025).

A seguir, será apresentado o objetivo deste experimento controle.

OBJETIVOS

Este teste controle objetivou avaliar **(i)** as diferenças no acesso e no processamento de itens lexicais concretos e abstratos por adultos, falantes nativos do PB; **(ii)** se há maior interesse pelas palavras concretas em relação às palavras abstratas; e **(iii)** contrastar dados do acesso e do processamento desses itens lexicais com o de crianças entre 24 e 30 meses, adquirindo o PB.

HIPÓTESES

Estudos com eletroencefalografia que investigaram o processamento de verbos concretos e abstratos (Huang et. al, 2010; Kounios; Holcomb, 1994; West; Holcomb, 2000), por adultos, apontam que palavras concretas tendem a evocar um N400 mais acentuado do que palavras abstratas. Essa diferença sugere que os mecanismos subjacentes ao acesso lexical dos dois tipos de palavras não são os mesmos, e que palavras abstratas podem ser mais difíceis de processar e interpretar do que palavras concretas.

Diante disso, as hipóteses deste experimento são: **(i)** adultos brasileiros serão capazes de interpretar palavras concretas e abstratas; **(ii)** haverá diferenças no acesso lexical entre nomes concretos e abstratos; **(iii)** haverá maior interesse atencional direcionado aos itens lexicais concretos em relação aos abstratos.

PREDIÇÃO

Os adultos, falantes nativos do PB, apresentarão maior proporção de fixações para as palavras concretas em comparação às abstratas.

MÉTODO

Para realizar este experimento, foi adotado o Paradigma Experimental de Fixação Preferencial do Olhar que mede, em tempo real, os padrões de olhar dos participantes em

resposta à fala. Essa abordagem, que utiliza o rastreamento ocular, é uma versão simplificada do Paradigma do Mundo Visual (Tanenhaus *et al.*, 1995), metodologia consagrada para monitorar como o processamento da linguagem interage com a percepção visual. É um procedimento simples, frequentemente usado com crianças pequenas, no qual ocorre a apresentação de duas imagens e de uma instrução auditiva, e cujos padrões visuais dos participantes são analisados para avaliar a compreensão da linguagem.

A respeito deste paradigma experimental, Berends *et al.* (2015, p. 3, tradução da autora) explicam que:

Os participantes escutam estímulos linguísticos e, simultaneamente, observam uma cena contendo vários objetos, enquanto seus movimentos oculares são monitorados. Os estímulos linguísticos geralmente estão relacionados a um ou mais objetos na cena, e a questão investigada é se, e quando, as pessoas olham para esses objetos. Quando se considera o tempo necessário para iniciar uma sacada (um movimento ocular para outra localização), o momento em que o olhar do ouvinte é direcionado para um objeto nomeado oferece uma estimativa precisa do instante em que a palavra foi reconhecida (Allopenna *et al.*, 1998). Manipular a relação entre os objetos e o *input* linguístico (por exemplo, tornando-os mais difíceis de distinguir devido a sons semelhantes na fala) permite aos pesquisadores testar teorias sobre como os ouvintes acessam informações em seus léxicos mentais. Essas informações podem ser inferidas pelos padrões de fixação ocular dos ouvintes ao longo do tempo.

Ademais, Reuter e Lew-Williams (2023, p. 2, tradução da autora) complementam que:

Em paradigmas de rastreamento ocular, os olhares direcionados a um referente visual são geralmente considerados como evidência da ativação de representações semânticas relevantes (*apud* Cooper, 1974; Fernald *et al.*, 2008; Tanenhaus *et al.*, 1995), e os movimentos oculares antecipatórios têm sido igualmente interpretados como evidência da pré-ativação de representações lexicais (*apud* Altmann & Kamide, 1999; Kamide, 2008).

Por essas razões, ao adotar o Paradigma do Mundo Visual, será possível obter medidas *online* referentes à movimentação ocular dos participantes para as imagens, enquanto eles escutam o estímulo sonoro, com o objetivo de avaliar se já identificam itens lexicais concretos e abstratos do PB. Esse paradigma se justifica como um procedimento metodológico adequado para a investigação aqui proposta, uma vez que, conforme explicam Huettig *et al.* (2011, p. 156, tradução da autora):

Os movimentos oculares mediados pela linguagem não refletem apenas o processamento linguístico ou apenas o processamento da cena visual, mas sim representações mentais continuamente atualizadas com base nas informações derivadas tanto do *input* linguístico quanto do *input* visual. O Paradigma do Mundo Visual tem se mostrado uma ferramenta muito útil para investigar o processamento cognitivo: as análises sintáticas em andamento dos participantes, suas previsões e representações de eventos podem ser inferidas a partir da direção de seus olhares.

Além desses fatores, este método permite sincronizar os padrões de olhar dos participantes com a escuta, sendo esses padrões codificados quadro a quadro (*frame by frame*). As latências de resposta de cada experimento podem ser medidas com precisão de milissegundos em múltiplos testes sobre múltiplos itens, com base em dados de milhares de quadros por experimento.

Nesta tarefa de Fixação Preferencial do Olhar, duas imagens são apresentadas lado a lado, e o participante ouve uma frase que designa uma delas (à esquerda ou à direita). A relevância da utilização dessa técnica, na investigação de fenômenos psicolinguísticos, está em avaliar as habilidades perceptuais e o processamento cognitivo no exato momento da exposição aos estímulos.

MATERIAIS

Para as palavras-alvo, foram selecionadas ilustrações do site *StorySet*, que disponibiliza imagens sem custo pelo link: [Free Student Illustrations to design | Storyset](#). Como, por vezes, não havia ilustração perfeita do que se procurava, algumas figuras usadas no teste foram adaptadas para ilustrar melhor o que a palavra deveria designar. Todas as imagens foram padronizadas no tamanho de 8,7 cm de comprimento e 8,47 cm de altura, e foram mantidas em tons pastéis (ver apêndice C).

Os estímulos auditivos foram gravados no *BabyLab*, em cabine acusticamente isolada, do *Laboratoire de Psychologie du Développement et de l'Éducation de l'enfant*, da Université Paris Cité, utilizando um gravador TASCAM do modelo Dr-70d. Todos os áudios que compõem o experimento foram tratados no programa *Adobe Audition* para diminuir ruídos e também para padronizar a duração deles.

Os *trials* de treinamento, i.e., aqueles que antecedem os experimentais, possuem uma duração padrão de 6.5 segundos. Todos os estímulos sonoros têm uma pausa até o início do estímulo como em “Olha!” aos 200 ms. Após a palavra, há uma pausa de 500 ms para iniciar o segundo áudio, por exemplo “Uau” e, no final, há o mesmo período de silêncio por 100 ms. Os *trials* experimentais, por sua vez, possuem uma duração de 6 segundos, com uma pausa até o *onset* do estímulo como em “Olha! Casaco! Viu! Casaco!” aos 200 ms e os mesmos 100 ms de silêncio no final.

Este experimento controle foi realizado no computador da pesquisadora — um MacBook Air (Modelo M1, 2020), tela de 13,3 polegadas (2560×1600). Ele foi feito na casa da própria pesquisadora ou na dos participantes, desde que houvesse um ambiente silencioso e sem distratores, como televisões ligadas, som alto ou outros estímulos visuais e/ou sonoros que pudessem ser competidores de atenção para o participante.

PROCEDIMENTOS

O experimento *Como adultos brasileiros processam palavras concretas e abstratas?* foi, conforme mencionado anteriormente, realizado no *Lookit* e tinha a identificação do *Laboratoire de Psychologie du Développement et de l'Éducation de l'enfant — LaPsyDé*, da Université Paris Cité.

A razão pela qual o *LaPsyDé* foi o laboratório vinculado à plataforma experimental não se restringe apenas ao fato de a pesquisadora ter realizado o seu período sanduíche lá, mas também ao fato de que, antes de iniciar o desenvolvimento dos testes, é necessário que a universidade interessada em utilizar a plataforma firme um acordo³³ com o MIT, instituição responsável pelo *Lookit*. Além disso, é preciso submeter previamente a pesquisa a ser desenvolvida na plataforma ao Comitê de Ética³⁴ da própria instituição interessada, a fim de comprovar ao MIT que o estudo foi devidamente aprovado para execução. Essas etapas foram realizadas logo após o início do meu doutorado sanduíche para viabilizar que o experimento começasse a ser montado durante o meu período no exterior.

Como forma de democratizar e promover a ciência, após o meu retorno ao Brasil, iniciamos o processo de filiação do Laboratório de Acesso Sintático — ACESIN LAB e da Universidade Federal do Rio de Janeiro ao *Lookit*. No entanto, há implicações jurídicas envolvidas nesse acordo, especificamente em relação à indenização de participantes em casos de possíveis danos materiais ou de outra natureza pela universidade interessada, que estão sendo (re)avaliadas pela UFRJ e redigidas para a apreciação do departamento jurídico do MIT.

³³Para mais informações a respeito da filiação de uma universidade ao *Lookit*, acesse a documentação da plataforma: <https://lookit.readthedocs.io/en/master/>.

³⁴O parecer favorável do *Comité d'Ethique de la Recherche*, da Universidade de Paris, sob o número IRB: 00012024-62, pode ser consultado no apêndice G. Além da aprovação do Comitê de Ética da Universidade de Paris, esta pesquisa já havia sido aprovada pelo Comitê de Ética em Pesquisa (CEP-IESC-UFRJ), da UFRJ, sob o registro (CAAE: 66650823.4.0000.5286, parecer nº.6.080.129).

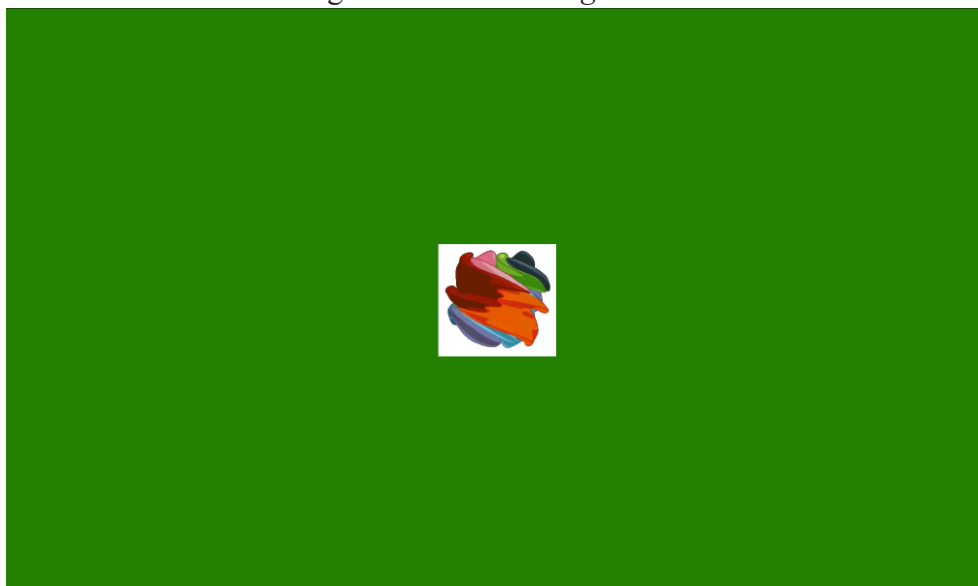
Em uma primeira tentativa de formalização do acordo, a direção da Faculdade de Letras informou que, em conformidade com a Legislação brasileira, não é permitido que a UFRJ se responsabilize por quaisquer questões que possam surgir. Por isso, a filiação do ACESIN lab da UFRJ ainda está em curso.

Diante dessas razões, o experimento foi, portanto, vinculado ao Laboratoire de Psychologie du Développement et de l'Éducation de l'enfant — LaPsyDé, da Université Paris Cité. Assim, a página de apresentação do experimento *Como adultos brasileiros processam palavras concretas e abstratas?* continha a identificação do LaPsyDé, a apresentação dos objetivos da pesquisa e informações básicas, como o caráter voluntário da participação, a duração do experimento e o e-mail da pesquisadora responsável.

Antes de iniciar o experimento, explicava ao participante que a atividade consistia em olhar duas imagens na tela do computador, uma à direita e outra à esquerda, e que ele ouviria uma palavra que designaria uma das imagens, estabelecendo uma relação mais concreta, como *gatinho* ou mais abstrata como *fofo*. Também dizia que a direção do seu olhar seria observada com o objetivo de avaliar para qual imagem ele observou, após a escuta do áudio. Posteriormente, fornecia o Termo de Consentimento Livre e Esclarecido para o participante, perguntava se havia alguma dúvida, solicitava que o participante colocasse o telefone no modo silencioso e o mantivesse guardado. Por fim, dizia que se retiraria da sala e o orientava a clicar no botão para iniciar o teste quando se sentisse à vontade.

Ao iniciar o teste, o participante via orientações para olhar atentamente para a tela, bem como os objetivos da pesquisa, e, em seguida, realizava um pequeno treinamento. Nesse momento, iniciavam-se as etapas experimentais e era apresentado um vídeo com um círculo colorido, acompanhado de um estímulo sonoro, o barulho de um sino. O círculo aparecia primeiro no centro da tela, depois à esquerda, à direita e, por fim, retornava ao centro. Em seguida, a mesma imagem surgia novamente no centro, e o fundo da tela ficava verde³⁵. Essa imagem (figura 1) era utilizada entre os *trials*, como um *attention grabber*, pelo período de um segundo antes do *trial* de familiarização ou dos experimentais começarem.

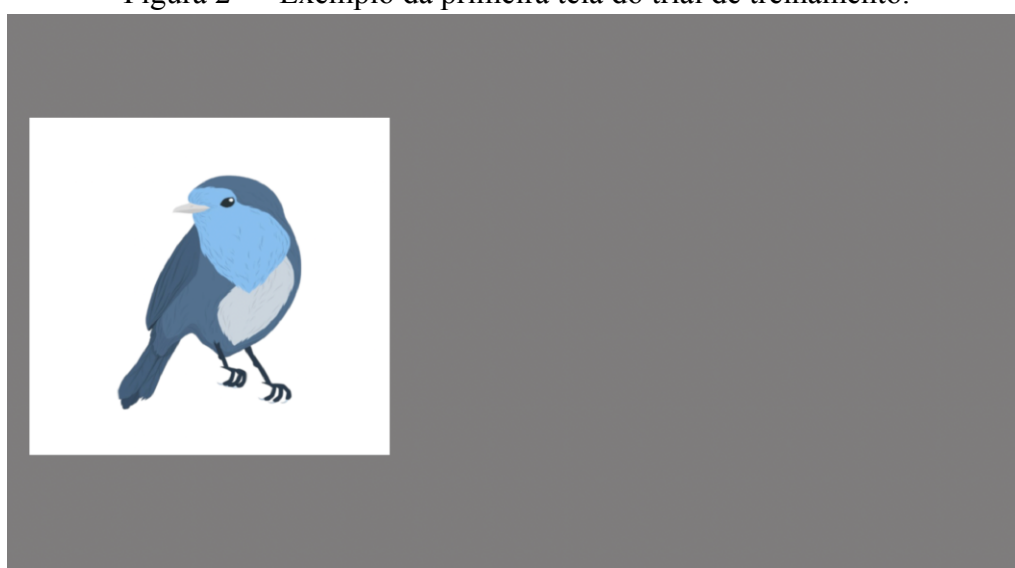
³⁵As gravações de todos os *trials*, fossem de familiarização ou os experimentais, eram feitas unitariamente sempre no início de um *trial* para outro. Portanto, neste momento em que a tela ficava verde, a gravação de cada *trial* era iniciada e, posteriormente, gravada no *Lookit*. Depois que a gravação iniciava ou terminava, aparecia a imagem que funcionava como *attention grabber* por um segundo.

Figura 1 — *Attention grabber*

Fonte: elaboração própria (2024).

Para permitir que o participante tenha tempo suficiente para inspecionar as imagens individualmente, cada imagem é primeiro apresentada sozinha por 3 segundos à esquerda e, após uma pausa de 500 ms com a tela cinza, a outra imagem aparece no lado direito da tela por 3s e um áudio neutro é reproduzido para cada uma dessas imagens (por exemplo: “Uau!” “Veja só!”; “Ih! Que legal”; “Nossa!” “Olha lá!”). Esses *trials* de treinamento possuem uma duração de 6500ms. A apresentação ocorre assim para cada um dos oito trials que compõem o teste, respeitando, logicamente, as imagens de cada *trial*. Um exemplo da ordem das telas está abaixo:

Figura 2 — Exemplo da primeira tela do trial de treinamento.



Fonte: elaboração própria (2024).

Depois de 3 segundos, a tela fica totalmente cinza por 500 ms. Conforme a imagem abaixo:

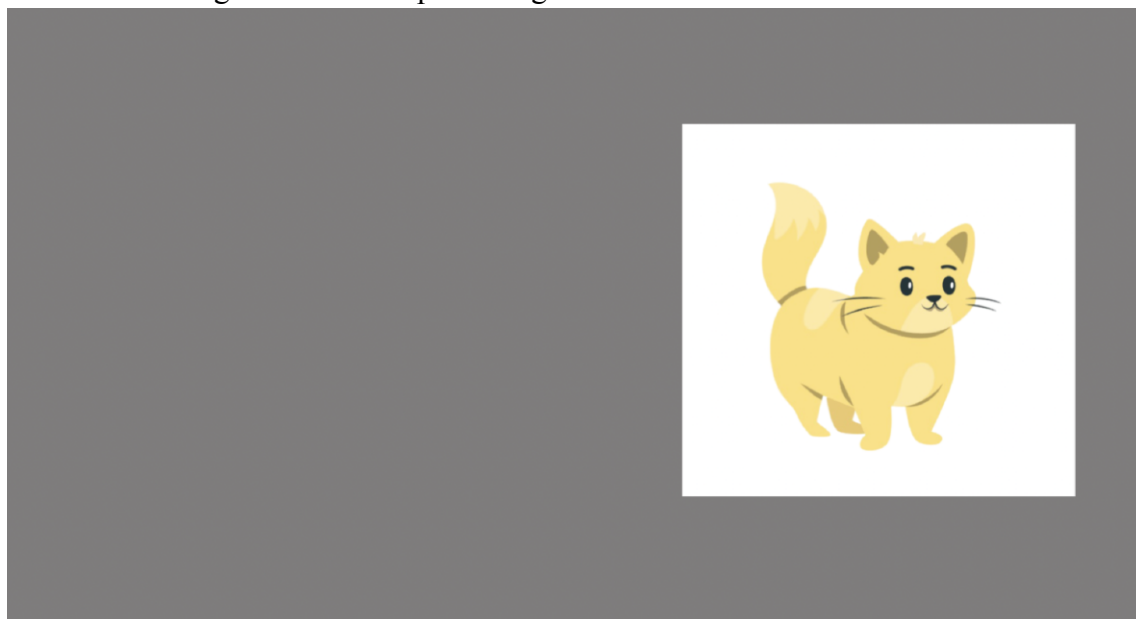
Figura 3 — Exemplo da tela intermediária dos trials de treinamento



Fonte: elaboração própria (2024).

Após este intervalo de 500 ms, surge a imagem da direita por 3 segundos. Como se vê abaixo:

Figura 4 — Exemplo da segunda tela do *trial* de treinamento

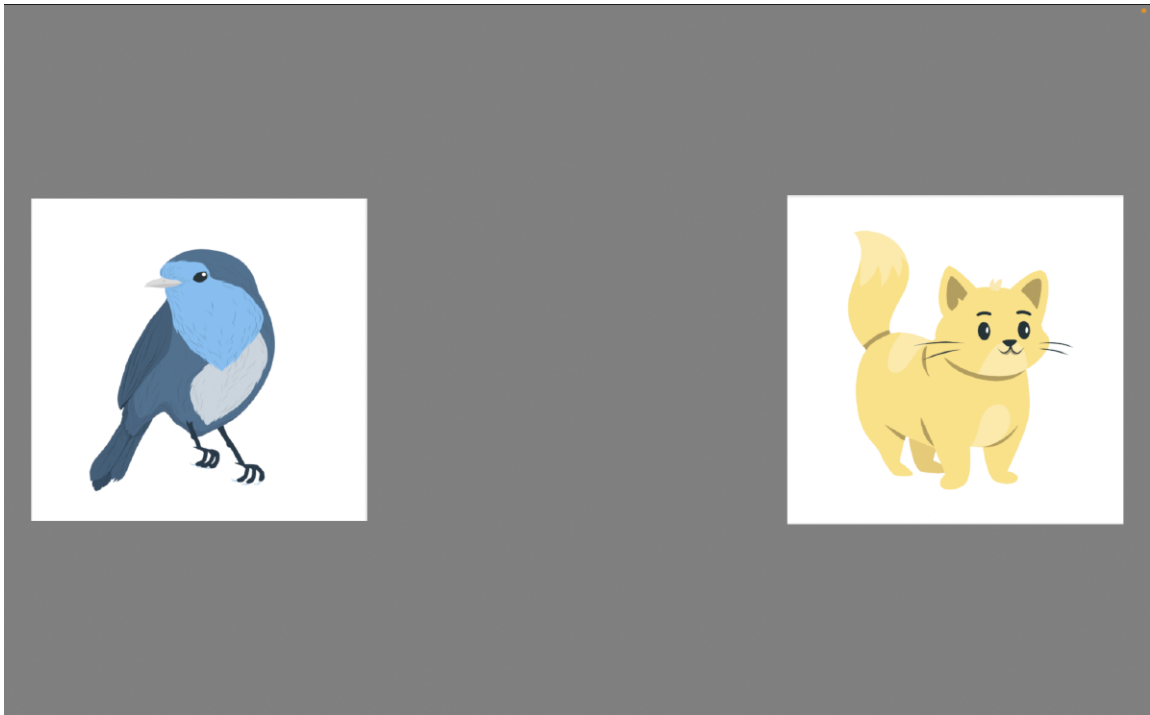


Fonte: elaboração própria (2024).

Esta imagem aparece por 3 segundos até que o primeiro trial de familiarização apareça na tela.

Nos dois *trials* de familiarização (Figura 5), o participante vê duas imagens pareadas (à esquerda ou à direita) e escuta o estímulo auditivo: “Olha! Passarinho! Viu Passarinho!” ou “Olha! Gatinho! Viu! Gatinho!” a depender da randomização. O objetivo dos *trials* de prática é apenas habituá-lo ao tipo de tarefa do experimento.

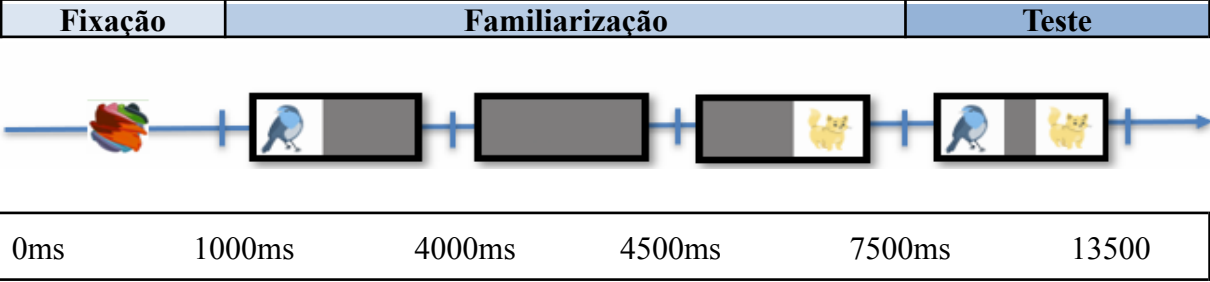
Figura 5 — Trial de familiarização



Fonte: elaboração própria (2024).

Para o segundo par de imagens que compõem a familiarização, conforme mencionado anteriormente, cada imagem era pré-apresentada, respeitando o tempo de 3 segundos para a imagem da esquerda, depois os 500 ms para a tela cinza, e mais 3 segundos para a imagem da direita. A cronologia do teste está representada abaixo na figura 6:

Figura 6 — Cronologia do experimento



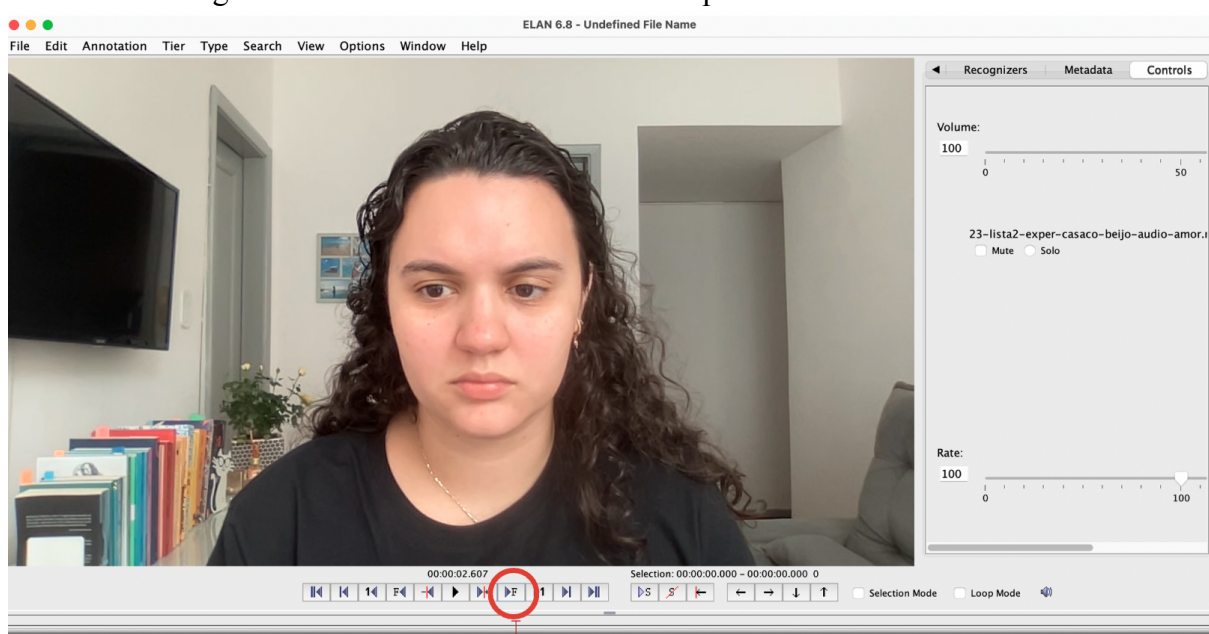
Fonte: elaboração própria (2025).

Após a etapa de familiarização, iniciam-se os *trials* experimentais. Os itens-alvo foram distribuídos em oito listas, conforme dito anteriormente (ver apêndice D), de modo que tanto a ordem dos *trials* quanto a posição das imagens (esquerda/direita) fossem contrabalançadas entre os participantes.

Cada *trial* experimental tem 6 segundos de duração. Entre essa pré-apresentação e o *trial* experimental, surge o vídeo do *attention grabber*, seguido sempre pelo *trial* de treinamento, i.e., aquela sequência em que as imagens aparecem separadamente, conforme foi apresentado acima para os *trials* de familiarização.

Para a análise dos dados, utilizamos o software ELAN — *Eudico Linguistic Annotator* — versão 6.8 — (Figura 7), criado pelo *Max Planck Institute for Psycholinguistics*, em Nijmegen, na Holanda. O programa, desenvolvido para fazer anotações manuais de arquivos de mídia, destina-se principalmente para análises linguísticas. A partir de um controle de mídia no estilo VCR³⁶, torna-se possível avançar quadro a quadro (*frame by frame*), conforme o botão destacado com o círculo vermelho, em cada vídeo, para analisar a direção do olhar do participante.

Figura 7 — Tela Inicial do ELAN — Experimento controle adulto



³⁶VCR refere-se a um conjunto de controles de mídia que imitam os botões e funcionalidades dos antigos gravadores de videocassete (VCRs). Esses controles geralmente incluem funções para iniciar, pausar, retroceder e avançar a reprodução do vídeo, permitindo, por exemplo, o avanço *frame a frame*. Essa interface é útil para análises detalhadas, pois possibilita ao usuário inspecionar cuidadosamente cada instante da gravação, algo fundamental em estudos que exigem precisão temporal.

Fonte: elaboração própria (2025)

Neste experimento, foi adotado um *time point* de 33 ms por *frame* (quadro) para a análise do rastreamento ocular, o qual foi computado em três níveis: *left*, *right* ou *look away*, em planilhas do Excel³⁷. Visto que cada *trial* experimental tinha 6 segundos de duração, havia 183 linhas para cada um deles, totalizando 1.098 linhas por participante.

Como mencionado na seção 5.2, um aspecto negativo em relação ao uso do *Lookit* é que essa análise do rastreamento ocular é feita manualmente, o que a torna substancialmente mais trabalhosa em comparação aos rastreadores convencionais. Porém, como também explicado na seção 5.2, isso não configura um problema em relação à confiabilidade dos dados, desde que sejam adotados alguns controles metodológicos.

Por exemplo, em tarefas de acesso lexical com o Paradigma do Mundo Visual, que não demandam alta resolução espacial (Erel et al., 2022), como os nossos experimentos, vídeos coletados remotamente, como pelo *Lookit*, podem ser anotados manualmente no ELAN com desempenho adequado (Kaduk et al., 2023; Scott et al., 2017; Scott; Schulz, 2017b; Slim et al., 2024).

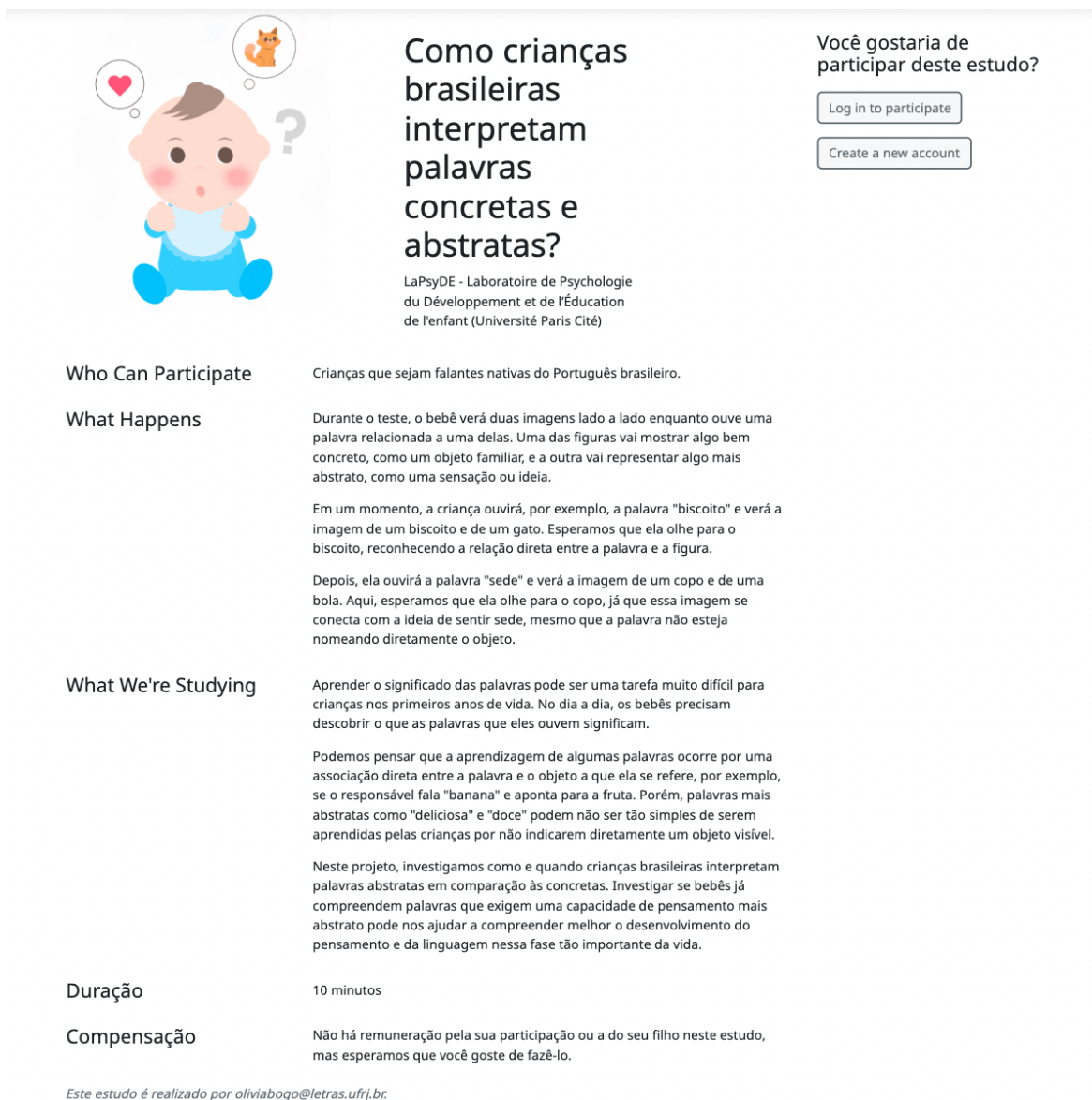
Para tanto, é preciso que a taxa de análise do frame seja ≥ 30 fps (≈ 33 ms por quadro) e haja sincronização temporal confiável com os eventos do experimento para todos os participantes. Seguindo esses critérios, a precisão temporal é suficiente para identificar o momento em que o participante direciona o olhar para um dos estímulos exibidos na tela.

³⁷ Para compreender melhor como o resultado de um participante foi codificado, basta consultar a planilha: <https://docs.google.com/spreadsheets/d/1ajtFJjM1eJNpwZLWrgz7-RpqbQkv5lpMP3r9xs7yFNU/edit?usp=sharing>.

5.3 COMO CRIANÇAS BRASILEIRAS INTERPRETAM PALAVRAS CONCRETAS E ABSTRATAS DO PORTUGUÊS BRASILEIRO?

O experimento *Como crianças brasileiras interpretam palavras concretas e abstratas?* também foi elaborado no *Lookit*. A tela de apresentação no *Lookit* era assim:

Figura 8 — Tela de apresentação do experimento infantil no *Lookit*



Como crianças brasileiras interpretam palavras concretas e abstratas?

LaPsyDE - Laboratoire de Psychologie du Développement et de l'Éducation de l'enfant (Université Paris Cité)

Você gostaria de participar deste estudo?

[Log in to participate](#)

[Create a new account](#)

Who Can Participate

Crianças que sejam falantes nativas do Português brasileiro.

What Happens

Durante o teste, o bebê verá duas imagens lado a lado enquanto ouve uma palavra relacionada a uma delas. Uma das figuras vai mostrar algo bem concreto, como um objeto familiar, e a outra vai representar algo mais abstrato, como uma sensação ou ideia.

Em um momento, a criança ouvirá, por exemplo, a palavra "biscoito" e verá a imagem de um biscoito e de um gato. Esperamos que ela olhe para o biscoito, reconhecendo a relação direta entre a palavra e a figura.

Depois, ela ouvirá a palavra "sede" e verá a imagem de um copo e de uma bola. Aqui, esperamos que ela olhe para o copo, já que essa imagem se conecta com a ideia de sentir sede, mesmo que a palavra não esteja nomeando diretamente o objeto.

What We're Studying

Aprender o significado das palavras pode ser uma tarefa muito difícil para crianças nos primeiros anos de vida. No dia a dia, os bebês precisam descobrir o que as palavras que eles ouvem significam.

Podemos pensar que a aprendizagem de algumas palavras ocorre por uma associação direta entre a palavra e o objeto a que ela se refere, por exemplo, se o responsável fala "banana" e aponta para a fruta. Porém, palavras mais abstratas como "deliciosa" e "doce" podem não ser tão simples de serem aprendidas pelas crianças por não indicarem diretamente um objeto visível.

Neste projeto, investigamos como e quando crianças brasileiras interpretam palavras abstratas em comparação às concretas. Investigar se bebês já compreendem palavras que exigem uma capacidade de pensamento mais abstrato pode nos ajudar a compreender melhor o desenvolvimento do pensamento e da linguagem nessa fase tão importante da vida.

Duração

10 minutos

Compensação

Não há remuneração pela sua participação ou a do seu filho neste estudo, mas esperamos que você goste de fazê-lo.

Este estudo é realizado por oliviabogo@letras.ufrj.br.



This material is based upon work supported by the National Science Foundation (NSF) under Grants 1429216, 1823919, and 2209756; the Center for Brains, Minds and Machines (CBMM), funded by NSF STC award CCF-1231216, and by an NSF Graduate Research Fellowship under Grant No. 1122374. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

[Privacidade](#)
[Entre em contato](#)
 Conecte-se: [f](#)

Fonte: Elaboração própria (2025).

5.3.1 ESTÍMULOS

Os estímulos do teste *Como crianças brasileiras interpretam palavras concretas e abstratas?* são idênticos ao do teste adulto descrito na seção 5.2.1.

5.3.2 PARTICIPANTES

Foram testadas 82 crianças, brasileiras nativas e residentes no Brasil, entre 24 e 30 meses de vida. Destas, 30 foram excluídas da amostra pelas seguintes razões: erro técnico na plataforma (5); coleta de dados realizada de forma inadequada pelo responsável, com o uso do celular ou sem filmar o rosto da criança (2); distração (5); interação de um adulto durante o teste (1); choro (1) e por não ter, ao menos, dois *trials* válidos em cada condição - concreta ou abstrata (16). A amostra final foi, então, composta por 52 crianças, entre 24 e 30 meses de vida ($M = 27,11$ meses; $DP = 2,11$), sendo 26 meninas e 26 meninos, com distribuição perfeitamente balanceada quanto ao gênero.

Do total de testes realizados, 18 foram feitos na própria plataforma *Lookit*, ou seja, feitos pelos próprios responsáveis das crianças. Os 34 testes restantes, aplicados pela pesquisadora, ocorreram em creches públicas, sendo uma localizada na cidade do Rio de Janeiro e cinco na cidade de Petrópolis (RJ), mediante autorização prévia da instituição e o Termo de Consentimento Livre e Esclarecido — TCLE (ver apêndice H) formalmente realizada pelos responsáveis legais dos menores, seja por escrito (quando a pesquisadora se dirigia à creche) ou registrada por vídeo na própria plataforma do *Lookit* (quando os responsáveis realizavam o teste em suas casas).

Os participantes foram distribuídos da seguinte forma entre uma das oito listas em que os estímulos foram randomizados, conforme se verifica na tabela 4:

Tabela 4 — Distribuição dos participantes entre as listas do experimento infantil

Número da lista	Total de participantes
01	6
02	7
03	6
04	6
05	6
06	6
07	9
08	6

Fonte: Elaboração própria (2025).

A respeito do número de participantes, é relevante citar uma publicação na *Society for Research in Child Development*, em que Dixon (2014) mapeou os 20 artigos mais influentes na área das ciências cognitivas sobre o desenvolvimento infantil para discutir o tamanho das amostras. De acordo com a autora, estudos relevantes, que impactaram a área com resultados estatisticamente significativos, incluíam de 12 a 24 participantes por condição, o que é adequado para detectar efeitos reais (Baillargeon, 1987; Hamlin et al., 2007 *apud* Dixon, 2014; Saffran et al., 1996). Nesta mesma direção, Oakes (2017) argumenta que amostras de 20 ou 24 bebês por condição são suficientes para fornecer estudos com poder estatístico adequado e estimativas precisas dos efeitos investigados.

Assim, em ambos os experimentos executados nesta tese, sobretudo neste com bebês, cuja coleta de dados é metodologicamente mais desafiadora, seguimos as melhores práticas publicadas internacionalmente na área quanto ao número mínimo de participantes por condição. Foram, portanto, incluídos ao menos 24 indivíduos em cada uma das duas condições experimentais, distribuídos entre oito listas experimentais, em um delineamento *within-subject*, de modo a assegurar a confiabilidade e a validade estatística dos resultados obtidos nos experimentos. Isso será melhor detalhado na seção de discussão dos resultados.

A seguir, serão apresentados os objetivos do experimento infantil.

OBJETIVOS

Esta pesquisa busca evidências experimentais sobre as etapas iniciais da aquisição linguística a fim de investigar como crianças adquirem e acessam itens lexicais conceitualmente distintos. Mais especificamente, o presente trabalho tem como objetivos: **(i)** verificar se crianças brasileiras, entre 24 e 30 meses de idade, adquirindo o PB, conseguem interpretar palavras concretas e abstratas; **(ii)** apontar as diferenças cognitivas no acesso lexical de nomes concretos e abstratos por crianças brasileiras; **(iii)** aferir se há maior interesse atencional por itens lexicais concretos em comparação aos abstratos, em virtude da natureza conceitual desses itens; **(iv)** contribuir para os estudos em AL com inventários de palavras sobre a aquisição vocabular de crianças, adquirindo o PB; **(v)** fornecer dados inéditos sobre a arquitetura cognitiva, especialmente sobre o processamento lexical de nomes concretos e abstratos por crianças brasileiras entre 24 a 30 meses; **(vi)** explorar a possibilidade de utilizar este teste como uma ferramenta linguística preditiva para transtornos do neurodesenvolvimento que acometem o processamento linguístico, sobretudo o abstrato; **(vii)** promover e democratizar a pesquisa em neurodesenvolvimento cognitivo no Brasil por meio da divulgação e ampliação da plataforma *Lookit*.

A seguir, serão descritas as hipóteses deste trabalho.

HIPÓTESES

A aquisição lexical é, por natureza, um processo abstrato e simbólico, independentemente de um item lexical específico possuir ou não um referente concreto no mundo real. Evidências empíricas indicam que existem diferenças neurocognitivas entre esses dois tipos de conceitos, tanto em adultos (Huang et. al, 2010; Kounios; Holcomb, 1994; West; Holcomb 2000;) quanto em crianças (Bergelson; Swingley, 2012, 2013; Reuter; Lew-Williams, 2023; Tincoff; Jusczyk, 1999).

Os estudos sobre aquisição de linguagem, por sua vez, especialmente os resenhados no capítulo três desta tese, destacam a capacidade das crianças de segmentar o fluxo contínuo da fala e atribuir significado a esses elementos. Essa habilidade ocorre independentemente dos itens em questão se referirem a elementos funcionais, como artigos, preposições e marcadores morfossintáticos, ou a itens lexicais que designam objetos ou conceitos no mundo real.

Ainda que exista uma vasta literatura acerca da sensibilidade de crianças a itens funcionais e lexicais (Saffran et al., 1996; Shady, 1996; Shi et al., 1999; Shi; Werker, 2001), que contribui para a discussão sobre a capacidade delas de categorizar elementos e adquirir a língua materna à qual estão expostas, poucos estudos abordaram especificamente a aquisição de conceitos concretos e abstratos por bebês, e nenhum investigou crianças adquirindo o PB.

Diante disso, as hipóteses experimentais deste estudo são: **(i)** crianças brasileiras, a partir dos 24 meses de idade, já serão capazes de identificar itens lexicais concretos e abstratos; **(ii)** haverá diferenças nos padrões temporais de fixação ocular entre itens lexicais concretos e abstratos; **(iii)** haverá maior interesse atencional por itens lexicais concretos em comparação aos abstratos. Isso porque, na condição abstrata, há uma pior imageabilidade.

A seguir, será apresentada a predição do experimento psicolinguístico.

PREDIÇÃO

As crianças, falantes nativas do PB, apresentarão maior proporção de fixações para as palavras concretas em comparação às abstratas.

MÉTODO

O método utilizado para o experimento das crianças foi o mesmo adotado para o teste adulto. Destaca-se, aqui, como o paradigma experimental de Fixação Preferencial do Olhar, enquanto uma versão simplificada do Paradigma do Mundo Visual (Tanenhaus *et al.*, 1995), é um procedimento adequado e frequentemente usado para investigar as etapas iniciais da aquisição de linguagem.

MATERIAIS

Os materiais utilizados para o experimento das crianças foram os mesmos elaborados para o teste de adultos.

PROCEDIMENTOS

O experimento *Como crianças brasileiras interpretam palavras concretas e abstratas?* também está hospedado no [Children Helping Science](#). Um site criado e gerenciado pelo *Parent Researcher Collaborative* (MIT News, 2020; Association For Psychological Science, 2020) que atua conjuntamente com o *Lookit* (Scott et al., 2017; Scott; Schulz, 2017) como uma plataforma para desenvolver e publicar testes online em ciências cognitivas, especialmente voltados para crianças.

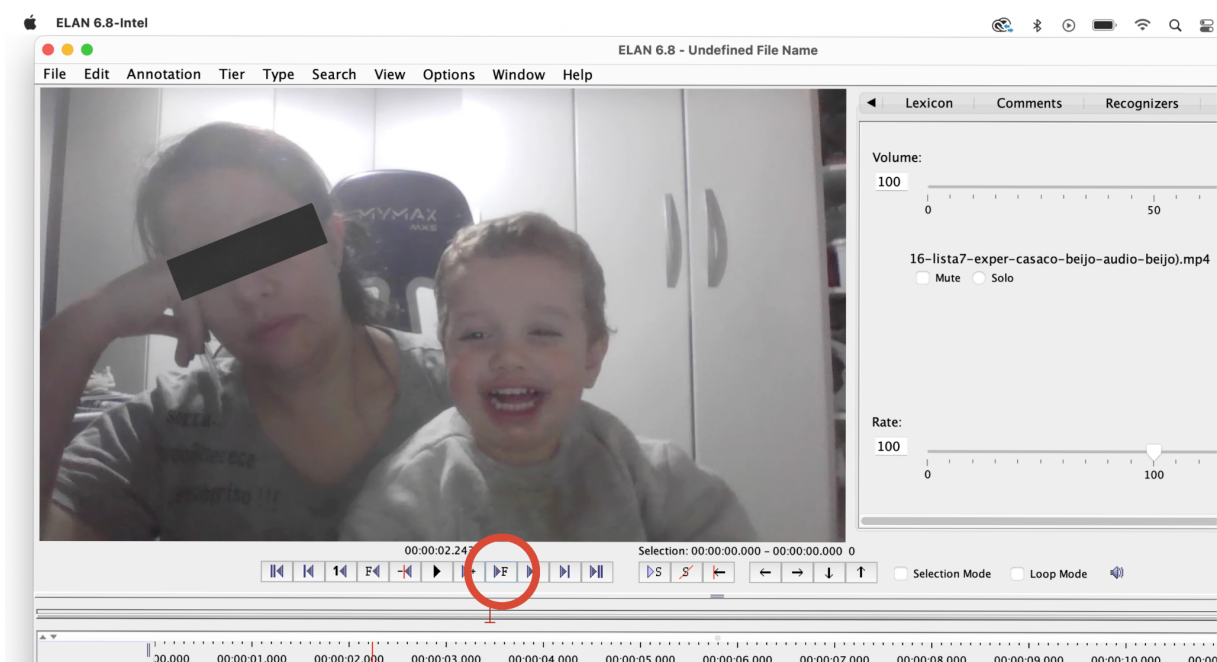
A página de apresentação do experimento *Como crianças brasileiras interpretam palavras concretas e abstratas?* tinha a identificação do *Laboratoire de Psychologie du Développement et de l'Éducation de l'enfant*, da Université Paris Cité, pelas razões já explicadas na parte de *procedimentos* do teste piloto, realizado com adultos. Além disso, havia a descrição da atividade, os objetivos do estudo, a informação de que a participação era voluntária e, por fim, a identificação da pesquisadora responsável pela atividade experimental na plataforma.

O responsável pelo bebê, ao iniciar o estudo, via a apresentação do experimento. Ao prosseguir, ele era direcionado para o Termo de Consentimento Livre e Esclarecido, no qual devia se gravar lendo as informações e manifestar seu consentimento para participação no experimento. Em seguida, ele era direcionado para a página de instruções de gravação do vídeo e era orientado a pegar o seu bebê no colo e a posicioná-lo de forma que, principalmente, o rosto da criança estivesse iluminado e centralizado na tela.

Também era pedido que, qualquer distração fosse retirada do ambiente, para que o bebê pudesse prestar atenção apenas no monitor em que a atividade experimental estivesse ocorrendo. Em seguida, aparecia uma mensagem, explicando que o bebê faria uma etapa de teste para se familiarizar com o tipo de tarefa. A partir desta etapa, todos os procedimentos seguiam o mesmo protocolo descrito para o experimento com adultos.

A análise dos dados do experimento das crianças também foi feita utilizando o software ELAN — *Eudico Linguistic Annotator* — versão 6.8 — (Figura 9), desenvolvido pelo *Max Planck Institute for Psycholinguistics*, em Nijmegen, na Holanda. Como explicado na parte de *Procedimentos* do teste com adultos, as anotações são feitas manualmente, a partir do avanço feito quadro a quadro (*frame by frame*), conforme indicado pelo botão destacado com o círculo vermelho, em cada vídeo, a fim de analisar a direção do olhar do participante.

Figura 9 — Tela Inicial do ELAN — Experimento infantil



Fonte: elaboração própria (2025).

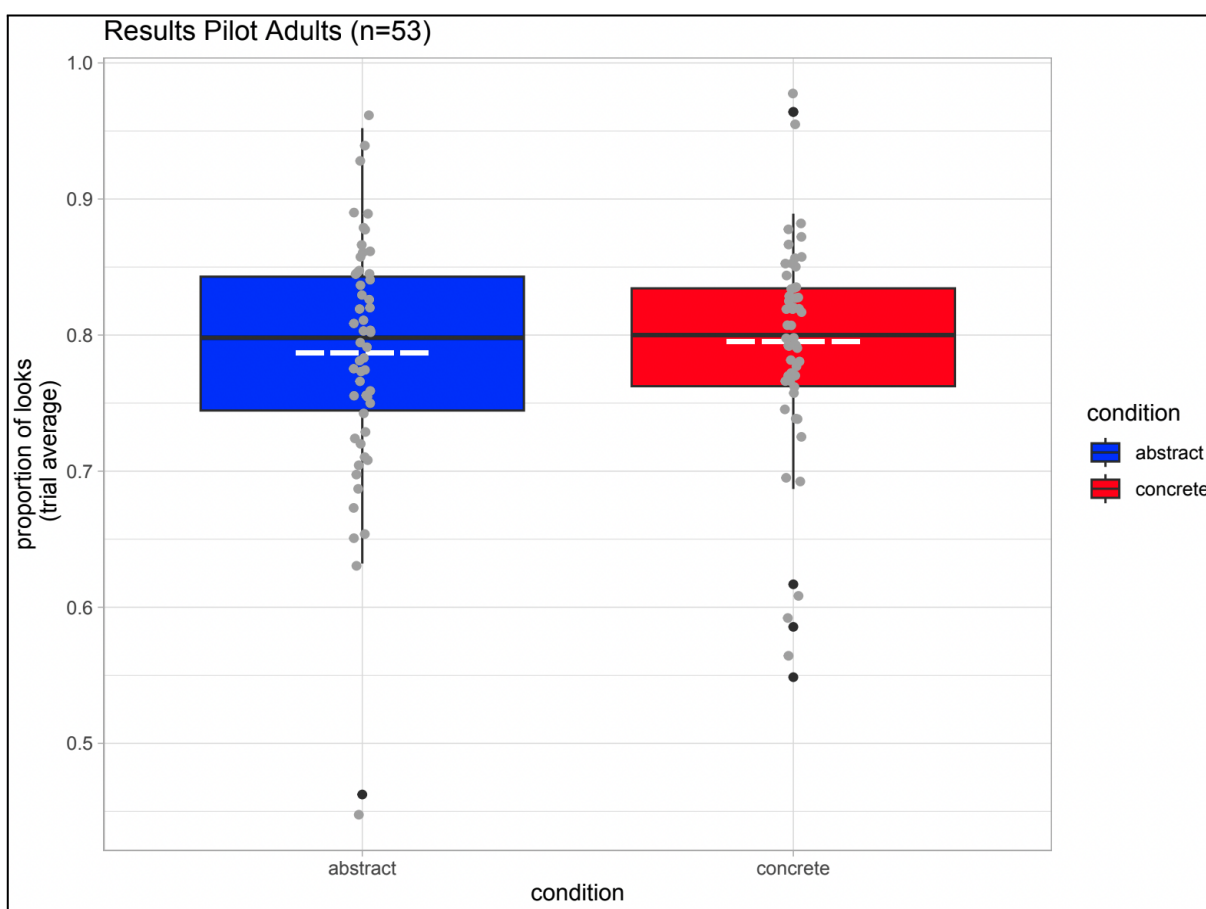
Neste experimento, foi adotado o mesmo *time point* de 33 ms por quadro (frame) para a análise do direcionamento do olhar da criança, que foi registrado em três níveis: *left*, *right* ou *look away*, e posteriormente computado em planilhas do Excel.

6 RESULTADOS

Para verificar possíveis diferenças no acesso lexical e no processamento de palavras concretas e abstratas, aplicamos o teste controle com 53 adultos, falantes nativos do PB. A proporção média de fixações foi analisada para comparar o desempenho global dos participantes em cada condição (concreta vs. abstrata) e o curso temporal da fixação visual ao longo do tempo. Esses dois tipos de análise permitem não só identificar distinções entre esses itens lexicais, como também observar quando essas diferenças emergem ao longo do processamento linguístico.

A proporção média de olhares para a imagem-alvo em cada condição (concreta vs. abstrata) dos 53 participantes do experimento está representada no *box-plot* (Gráfico 8) abaixo.

Gráfico 8 — Proporção média de olhares em cada condição (concreta e abstrata) — experimento adulto



Fonte: Elaboração própria (2025).

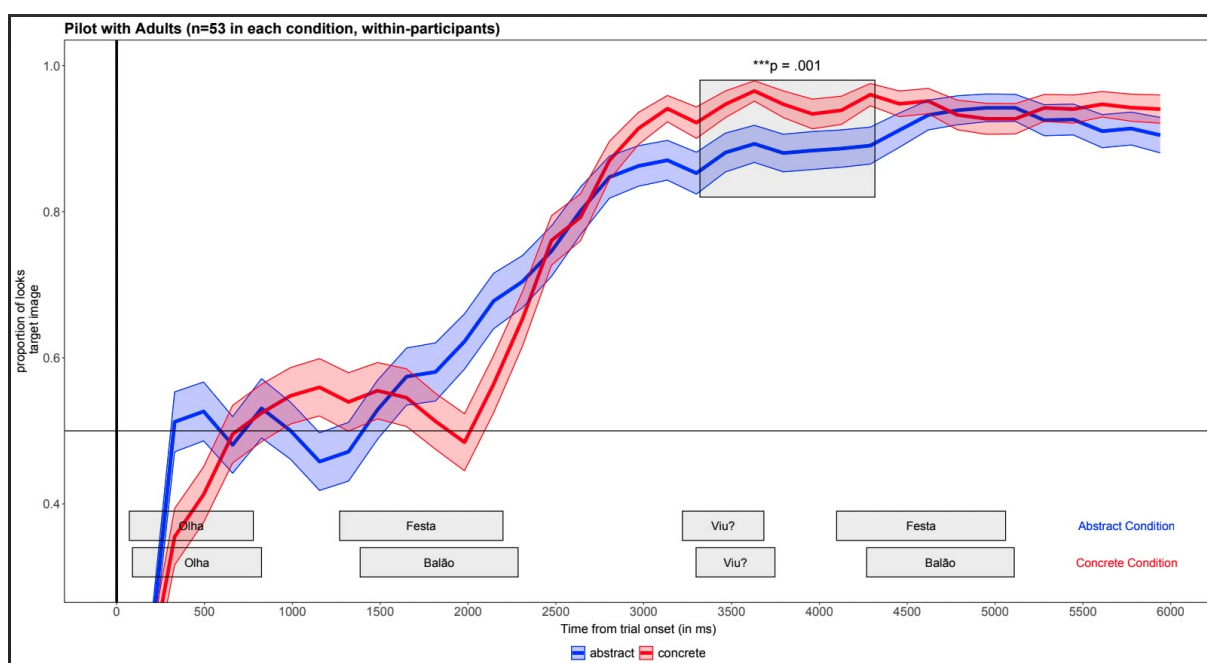
Esse *box-plot* indica a distribuição da média de olhar dos participantes para a imagem-alvo ao longo de todos os *trials* em cada condição. A análise indica que a distribuição das fixações é semelhante entre as condições concreta e abstrata, com ambas acima do nível da chance (50%), embora sem uma diferença evidente entre elas.

As médias das proporções de fixação, representadas pelas linhas brancas tracejadas dentro das caixas, indicam uma leve preferência no processamento de palavras concretas pelos participantes. A mediana, indicada pela linha preta, apresenta valores semelhantes para ambas as condições, mas com uma leve tendência à condição concreta. Além disso, o intervalo interquartil (IQR), que corresponde à distância entre o primeiro e o terceiro quartil e reflete a dispersão dos dados na região central da distribuição, aponta maior dispersão de *outliers* na condição abstrata.

Uma análise de variância (ANOVA) de medidas repetidas foi conduzida para avaliar o efeito condição experimental (concreto vs. abstrato) sobre a proporção média de olhares ao longo dos *trials*. Os resultados indicaram que não houve diferenças significativas entre as condições, $F(1, 52) = 0,35$, $p = 0.559$, $\eta^2 = 0,007$, o que significa o tipo de estímulo não influenciou a reação do participante.

Também foi conduzida uma análise da proporção do olhar ao longo dos *trials* experimentais (Gráfico 9). Essa análise possibilita acompanhar o processamento lexical ao longo do tempo de apresentação dos estímulos, conforme representado abaixo:

Gráfico 9 — Diferenças temporais nos padrões de fixação ocular entre a condição concreta e a abstrata (*Cluster Permutation*) — experimento adulto



Fonte: Elaboração própria (2025).

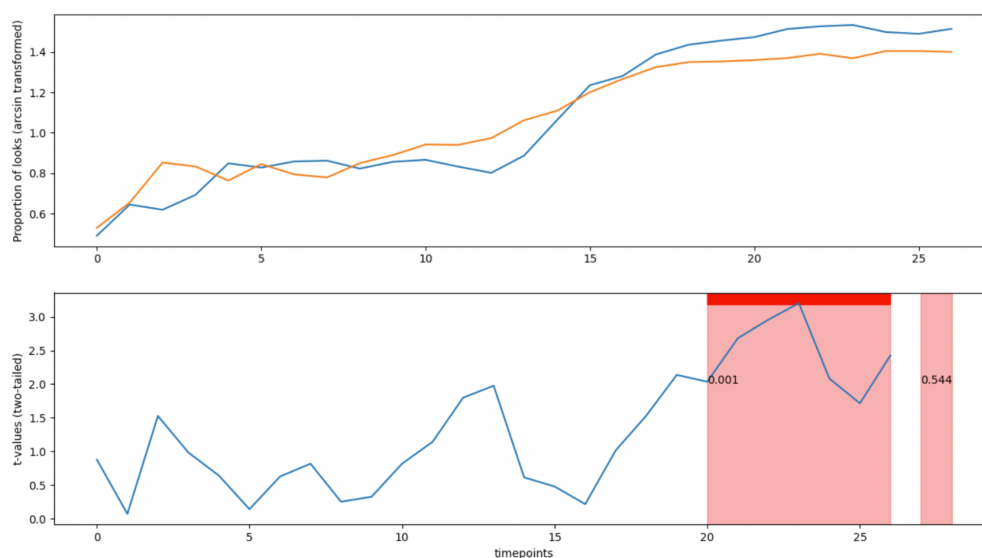
Inicialmente, observamos que a proporção de fixações na imagem-alvo está em torno de 50% para ambas as condições, sugerindo que os participantes ainda estão ouvindo e processando o estímulo auditivo, sem uma preferência clara.

Aos 2000 ms, após o início do *trial*, percebemos uma queda sutil na fixação das palavras concretas, o que pode indicar maior facilidade em reconhecê-las e, consequentemente, uma pequena perda de interesse na manutenção da atenção ao estímulo pelos participantes.

Em seguida, vemos um aumento gradual na proporção de fixações em ambas as condições, com uma leve vantagem para a condição concreta em relação à abstrata ao longo do tempo. Essa diferença visual se torna mais acentuada entre 3000 ms e 4000 ms, intervalo em que foi encontrado um *cluster* significativo.

Para avaliar a significância estatística das diferenças observadas entre as condições ao longo do tempo, foi aplicada uma análise do tipo *Cluster Permutation*³⁸, apresentada no gráfico 10 abaixo.

Gráfico 10 — Trajetória temporal da proporção de olhares por condição (concreta e abstrata) — experimento adulto



Fonte: Elaboração própria (2025).

A análise revelou um *Cluster* significativo ($p = 0.001$) no intervalo entre aproximadamente o 20º e o 25º ponto de tempo analisado, correspondente a cerca de 3000–4000 ms após o início do *trial*. Nesse intervalo, em consonância com os resultados previamente reportados no gráfico 9, a diferença observada indica uma preferência dos participantes pela condição concreta em relação à abstrata. Esse efeito, destacado em vermelho no gráfico, corresponde ao cluster significativo identificado pela análise de *Cluster Permutation*.

Essa diferença pode ser explicada devido à maior imageabilidade desses itens lexicais, dado que as palavras concretas são mais facilmente associadas a imagens mentais. Esse fator

³⁸A *cluster-based permutation analysis* é um teste estatístico não paramétrico utilizado para identificar efeitos significativos em dados com múltiplas dimensões. De acordo com Maris e Oostenveld (2007, p. 181, tradução da autora): “O teste de permutação baseado em *clusters* para análises com múltiplos sensores é muito semelhante ao utilizado para análises com um único sensor. Na verdade, o cálculo da estatística de teste difere em apenas um aspecto: em vez de agrupar as amostras de tempo selecionadas em conjuntos conectados com base na adjacência temporal, as amostras (sensor, tempo) são agrupadas com base na adjacência espacial e temporal.”

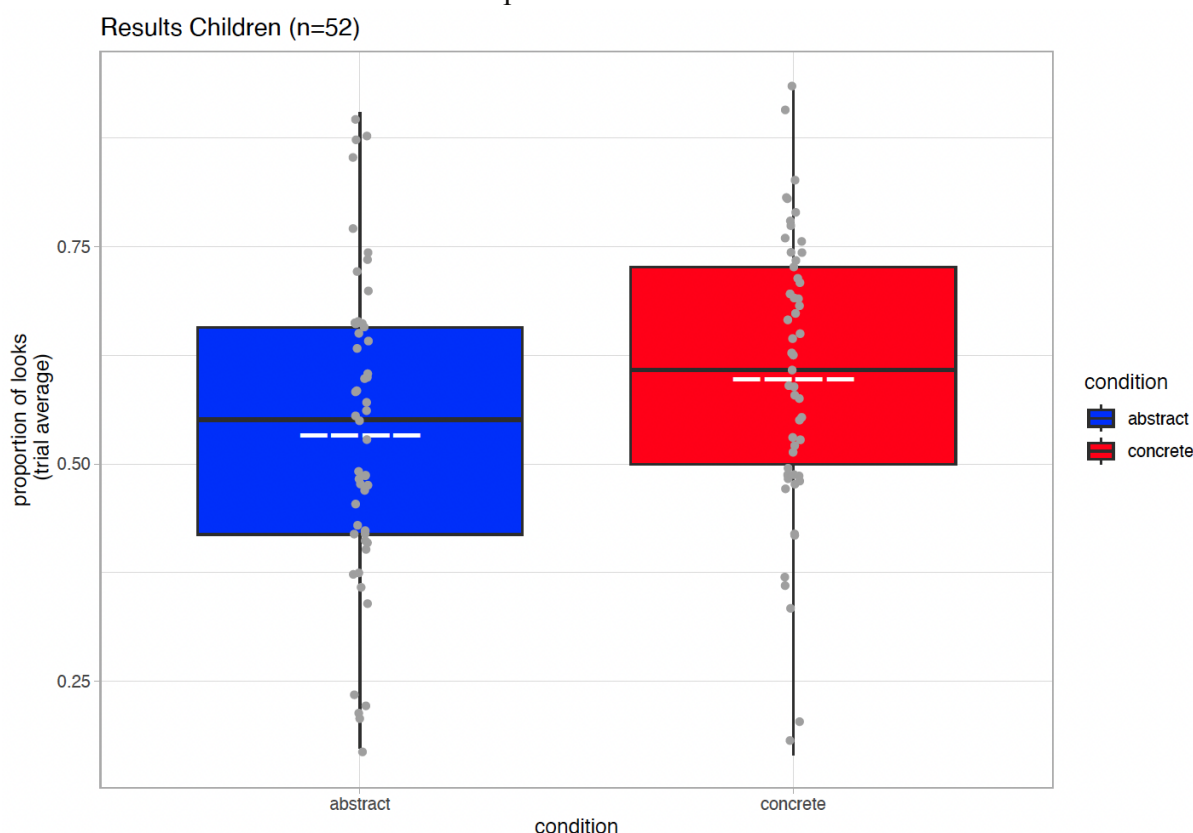
pode facilitar a ativação de representações visuais e semânticas, bem como a associação entre o estímulo auditivo e o visual correspondente, refletindo a vantagem da condição concreta em relação à abstrata, como apresentada no gráfico 10. Outro fator que pode influenciar esse resultado é a sinestesia, já que palavras concretas tendem a ativar, não só áreas já mapeadas para a linguagem, como múltiplas áreas sensoriais do cérebro, como as sensório-motoras e mnemônicas. Isso significa que palavras concretas podem evocar experiências multimodais, o que pode favorecer uma vantagem temporal na sua identificação.

Em relação ao teste infantil, tínhamos como objetivos com esta pesquisa: **(i)** verificar se crianças brasileiras, entre 24 e 30 meses de idade, adquirindo o PB, já interpretam palavras concretas e abstratas; **(ii)** indicar diferenças no acesso lexical de nomes concretos e abstratos por crianças brasileiras nas etapas iniciais da AL; **(iii)** aferir se há maior interesse atencional por itens lexicais concretos em comparação aos abstratos, em função da natureza conceitual desses itens; **(iv)** contribuir para os estudos em AL com inventários de palavras sobre a aquisição vocabular de crianças brasileiras; **(v)** publicar dados empíricos inéditos sobre a arquitetura cognitiva, especialmente sobre o processamento lexical de nomes concretos e abstratos por crianças entre 24 e 30 meses, adquirindo o PB.

Além disso, outros objetivos deste trabalho são: **(vi)** verificar a possibilidade de utilizar este teste como uma ferramenta linguística preditiva para transtornos do neurodesenvolvimento que acometem o processamento linguístico, principalmente o abstrato; e **(vii)** promover e democratizar a pesquisa em neurodesenvolvimento cognitivo no Brasil por meio da divulgação e da ampliação de outros grupos de pesquisa que possam usar o *Lookit* como plataforma experimental. Esses últimos dois pontos serão discutidos posteriormente, após a apresentação dos gráficos.

A fim de verificar se crianças brasileiras, entre 24 e 30 meses de idade, já interpretam itens lexicais concretos e abstratos, foi realizada a análise da proporção média de olhares para a imagem-alvo nas duas condições experimentais (concreta e abstrata), conforme apresentado no *boxplot* abaixo (Gráfico 11).

Gráfico 11 — Proporção média de olhares em cada condição (concreta e abstrata) — experimento infantil



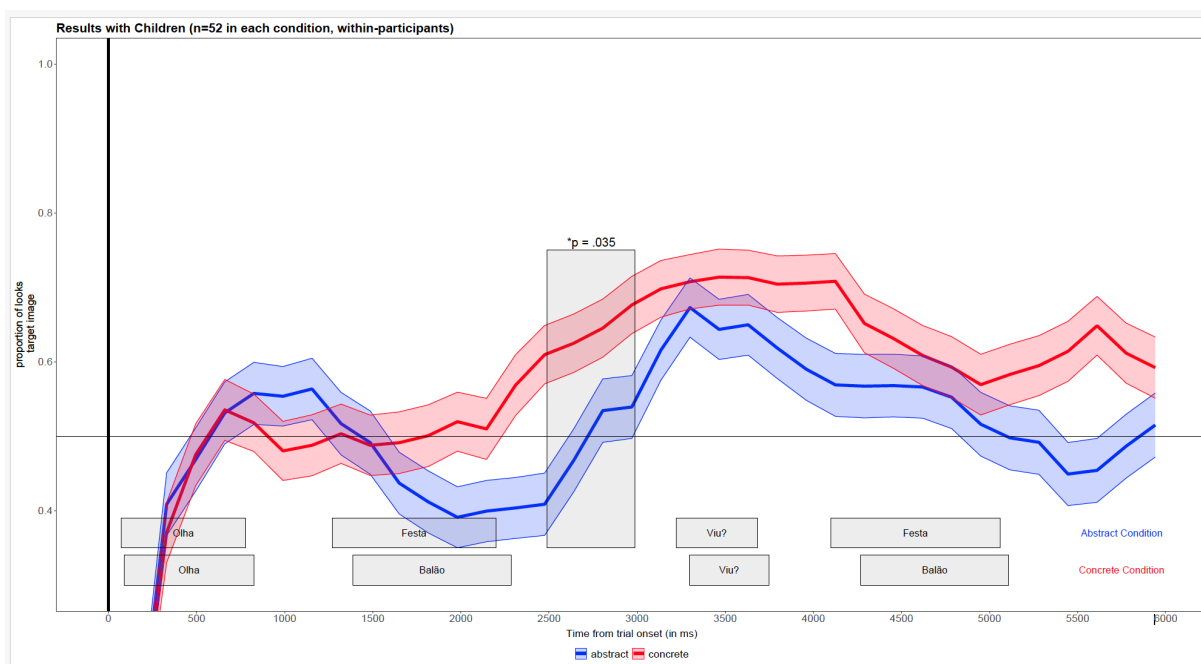
Fonte: Elaboração própria (2025).

De acordo com o gráfico, ambas as condições estão acima de 50%, i.e., acima do nível da chance, sendo a condição abstrata ligeiramente inferior em relação à concreta. Isso indica que crianças brasileiras entre 24 e 30 meses acessam e interpretam corretamente tanto as palavras concretas quanto as abstratas.

Uma análise de variância (ANOVA), considerando o tempo total de fixação ao alvo (*total looking time*) nas duas condições experimentais (concreta e abstrata), não revelou diferença estatisticamente significativa entre as condições. O efeito foi marginalmente significativo ($F(1,51) = 3.31, p = .075$). Além da ANOVA, também foi realizada uma análise baseada em *Cluster Permutation*, como no teste dos adultos, como um método mais sensível para identificar diferenças pontuais em experimentos com rastreamento ocular.

No gráfico 12, abaixo, está detalhado o resultado da análise de *Cluster Permutation*, evidenciando o período de tempo em que as curvas diferem de forma significativa.

Gráfico 12 — Diferenças temporais nos padrões de fixação ocular entre a condição concreta e a abstrata (*Cluster Permutation*) — experimento infantil

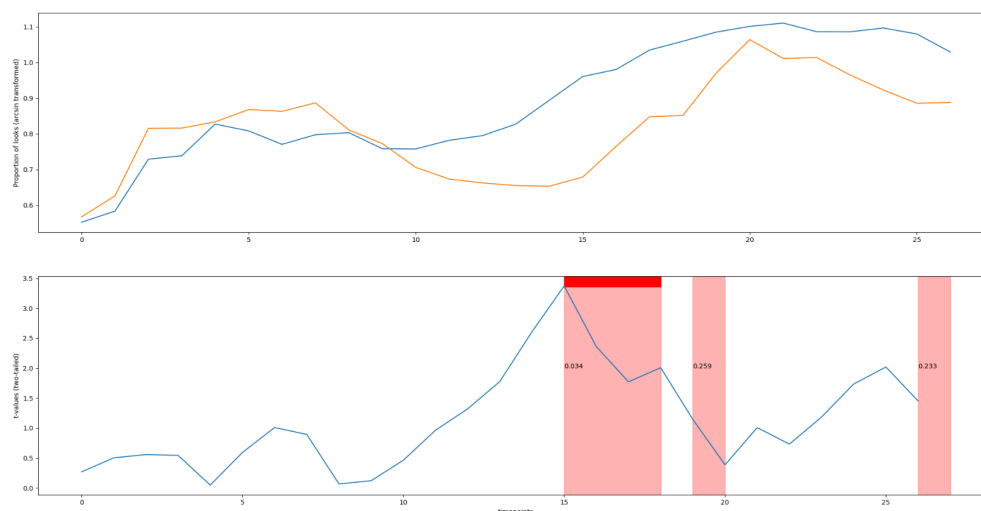


Fonte: Elaboração própria (2025).

O primeiro *cluster*, destacado em cinza, apresentou $p = .034$, o que indica uma diferença estatisticamente significativa entre as condições. Esse resultado reflete um efeito temporal para as palavras concretas, com maior engajamento visual dos participantes nessa condição.

O Gráfico 13, a seguir, apresenta a trajetória média de olhares para o alvo ao longo dos 6000 ms de cada *trial* experimental, nas duas condições (concreta e abstrata), no teste infantil.

Gráfico 13 — Trajetória temporal da proporção de olhares por condição (concreta e abstrata) — experimento infantil



Fonte: Elaboração própria (2025).

O gráfico apresenta a trajetória temporal da proporção de olhares para o alvo na condição concreta e na abstrata. No painel superior, a linha azul representa a condição concreta e a linha laranja representa a condição abstrata. Ambas permanecem acima da linha de chance (50%), o que indica que as crianças brasileiras, entre 24 e 30 meses, já direcionaram o olhar para o alvo de maneira consistente em ambas as condições.

Esse resultado demonstra que elas foram capazes de acessar esses itens lexicais e interpretar corretamente o referente após a escuta do estímulo auditivo. Mesmo os itens lexicais abstratos, que não possuem uma representação imagética direta, despertaram interesse visual sistemático, atendendo ao objetivo (i) da pesquisa, que era verificar se crianças nessa faixa etária já interpretam palavras concretas e abstratas.

No entanto, é importante destacar a diferença em relação à distribuição temporal do interesse visual entre as duas condições. A curva da condição concreta (representada pela cor azul) apresenta uma ascensão mais precoce e acentuada na proporção de olhares ao alvo, no intervalo entre aproximadamente o 15º e o 20º ponto de tempo analisado, correspondente a cerca de 2500–3000 ms após o início do *trial*, enquanto a curva da condição abstrata (em laranja) mostra um crescimento mais gradual ao longo do tempo. Essa diferença indica que o interesse visual pelas palavras concretas emergiu antes e de forma mais acentuada em relação às palavras abstratas. Esse padrão também foi observado no experimento com adultos.

Esse comportamento visual associado a cada tipo de item lexical vai diretamente ao encontro dos objetivos **(ii)** e **(iii)** deste estudo, os quais buscavam identificar possíveis diferenças cognitivas entre as condições e verificar se a identificação de palavras concretas seria primária em relação à de palavras abstratas. Esse achado é interessante, pois ele corrobora as nossas hipóteses de que não só as crianças nessa faixa etária são capazes de interpretar palavras concretas e abstratas, como também evidenciam diferenças entre esses itens quanto ao tempo de ativação, e não quanto à ausência de representação mental desses itens lexicais. A respeito dessa diferença, o painel inferior do gráfico apresenta a análise estatística por *Cluster Permutation*, utilizada para testar a significância das divergências entre as curvas de proporção de olhares ao longo do tempo.

Nossos resultados contribuem, também, para o desenvolvimento de inventários de aquisição lexical baseados em dados experimentais, conforme delineado no objetivo **(iv)**, e para a publicação de achados inéditos a respeito da arquitetura cognitiva, como previsto no objetivo **(v)**, especialmente sobre o processamento lexical de nomes concretos e abstratos por crianças entre 24 a 30 meses, adquirindo o PB. Esses dados permitem descrever com maior precisão o curso do desenvolvimento vocabular infantil, oferecendo uma base empírica independente da percepção dos responsáveis sobre o léxico das crianças. Dessa forma, nossos resultados contribuem para que pesquisadores, sobretudo da área de desenvolvimento cognitivo infantil, possam dispor de parâmetros mais robustos e condizentes com o repertório linguístico real das crianças nessa faixa etária.

Outra questão é que a variabilidade individual observada nas distribuições do *boxplot* (gráfico 11) pode ter implicações aplicadas, caso o teste venha a ser explorado como ferramenta preditiva, conforme nosso objetivo **(vi)**. A trajetória de crescimento do interesse visual e a proporção de olhares ao alvo em cada condição podem servir como indicadores comportamentais sensíveis à forma como as crianças direcionam a atenção a diferentes tipos de conceitos, em função da sua natureza conceitual. Esses parâmetros podem, em estudos futuros, ser explorados como possíveis marcadores de risco para transtornos do neurodesenvolvimento que afetam o processamento linguístico, sobretudo o de palavras mais abstratas. Essa questão será melhor detalhada na parte de encaminhamentos futuros desta pesquisa.

Por fim, os resultados deste experimento apontam para a viabilidade de medir efeitos temporais sutis a partir de experimentos remotos com o uso do *Lookit*. Isso reforça, conforme descrito no nosso objetivo **(vii)**, o potencial da plataforma para ampliar amostras de pesquisas

em desenvolvimento cognitivo no Brasil, com dados compatíveis com a realidade socioeconômica do nosso país.

É importante ressaltar que consideramos adequado fazer a análise dos resultados utilizando tanto a ANOVA, prática comum na área, quanto a análise baseada em *Clusters*. A justificativa para a realização de ambos os testes é que poderia haver um efeito temporal muito sutil que não apareceria na média geral. Assim, com o objetivo de identificar possíveis efeitos temporais que poderiam ser mascarados pela análise de médias, optamos também pela realização da análise de *Cluster*, mais adequada metodologicamente para experimentos com rastreamento ocular.

Essa decisão foi tomada, também, a partir da identificação de uma limitação na análise de dados em estudos com crianças, como os de Bergelson e Swingley (2012; 2013), reportados nesta tese. Nesses trabalhos, os autores compararam apenas as médias das condições por meio de ANOVAs, sem explorar de forma detalhada o curso temporal do processamento lexical em experimentos com rastreamento ocular.

Uma limitação deste tipo de análise diz respeito à possibilidade de resultados com baixo poder estatístico (*underpowered*), o que pode comprometer a confiabilidade dos resultados. Isso ocorre porque, além do número reduzido de participantes em experimentos com crianças, o qual foi convencionado, sem considerar estimativas formais de poder estatístico, muitos desses estudos contam com poucos itens por condição. Isso é problemático, uma vez que pode resultar em análises com baixo poder estatístico e em maior risco de ocorrência de erro tipo II, i.e., a não detecção de diferenças reais entre as condições experimentais, ou seja, falsos negativos (Oakes, 2017).

De forma a evitar esses efeitos, conduzimos a análise temporal (*time course*) baseada em *clusters*, com percentis de tempo, a fim de realizar uma descrição mais robusta para o tipo de experimento executado. Além disso, mantivemos a amostra compatível com as melhores práticas publicadas nos estudos em desenvolvimento infantil, com um número mínimo de 24 participantes por condição.

Tendo o nosso experimento duas condições, i.e., uma concreta e uma abstrata, temos um estudo com mais 48 participantes, distribuídos entre as condições com o número ideal de, pelo menos, 24 indivíduos por condição. Isso permite que o estudo, mesmo em amostras de difícil obtenção, como as que envolvem bebês, apresente alto poder estatístico, reduzindo o risco de erros do tipo II (Dixon, 2014; Lakens, 2014; Oakes, 2017).

É ainda pertinente destacar que, por o experimento ter sido executado de forma remota, possíveis atrasos em virtude da conexão não se configuram como um problema para o tipo de teste e de análise adotados neste estudo. A respeito de possíveis *delays* gerados em testes *online*, o próprio *Lookit*³⁹, a plataforma experimental, esclarece em seu manual (referência do manual - tradução da autora):

A medição do tempo de olhar tem início apenas quando o vídeo começa, e não durante o estabelecimento da conexão de vídeo. A medição do tempo de olhar começa somente quando as imagens são exibidas (ou quando o temporizador de “atraso” começa a contar, no caso de imagens apresentadas com atraso — mas, por exemplo, não durante a conexão da webcam). Desvios do olhar no início da apresentação não são contabilizados. Caso a criança não esteja olhando no momento inicial, a medição tem início assim que ela direciona o olhar pela primeira vez.

Assim, a contagem de tempo em estudos com rastreamento ocular só se inicia quando o vídeo/estímulo é apresentado, de modo que eventuais atrasos de conexão não interferem na janela de tempo de exibição das condições. Essa diferença pode, então, ser tratada como ruído, i.e., uma variação aleatória, ou como *lag*, em que o trecho inicial é ligeiramente anterior ao *onset* do estímulo auditivo.

Em um estudo conduzido por Becker et al (2021)⁴⁰, com experimentos executados com vídeos, o *lag* observado ficou entre 150–450ms para todos os *trials*. Dado que o nosso *design* é do tipo *within-subjects*, qualquer atraso residual afeta as condições igualmente. Por essa razão, esse atraso pode ser considerado um ruído não diferencial, em que a variância da duração do *trial* aumenta, mas isso não favorece uma condição em detrimento da outra.

Por fim, alguns outros critérios metodológicos conservadores em experimentos com rastreamento ocular foram adotados a fim de garantir a confiabilidade dos nossos resultados, como a padronização das imagens em relação ao tamanho e às cores, para evitar algum tipo de preferência visual pelas crianças, e a apresentação de um *attention grabber*, antes de cada *trial* experimental, a fim de assegurar que a criança iniciasse o *trial* com o olhar fixado no centro da tela, evitando vieses na direção do olhar (ou à direita ou à esquerda).

Em relação aos resultados, o curso temporal dos olhares foi analisado com janelas de tempo padronizadas (33 ms) e apenas crianças com, no mínimo, dois *trials* válidos por condição (concreta e abstrata) foram incluídas na análise (Delle Luche, 2015).

³⁹Para mais informações, consultar o manual do *Lookit*, disponível em: <https://lookit.readthedocs.io/projects/frameplayer/en/latest/components/exp-lookit-images-audio-infant-control/doc.html#when-looking-time-is-measured>.

⁴⁰Para mais detalhes sobre a acurácia de tempo em estudos do *Lookit*, recomendamos a leitura desta parte do manual, disponível em: <https://lookit.readthedocs.io/en/master/researchers-lag-issues.html>.

No próximo capítulo, serão discutidos os resultados de cada questionário realizado para a seleção dos itens lexicais, bem como dos dois experimentos — o controle e o infantil.

7 DISCUSSÃO DOS RESULTADOS

O questionário *Quais palavras sua criança já aprendeu?* foi elaborado com base em critérios metodológicos específicos da pesquisa, como explicado anteriormente. Diante da escassez de dados nos inventários de aquisição lexical acerca de quais palavras abstratas as crianças já compreendem ou compreendem e dizem, optamos pela criação de um novo questionário que permitisse a inclusão de itens lexicais de natureza mais abstrata e evitasse uma definição arbitrária por parte da pesquisadora.

Por isso, essa readaptação foi inspirada em modelos consagrados de inventários de desenvolvimento vocabular utilizados internacionalmente, como o *CDI* e o *Wordbank*. As palavras foram, portanto, selecionadas com base nas respostas fornecidas pelos próprios responsáveis de crianças entre 24 e 30 meses, a partir do questionário: *Quais palavras sua criança já aprendeu?*.

A partir desse levantamento, foi possível selecionar os itens lexicais que estivessem acima do nível da chance, ou seja, 50% para a classe *compreende e diz* de crianças brasileiras entre 24 e 30 meses para o experimento psicolinguístico. A aplicação de inventários de palavras é uma prática comum no campo da aquisição de linguagem, uma vez que confere maior robustez à seleção dos estímulos experimentais.

Desse modo, as palavras que compuseram o teste não foram escolhidas de modo aleatório, mas selecionadas a partir dessa adaptação de um instrumento de avaliação do desenvolvimento linguístico, preenchido pelos próprios responsáveis, sobre o vocabulário expressivo dessas crianças. Com a aplicação dos dois formulários, não só foi possível analisar a taxa de produção dessas palavras por crianças na faixa etária de interesse, mas também verificar como falantes nativos do português brasileiro categorizam esses itens em termos de concretude e abstração.

Conforme explicado ao longo desta tese, a definição de concretude e abstração de itens lexicais foi um ponto sensível, durante todo o desenvolvimento desta pesquisa. Essa questão tem sido abordada desde a filosofia até os estudos empíricos de neuroimagem, a fim de compreender como conceitos concretos e abstratos são representados e processados pela mente/cérebro.

Dada a natureza imprecisa dessa dimensão, tornou-se igualmente desafiador estabelecer critérios objetivos para a categorização dessas palavras. Por isso, julgamos pertinente elaborar, também, o questionário *Essa palavra é concreta ou abstrata?*, para que

adultos brasileiros realizassem a categorização de itens lexicais em relação à concretude/abstração. Dessa forma, os inventários elaborados constituíram soluções para equilibrar os dados da produção vocabular infantil e os dados da categorização por falantes adultos do PB. A partir deles, foi possível garantir coerência na definição e na seleção dessas palavras para os experimentos.

Além de cumprir o objetivo principal, esses dados ultrapassam os limites da pesquisa aqui desenvolvida, podendo contribuir para a ampliação do CDI brasileiro, com a inclusão de itens lexicais mais abstratos, bem como para a complementação de outros bancos de dados, como o LexporBR e o *Wordbank*, especialmente este último, que ainda não dispõe de dados sobre o PB. Ademais, o mapeamento desses conceitos pode servir como um parâmetro psicométrico em estudos futuros de neurociência cognitiva e de aquisição da linguagem, seja ela natural ou artificial.

Em relação ao experimento controle, realizado com adultos, falantes nativos do PB, a análise de *Cluster Permutation* apontou resultados favoráveis à hipótese experimental de que adultos brasileiros acessam e processam palavras concretas e abstratas de forma distinta. Esses achados vão ao encontro de evidências de estudos anteriores que apontam para uma distinção neurocognitiva no processamento de conceitos concretos por adultos, em tarefas com eletroencefalografia e com rastreamento ocular com bebês (Bergelson; Swingley, 2012, 2013; Huang *et al.*, 2010; Kounios; Holcomb, 1994).

No teste infantil, os dados obtidos são compatíveis com as nossas hipóteses de que crianças brasileiras, de 24 a 30 meses de idade, já são capazes de interpretar tanto conceitos concretos quanto abstratos, mas demonstraram maior interesse nos itens concretos, assim como os adultos. Embora as palavras abstratas já estejam presentes nas etapas iniciais da AL em crianças brasileiras, é interessante destacar que elas tiveram menos fixações nos conceitos abstratos do que nos concretos. Esses resultados corroboram as discussões trazidas ao longo deste trabalho sobre como as representações conceituais são acessadas e processadas cognitivamente.

A interpretação para a diferença no tempo de reação dos participantes, com significância estatística para as palavras concretas, pode ser explicada pelo fato de que tanto a imageabilidade quanto a experiência sensório-motora influenciam diretamente a associação palavra-conceito e a sua interpretação. Essa vantagem pode resultar da ativação de múltiplas áreas sensório-motoras durante o processamento de conceitos concretos, como mencionado na revisão bibliográfica deste trabalho, o que ocorre de forma distinta ou reduzida para palavras

abstratas, que possivelmente requerem a ativação de redes semânticas distintas (Brennan, 2022; Pulvermüller, 2013).

A identificação mais tardia das palavras abstratas pode estar associada à maior ativação de redes neurais associativas distribuídas no cérebro, as quais são mais dependentes de processos inferenciais e de suporte contextual (Bedny; Saxe, 2012; Crutch *et al.*, 2009). Palavras como *quente*, *medo*, *fome* podem, por exemplo, demandar a integração de informações sensoriais não visuais e de experiências sociais, o que pode tornar o mapeamento entre palavra e conceito mais complexo, conforme apontado por Paivio (1990) e por estudos recentes em cognição corporificada, como os resenhados na revisão sistemática deste trabalho (Borghesani; Piazza, 2017; Dreyer; Pulvermüller, 2018; Fernandino, 2015; Fernandino *et al.*, 2015; Kiefer; Pulvermüller, 2012, 2013; Reilly *et al.*, 2016; Sakreida *et al.*, 2013; Zdrzilova, 2017)

Alguns aspectos como a experiência sensorial, a tangibilidade dos conceitos, os mecanismos inferenciais e informações provenientes de outros domínios cognitivos podem atuar como facilitadores importantes no acesso às representações mentais/abstratas, conforme apontam alguns estudos com populações atípicas, como aqueles com crianças cegas (Bedny *et al.*, 2011; Landau; Gleitman, 1985), mencionados no início deste trabalho.

Os dados aqui apresentados reforçam que, tanto para adultos expostos à língua e com a gramática já consolidada quanto para crianças nas etapas iniciais da AL, o processamento de palavras abstratas é mais custoso cognitivamente do que o de concretas. É interessante destacar que, entre as crianças, a fixação visual em itens abstratos foi menor do que entre os adultos, o que sugere menor facilidade na identificação do referente nessa condição.

Esses dados são relevantes para avaliar como a experiência com o meio e outras cognições podem influenciar o desenvolvimento linguístico e cognitivo dos bebês, conforme sugerido por Borghi *et al.* (2018). É possível que conceitos que depreendam noções relacionadas a experiências sensoriais, fisiológicas e sociais, como *carinho*, *fome*, *alegria* e *vazio*, ativem redes semânticas mais amplas e, por isso, sejam acessados mais tardiamente do que itens lexicais que referenciam objetos concretos no mundo.

A partir das amostras do teste infantil, pretendíamos avaliar se o contexto socioeconômico das crianças impactaria a aquisição vocabular e o desenvolvimento cognitivo. Isso porque diversos estudos, inclusive com crianças brasileiras (Piccolo *et al.*, 2016b), apontam que aspectos socioeconômicos, como o nível de escolaridade dos responsáveis, a exposição a livros, a programas culturais ou a atividades e jogos que estimulem o

desenvolvimento cognitivo, estão associados a um melhor desempenho em testes linguísticos e em outras cognições já aos 18 e aos 24 meses de idade (Crosnoe *et al.*, 2010; Fernald *et al.*, 2013; Hoff, 2003; Tucker-Drob, 2011; Turkheimer *et al.*, 2003). Diante dessas evidências, tínhamos interesse em correlacionar os resultados de grupos de crianças provenientes de contextos menos e mais favorecidos socioeconomicamente, como os de creches públicas e privadas, a fim de realizar uma análise comparativa do desempenho no teste.

No entanto, a adesão de instituições particulares foi muito baixa, e os dados provenientes das crianças que realizaram o experimento pela plataforma do *Lookit* foram quantitativamente inferiores aos coletados nas creches públicas. Acreditamos que, mesmo com o envio do Termo de Consentimento Livre e Esclarecido para a autorização prévia dos responsáveis, as instituições particulares de ensino ficaram receosas por causa da gravação do rosto das crianças. Em função disso, não conseguimos fazer essa análise comparativa dos resultados a fim de investigar de que modo fatores sociais, educacionais e ambientais poderiam influenciar o desempenho dos participantes.

Em suma, os resultados do teste infantil corroboram as hipóteses delineadas nesta pesquisa acerca do acesso e das diferenças no processamento lexical de palavras concretas e abstratas por crianças brasileiras. Com esses dados empíricos, é possível fornecer evidências para a elaboração ou complementação de inventários de palavras sobre o PB, de forma a mapear as etapas da aquisição lexical por idade, sem que as predições sobre o desenvolvimento linguístico-cognitivo sejam feitas baseadas na intuição dos pais sobre quais palavras os filhos já compreendem ou compreendem e dizem.

Além disso, a partir dos resultados obtidos com crianças neurotípicas, investigamos a possibilidade de utilizar este teste como uma ferramenta linguística preditiva. A proporção de olhares para os alvos, nas duas condições experimentais, pode servir como um biomarcador comportamental em estudos comparativos futuros com crianças neurodivergentes, especialmente aquelas dentro do Transtorno do Espectro Autista (TEA), que tendem a apresentar maior dificuldade no processamento de conceitos mais abstratos.

Investigações sobre processos mentais e transtornos psicopatológicos que utilizam a linguagem como parâmetro têm possibilitado a identificação de marcadores biológicos e sociais por meio da observação de protocolos naturalistas em diferentes contextos. Um exemplo é a análise a partir do *SpeechGraphs*, que possibilitou identificar padrões característicos de patologias como a esquizofrenia, o transtorno bipolar e alguns tipos de demência, a partir da fala espontânea dos indivíduos (Chaves; Rodrigues, 2022; Chaves et al.,

2023; Mota, 2023). No entanto, ainda não há estudos que explorem esse tipo de abordagem em transtornos do neurodesenvolvimento, como o Transtorno do Espectro Autista (TEA). Nesse sentido, a ampliação desse modelo para investigações com crianças em diferentes estágios do desenvolvimento pode contribuir para a identificação precoce de alterações cognitivas.

Diante desse cenário, destacamos também o potencial do Lookit, plataforma utilizada nesta pesquisa, para a execução de experimentos em ciências cognitivas de forma naturalística e escalável, permitindo a coleta de amostras mais heterogêneas e representativas da população brasileira.

No capítulo seguinte, serão descritas as considerações finais desta pesquisa e os encaminhamentos futuros.

8 CONSIDERAÇÕES FINAIS

Esta investigação buscou compreender as etapas iniciais da AL, oferecendo evidências empíricas sobre o acesso e o processamento de palavras concretas e abstratas tanto por crianças brasileiras entre 24 e 30 meses de idade quanto por adultos.

Embora tenha sido realizada uma extensa revisão sistemática da literatura acerca da aquisição de nomes concretos e abstratos, não foram encontradas pesquisas sobre o PB. Apenas o estudo de Gro (2007) investigou *priming* semântico em crianças norueguesas de 24 meses. Além disso, estudos como os de Bergelson e Swingley (2012, 2013, 2015), Lew-Williams (2019) e Reuter *et al.* (2023), embora não mapeados na revisão sistemática, foram incluídos na bibliografia da tese por tratarem da aquisição lexical e por serem relevantes tanto para o delineamento do *design* experimental quanto para a definição da faixa etária das crianças participantes deste trabalho.

Diante dessa escassez de estudos na área de AL, o presente estudo preenche uma lacuna importante na literatura em neurociência da linguagem ao fornecer evidências empíricas inéditas sobre o acesso e o processamento de palavras concretas e abstratas, por crianças brasileiras entre 24 e 30 meses de vida. Além disso, essa pesquisa também contribui para a área de AL por meio da elaboração de um inventário de palavras nos moldes do CDI, incluindo palavras abstratas, e com as definições de concretude/abstração, realizadas por falantes nativos do PB.

Outro aspecto inovador deste trabalho é o fato de ter sido o primeiro experimento realizado por pesquisadores brasileiros na plataforma do *Lookit*. Ao divulgar essa ferramenta, que alia confiabilidade científica e baixo custo operacional (é preciso apenas conexão à internet), buscamos incentivar a execução de experimentos em psicolinguística de forma remota no Brasil, sobretudo com populações mais difíceis de serem recrutadas, como crianças. Isso traz implicações relevantes para o contexto científico, uma vez que amplia a representatividade de amostras populacionais, sobretudo de crianças brasileiras, como também fortalece a produção de mais estudos conduzidos por pesquisadores brasileiros.

Além disso, este estudo propõe a aplicação desse experimento como um possível instrumento de avaliação preditiva de habilidades linguísticas para mapear possíveis *déficits* cognitivos-linguísticos, ainda na primeira infância, por meio de um protocolo experimental de fácil aplicação e de baixo custo, adequado ao contexto socioeconômico brasileiro. Nessa perspectiva, nossos resultados reforçam a importância de ampliar bases de dados sobre o

desenvolvimento lexical infantil no Brasil, capazes de orientar tanto a pesquisa básica quanto aplicações educacionais e clínicas sobre o desenvolvimento linguístico-cognitivo.

A partir de evidências empíricas sobre a arquitetura cognitiva de itens concretos e abstratos, esta tese contribui para uma ciência mais inclusiva e contextualizada. Nesse sentido, buscamos promover um impacto social no fazer ciência ao aproximar a produção acadêmica de questões ainda pouco exploradas sobre o desenvolvimento cognitivo, compatíveis com a realidade socioeconômica do nosso país.

Como fio condutor para os encaminhamentos, nossos achados convergem em duas frentes complementares. Em adultos falantes nativos do PB, a análise por *Cluster Permutation* revelou latências menores e processamento mais eficiente para itens concretos do que para abstratos, alinhando-se a evidências de que conceitos concretos recrutam redes sensório-motoras mais diretamente acessíveis. Em crianças brasileiras entre 24 e 30 meses, observamos um padrão convergente, com maior interesse e mais fixações para os itens concretos, sugerindo que a imagem mental e a experiência sensório-motora facilitam o mapeamento palavra–conceito nas etapas iniciais da aquisição, embora já sejam capazes de interpretar tanto as palavras concretas quanto as abstratas.

Em contraste, a identificação relativamente mais tardia de palavras abstratas é compatível com a necessidade de integrar pistas contextuais e de recrutar redes semânticas distribuídas, processos mais dependentes de inferência. Embora tivéssemos planejado testar o impacto do contexto socioeconômico, a baixa adesão de creches privadas impediu a comparação sistemática entre grupos; ainda assim, o conjunto de resultados indica que a concretude confere uma vantagem robusta tanto no acesso quanto na interpretação lexical em ambos os grupos etários. É a partir desse quadro, marcado por uma assimetria estável pró-concreto e por um maior custo cognitivo para o abstrato, que delineamos, no que segue, as prioridades empíricas e metodológicas para os próximos estudos.

8.1 ENCAMINHAMENTOS FUTUROS

Durante a elaboração deste trabalho, a complexidade do acesso lexical e das representações mentais envolvidas na aquisição de itens lexicais concretos e abstratos evidenciou, a partir de dados empíricos sobre o desenvolvimento cognitivo típico de crianças brasileiras entre 24 e 30 meses, o papel do experimento aqui desenvolvido como uma possível

ferramenta de investigação comparativa do desempenho entre indivíduos neurotípicos e neurodivergentes.

À luz desses achados, levantamos a possibilidade de utilizar este experimento como um teste de habilidade preditiva linguística para crianças neurodivergentes (American Psychiatric Association, 2023). Isso porque alguns transtornos do neurodesenvolvimento se caracterizam por limitações no desenvolvimento ou por diferenças nos processos cerebrais, com prejuízos na comunicação e na interação social dos indivíduos, especialmente na compreensão e na interpretação de conceitos abstratos, como ocorre no Transtorno do Espectro Autista (TEA).

Outra questão pertinente é que, em geral, esses *déficits* se manifestam precocemente, antes mesmo do início da vida escolar. De acordo com Peyre et al. (2017), os marcos de linguagem aos 24 meses explicam cerca de 20% da variação no QI de crianças entre 5 e 6 anos de idade, sendo a linguagem o melhor preditor de QI futuro, superando habilidades motoras e sociais. Em contrapartida, crianças que apresentam atrasos nesses marcos tendem a exibir escores de QI mais baixos em etapas posteriores do desenvolvimento.

O período entre 12 e 24 meses é, portanto, marcado por aumentos significativos na compreensão e na ampliação do vocabulário (Fernald et al., 2006, 2008; Marchman; Fernald, 2008). Diante dessas evidências, as habilidades linguísticas são um dos indicadores mais sensíveis para a identificação precoce de *déficits* cognitivos, com maior precisão a partir dos dois anos de idade. Por isso, contrastar dados do desenvolvimento linguístico-cognitivo de crianças neurotípicas entre 24 e 30 meses de vida, como realizado neste trabalho, com os de indivíduos neurodivergentes, na mesma faixa etária, pode constituir uma ferramenta promissora para a identificação de um possível transtorno do neurodesenvolvimento.

A análise do desenvolvimento vocabular e das habilidades cognitivas associadas ao pensamento abstrato, no âmbito do acesso lexical e do mapeamento vocabular, logo nos dois primeiros anos de vida de uma criança, pode desempenhar um papel crucial na investigação de possíveis atrasos de fala. Além disso, esse tipo de análise pode facilitar a implementação de intervenções precoces durante períodos críticos do desenvolvimento, em que há maior plasticidade neural e, portanto, maior potencial de sucesso terapêutico (Nelson, 2000; Werker; Hensch, 2015).

Essas medidas não apenas contribuem para uma melhor qualidade de vida dos indivíduos, como previnem danos cognitivos mais graves. Esses resultados têm sido evidenciados em estudos envolvendo diferentes transtornos do neurodesenvolvimento e

psiquiátricos (Malcorra et al., 2024; Mota et al., 2017; Mota et al., 2023; Tenenbaum et al., 2024), inclusive em tarefas que utilizaram o paradigma da Fixação Preferencial do Olhar, semelhantes ao experimento aqui realizado.

Portanto, os desdobramentos desta tese apontam o potencial da psicolinguística para integrar a ciência básica e a aplicada, com implicações intersetoriais relevantes. Os dados apresentados podem contribuir significativamente para a elaboração de políticas públicas no âmbito da educação e da saúde, oferecendo parâmetros objetivos sobre o desenvolvimento linguístico-cognitivo e promovendo práticas educacionais e terapêuticas baseadas em evidências científicas.

REFERÊNCIAS

- ALVES, M. V.; EKUNI, R.; HERMIDA, M. J.; VALLE-LISBOA, J. (ed.). *Cognitive sciences and education in non-WEIRD populations: a Latin American perspective*. Springer Nature, 2022.
- AMEDI, A.; FLOEL, A.; KNECHT, S.; ZOHARY, E.; COHEN, L. G. Transcranial magnetic stimulation of the occipital pole interferes with verbal processing in blind subjects. *Nature Neuroscience*, v. 7, n. 11, p. 1266–1270, 2004. DOI: [10.1038/nn1328](https://doi.org/10.1038/nn1328).
- AMERICAN PSYCHIATRIC ASSOCIATION. *Manual diagnóstico e estatístico de transtornos mentais: DSM-5-TR*. 5. ed., texto revisado. Porto Alegre: Artmed, 2023.
- ANTONUCCI, S. M.; ALT, M. A lifespan perspective on semantic processing of concrete concepts: does a sensory/motor model have the potential to bridge the gap?. *Cognitive, Affective, & Behavioral Neuroscience*, v. 11, p. 551–572, 2011. DOI: <https://doi.org/10.3758/s13415-011-0053-y>.
- ARNAULD, A.; LANCELOT, C. *Gramática de Port-Royal*. São Paulo: Martins Fontes, 2001.
- ASSOCIATION FOR PSYCHOLOGICAL SCIENCE. *Recruit online participants with Children Helping Science*. *Psychological Science*, 2020. Disponível em: <https://www.psychologicalscience.org/publications/observer/announcements/recruit-online-participants-with-children-helping-science.html>. Acesso em: 15 jul. 2025.
- ATTIG, M.; WEINERT, S. What impacts early language skills? Effects of social disparities and different process characteristics of the home learning environment in the first 2 years. *Frontiers in Psychology*, [S.l.], v. 11, p. 1–14, 2020. DOI: <https://doi.org/10.3389/fpsyg.2020.557751>.
- AYERS, M. *Locke: ideias e coisas*. Tradução de José Oscar de Almeida Marques. São Paulo: Editora Unesp, 2001.
- BAHN, D.; VESKER, M.; GARCÍA ALANIS, J. C.; SCHWARZER, G.; KAUSCHKE, C. Age-dependent positivity-bias in children's processing of emotion terms. *Frontiers in Psychology*, v. 8, p. 1268, 2017. DOI: <https://doi.org/10.3389/fpsyg.2017.01268>.
- BÁNKI, A.; DE ECCHER, M.; FALSCHLEHNER, L.; HOEHL, S.; MARKOVA, G. Comparing online webcam-and laboratory-based eye-tracking for the assessment of Infants' audio-visual synchrony perception. *Frontiers in Psychology*, v. 12, p. 1–19, 2022. Disponível em: <https://doi.org/10.3389/fpsyg.2021.733933>.
- BEDNY, M.; PASCUAL-LEONE, A.; DODELL-FEDER, D.; FEDORENKO, E.; SAXE, R. Language processing in the occipital cortex of congenitally blind adults. *Proceedings of the National Academy of Sciences of the United States of America*, Washington, v. 108, n. 11, p. 4429–4434, 15 mar. 2011. DOI: <https://doi.org/10.1073/pnas.1014818108>.

BEDNY, M.; SAXE, R. Insights into the origins of knowledge from the cognitive neuroscience of blindness. *Cognitive Neuropsychology*, v. 29, n. 1–2, p. 56–84, 2012. Disponível em: <https://doi.org/10.1080/02643294.2012.713342>.

BERGELSON, E.; SWINGLEY, D. At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, v. 109, n. 9, p. 3253–3258, 2012. DOI: <https://doi.org/10.1073/pnas.1113380109>.

BERGELSON, E.; SWINGLEY, D. The acquisition of abstract words by young infants. *Cognition*, v. 127, n. 3, p. 391–397, 2013. DOI: <https://doi.org/10.1016/j.cognition.2013.02.011>.

BERGELSON, E.; SWINGLEY, D. Early word comprehension in infants: replication and extension. *Language Learning and Development*, v. 11, n. 4, p. 369–380, 2015. DOI: [10.1080/15475441.2014.979387](https://doi.org/10.1080/15475441.2014.979387).

BERENDS, S. M.; BROUWER, S. M.; SPRENGER, S. A. Eye-tracking and the visual world paradigm. In: SCHMID, M. S.; KOOPMANS-VAN BEINUM, F. (org.). *Designing research on bilingual development*. Cham: Springer, 2016. p. 55–80. DOI: https://link.springer.com/chapter/10.1007/978-3-319-11529-0_5.

BINDER, J. R.; DESAI, R. H.; GRAVES, W. W.; CONANT, L. L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, v. 19, n. 12, p. 2767–2796, 2009. DOI: <https://doi.org/10.1093/cercor/bhp055>

BINDER, J. R.; DESAI, R. H. The neurobiology of semantic memory. *Trends in Cognitive Sciences*, v. 15, n. 11, p. 527–536, 2011. DOI: <https://doi.org/10.1016/j.tics.2011.10.001>

BŁASZCZAK, J.; KLIMEK-JANKOWSKA, D. Noun and verb in the mind: An interdisciplinary approach. In: BŁASZCZAK, J.; KLIMEK-JANKOWSKA, D.; MIGDAŁSKI, K. *How categorical are categories?: new approaches to the old questions of noun, verb, and adjective*. Berlin: De Gruyter Mouton, 2015. p. 75–112. DOI: <https://doi.org/10.1515/9781614514510>.

BOAS, F. *Antropologia cultural*. Trad. Celso Castro. Rio de Janeiro: Zahar Editora, 2007.

BORGHESANI, V.; PIAZZA, M. The neuro-cognitive representations of symbols: the case of concrete words. *Neuropsychologia*, [S. l.], v. 105, p. 4–17, 2017. DOI: <https://doi.org/10.1016/j.neuropsychologia.2017.06.026>.

BORGHI, A. M.; BARCA, L.; BINKOFSKI, F.; TUMMOLINI, L. *Varieties of abstract concepts: development, use and representation in the brain*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, [S.l.], v. 373, 2018. DOI: [http://doi.org/10.1098/rstb.2017.0121](https://doi.org/10.1098/rstb.2017.0121).

BORGHI, A. M.; BARCA, L.; BINKOFSKI, F.; CASTELFRANCHI, C.; PEZZULO, G.; TUMMOLINI, L. Words as social tools: flexibility, situatedness, language and sociality in abstract concepts: reply to comments on "Words as social tools: language, sociality and inner grounding in abstract concepts". *Physics of Life Reviews*, [S.l.], v. 29, p. 178–184, 2019. DOI:

<https://doi.org/10.1016/j.plrev.2019.06.004>.

BORGHI, A. M.; BINKOFSKI, F. *Words as social tools: an embodied view on abstract concepts*. Nova Iorque: Springer New York, 2014. DOI: 10.1007/978-1-4614-9539-0.

BORGHI, A. M. *et al.* Abstract concepts, language and sociality: from acquisition to inner speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 373, n. 1752, p. 20170134, 2018. DOI: <https://doi.org/10.1098/rstb.2017.0134>.

BOOTH, A. E.; WAXMAN, S. Object names and object functions serve as cues to categories for infants. *Developmental Psychology*, Washington, v. 38, n. 6, p. 948–957, 2002. DOI: <https://doi.org/10.1037/0012-1649.38.6.948>.

BOWERS, J. M. *Hemispheric difference in semantic priming for early and late age of acquisition words*. 2008. Dissertação (Mestrado) – Oklahoma State University, 2008.

BRENNAN, J. R. *Language and the brain: a slim guide to neurolinguistics*. Oxford: Oxford University Press, 2022.

BRYSSBAERT, M.; BUCHMEIER, M.; CONRAD, M.; JACOBS, A. M.; BOLTE, J.; BÖHL, A. The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, v. 58, p. 412–424, 2011. DOI: <https://doi.org/10.1027/1618-3169/a000123>.

BRUHL, A. *The role of situations and thematic reorganization in the conceptual processing of abstract concepts*. 2014. Tese (Doutorado) – Lyon 2, 2014.

CAUVET, E.; LIMISSURI, R.; MILLOTTE, S.; SKORUPPA, K.; CABROL, D.; CHRISTOPHE, A. Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development*, v. 10, n. 1, p. 1–18, 2013. Disponível em: <https://doi.org/10.1080/15475441.2012.757970>.

CHAVES, M. F. F.; RODRIGUES, C. *Linguistic marker of schizotypy: a study on nominal reference*. **Revista da ABRALIN**, [S. l.], v. 21, n. 1, 2022. DOI: 10.25189/abralin.v21i1.2092. Disponível em: <https://revista.abralin.org/index.php/abralin/article/view/2092>

CHAVES, M. F.; RODRIGUES, C.; RIBEIRO, S.; MOTA, N. B.; COPELLI, M. *Grammatical impairment in schizophrenia: An exploratory study of the pronominal and sentential domains*. **PLoS ONE**, San Francisco, v. 18, n. 9, e0291446, 2023. DOI: <https://doi.org/10.1371/journal.pone.0291446>

CHOMSKY, N. *Linguística cartesiana*. Petrópolis: Vozes, 1972.

CHOMSKY, N. *Aspects of the Theory of Syntax*. Cambridge: MIT press, 1965.

CORRÊA, L.; NAME, M. The processing of determiner – noun agreement and the identification of the gender of nouns in the early acquisition of Portuguese. *Journal of Portuguese Linguistics*, v. 2, n. 1, p. 19–43, 2003. DOI: 10.5334/jpl.34.

COTTINGHAM, J. *Descartes*. São Paulo: Editora UNESP, 1999.

CUARTAS, J.; McCOY, D. C.; TORRES, I.; BURGHARDT, L.; SHONKOFF, J. P.; YOSHIKAWA, H. *The developmental consequences of early exposure to climate change-related risks*. *Child Development Perspectives*, v. 18, n. 3, p. 145–154, 2024. DOI: <https://doi.org/10.1111/cdep.12503>

CROSNOE, R.; LEVENTHAL, T.; WIRTH, R. J.; PIERCE, K. M.; PIANTA, R. C.; NICHD EARLY CHILD CARE RESEARCH NETWORK. Family socioeconomic status and consistent environmental stimulation in early childhood. *Child Development*, [S.l.], v. 81, n. 3, p. 972–987, 2010. DOI: <https://doi.org/10.1111/j.1467-8624.2010.01446.x>.

CRUTCH, S. J.; CONNELL, S.; WARRINGTON, E. K. The different representational frameworks underpinning abstract and concrete knowledge: evidence from odd-one-out judgements. *Quarterly Journal of Experimental Psychology*, v. 62, n. 7, p. 1377–1390, 2009. DOI: <https://doi.org/10.1080/17470210802483834>.

De BONI, L. A. *Lógica e linguagem na Idade Média*. Porto Alegre: EDIPUCRS, 1995.

De CARVALHO, A.; DAUTRICHE, I.; LIN, I.; CHRISTOPHE, A. Phrasal prosody constrains syntactic analysis in toddlers. *Cognition*, [S. l.], v. 163, p. 67–79, 2017. DOI: <https://doi.org/10.1016/j.cognition.2017.02.018>.

De CARVALHO, A.; HE, A. X.; LIDZ, J.; CHRISTOPHE, A. Prosody and function words cue the acquisition of word meanings in 18-month-old infants. *Psychological Science*, v. 30, n. 3, p. 319–332, 2019. DOI: <https://doi.org/10.1177/0956797618814131>.

De CARVALHO, C. de. *Para compreender Saussure*. Rio de Janeiro: Editora Rio, 1976.

De VEGA, M.; GRAESSER, A. C.; GLENBERG, A. M. Reflecting on the debate. In: de VEGA, M.; GLENBERG, A.; GRAESSER, A. C. (ed.). *Symbols and embodiment: debates on meaning and cognition*. Oxford: Oxford University Press, 2008. p. 397–440.

De SAUSSURE, F. *Curso de linguística geral*. Payot Paris, 1916.

DÓCZI, B.; KORMOS, J. *Longitudinal developments in vocabulary knowledge and lexical organization*. Oxford: Oxford University Press, 2016.

DREYER, F. R.; Pulvermüller, F. Abstract semantics in the motor system?—An event-related fMRI study on passive reading of semantic word categories carrying abstract emotional and mental meaning. *Cortex*, v. 100, p. 52–70, 2018. DOI: <https://doi.org/10.1016/j.cortex.2017.10.021>.

ELAN. *Version 6.8* [Software]. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive, 2024.

EREL, Y.; POTTER, C. E.; JAFFE-DAX, S.; LEW-WILLIAMS, C.; BERMANO, A. H. iCatcher: A neural network approach for automated coding of young children's eye

movements. *Infancy*, v. 27, n. 4, p. 765–779, 2022. DOI: <https://doi.org/10.1111/infa.12468>.

ESTIVALET, G. L.; HARTMANN, N. S.; MARQUIAFANEL, V.; LUKASOVA, K.; CARTHERY-GOULART, M. T.; ALUÍSIO, S. M. *Léxico do português infantil*, 2017. Disponível em: <https://www.lexicodoportugues.com/infantil/>.

FERNALD, A.; MARCHMAN, V. A.; WEISLEDER, A. *SES differences in language processing skill and vocabulary are evident at 18 months*. *Developmental Science*, [S.l.], v. 16, n. 2, p. 234–248, 2013. DOI: <https://doi.org/10.1111/desc.12019>.

FERNALD, A.; PERFORS, A.; MARCHMAN, V. A. Picking up speed in understanding: speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, v. 42, n. 1, p. 98–116, 2006. DOI: <https://doi.org/10.1037/0012-1649.42.1.98>.

FERNANDINO, L.; HUMPHRIES, C. J.; SEIDENBERG, M. S.; GROSS, W. L.; CONANT, L. L.; BINDER, J. R. Predicting brain activation patterns associated with individual lexical concepts based on five sensory-motor attributes. *Neuropsychologia*, v. 76, p. 17–26, 2015. DOI: <https://doi.org/10.1016/j.neuropsychologia.2015.04.009>.

FERNANDINO, L.; BINDER, J. R.; DESAI, R. H.; PENDL, S. L.; HUMPHRIES, C. J.; GROSS, W. L.; CONANT, L. L.; SEIDENBERG, M. S. Concept representation reflects multimodal abstraction: a framework for embodied semantics. *Cerebral Cortex*, v. 26, n. 5, p. 2018–2034, 2016. DOI: <https://doi.org/10.1093/cercor/bhv020>.

FERRÉ, P.; VENTURA, D.; COMESAÑA, M.; FRAGA, I. The role of emotionality in the acquisition of new concrete and abstract words. *Frontiers in Psychology*, v. 6, p. 976, 2015. DOI: <https://doi.org/10.3389/fpsyg.2015.00976>.

FRANÇA, A. I. *Concatenações linguísticas: estudo de diferentes módulos cognitivos na aquisição e no córtex*. 2002. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 2002.

FRANÇA, A. I.; FERRARI, L.; MAIA, M. *A linguística no século XXI: convergências e divergências no estudo da linguagem*. São Paulo: Contexto, 2016. 224 p.1

FROTA, S.; NAME, C. Questões de percepção em língua materna. In: FREITAS, Maria João; SANTOS, Ana Lúcia (org.). *Aquisição de língua materna e não materna: questões gerais e dados do português*. Berlin: Language Science Press, 2017. p. 35–50. DOI: <https://doi.org/10.5281/zenodo.889261>.

GANDÍA, M. A. *On word definition in children and adults: effects of word category and level of abstraction*. Universitat de Barcelona (Spain), 2016.

GLEITMAN, L. The structural sources of verb meanings. *Language acquisition*, v. 1, n. 1, p. 3–55, 1990. DOI: https://doi.org/10.1207/s15327817la0101_2.

GLEITMAN, L. R.; CASSIDY, K.; NAPPA, R.; PAPAFRAGOU, A.; TRUESWELL, J. C. Hard Words. *Language Learning and Development*, v. 1, n. 1, p. 23–64, 2005. DOI: https://doi.org/10.1207/s15473341l1d0101_4.

GLEITMAN, L. R.; TRUESWELL, J. C. Easy words: reference resolution in a malevolent referent world. *Topics in Cognitive Science*, v. 12, n. 1, p. 22–47, 2020. DOI: <https://doi.org/10.1111/tops.12352>.

HAERI, A. M. *Conceptual representation in human cognitive system: differences between concrete and abstract words and concepts*. 2017. Tese (Doutorado) – Università di Pisa, 2017.

HENNINGSEN-SCHOMERS, M. R.; Pulvermüller, F. Modelling concrete and abstract concepts using brain-constrained deep neural networks. *Psychological Research*, v. 86, n. 8, p. 2533–2559, 2022. DOI: <https://doi.org/10.1007/s00426-021-01591-6>.

HENRICH, J.; HEINE, S. J.; NORENZAYAN, A. The weirdest people in the world? *Behavioral and Brain Sciences*, v. 33, n. 2–3, p. 61–135, 2010. DOI: <https://doi.org/10.1017/S0140525X0999152X>.

HENSCH, T. K. Critical period regulation. *Annual Review of Neuroscience*, [S.l.], v. 27, p. 549–579, 2004. DOI: <https://doi.org/10.1146/annurev.neuro.27.070203.144327>.

HILDEBRANDT, R. *Investigating emotion-label and emotion-laden words in a semantic satiation paradigm*. 2020. Dissertação (Mestrado) – Rochester Institute of Technology, 2020.

HINOJOSA, J. A.; MORENO, E. M.; FERRÉ, P. Affective neurolinguistics: towards a framework for reconciling language and emotion. *Language, Cognition and Neuroscience*, v. 35, n. 7, p. 813–839, 2019. DOI: <https://doi.org/10.1080/23273798.2019.1620957>.

HOFF, E. The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, [S.l.], v. 74, n. 5, p. 1368–1378, 2003. DOI: <https://doi.org/10.1111/1467-8624.00612>.

HOWELL, S. R. *Sensorimotor representations of meaning in early language acquisition*. 2006. Tese (Doutorado) – Library and Archives Canada. Bibliothèque et Archives Canada, Ottawa, 2006.

HOWELL, S. R.; JANKOWICZ, D.; BECKER, S. A model of grounded language acquisition: sensorimotor features improve lexical and grammatical learning. *Journal of Memory and Language*, v. 53, n. 2, p. 258–276, 2005. DOI: <https://doi.org/10.1016/j.jml.2005.03.002>.

HUANG, H. W.; FEDERMEIER, K. D. Imaginative language: what event-related potentials have revealed about the nature and source of concreteness effects. *Language and Linguistics*, v. 16, n. 4, p. 503–515, 2015. DOI: [10.1177/1606822X15583233](https://doi.org/10.1177/1606822X15583233).

HUETTIG, F.; ROMMERS, J.; MEYER, A. S. Using the visual world paradigm to study language processing: a review and critical evaluation. *Acta Psychologica*, v. 137, n. 2, p. 151–171, 2011. DOI: <https://doi.org/10.1016/j.actpsy.2010.11.003>.

HUANG, H.-W.; LEE, C.-L.; FEDERMEIER, K. D. Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts.

NeuroImage, v. 49, n. 1, p. 1116–1123, 2010. DOI: <https://doi.org/10.1016/j.neuroimage.2009.07.031>.

HUTTON, J. S.; DUDLEY, J.; HOROWITZ-KRAUS, T.; DEWITT, T.; HOLLAND, S. K. Associations between home literacy environment, brain white matter integrity and cognitive abilities in preschool-age children. *Acta Paediatrica*, v. 109, n. 7, p. 1376–1386, 2020. DOI: 10.1111/apa.15124. Disponível em: <https://doi.org/10.1111/apa.15124>.

JACKENDOFF, R. What is a concept, that a person may grasp it? In: MARGOLIS, E; LAURENCE, S. (ed.). *Concepts: core readings*. Cambridge, MA: MIT Press, 1999. p. 305–333.

JIROUT, J.; LOCASALE-CROUCH, J.; TURNBULL, K.; GU, Y.; CUBIDES, M.; GARZIONE, S.; EVANS, T. M.; WELTMAN, A. L.; KRANZ, S. *How lifestyle factors affect cognitive and executive function and the ability to learn in children*. *Nutrients*, v. 11, 1953, 2019. DOI: 10.3390/nu11081953. Disponível em: <https://doi.org/10.3390/nu11081953>

KADUK, T.; GOEKE, C.; FINGER, H.; KÖNIG, P. Webcam eye tracking close to laboratory standards: comparing a new webcam-based system and the EyeLink 1000. *Behavioral Research Methods*, v. 56, n. 5, p. 5002–5022, ago. 2023. DOI: <https://doi.org/10.3758/s13428-023-02237-8>.

KIDD, E.; GARCIA, R. How diverse is child language acquisition research? *First Language*, v. 42, n. 6, p. 703–735, 2022. DOI: 10.1177/014272372110664.

KIEFER, M.; Pulvermüller, F. Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex*, v. 48, n. 7, p. 805–825, 2012. DOI: <https://doi.org/10.1016/j.cortex.2011.04.006>.

KIM, J. M.; SIDHU, D. M.; PEXMAN, P. M. Effects of emotional valence and concreteness on children's recognition memory. *Frontiers in Psychology*, v. 11, 615041, 2020. DOI: <https://doi.org/10.3389/fpsyg.2020.615041>.

KOUNIOS, J.; HOLCOMB, P. J. Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, v. 20, n. 4, p. 804, 1994. DOI: <https://doi.org/10.1037//0278-7393.20.4.804>.

KUTAS, M.; FEDERMEIER, K. D. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, v. 62, n. 1, p. 621–647, 2011. DOI: <https://doi.org/10.1146/annurev.psych.093008.131123>.

LANDAU, B.; GLEITMAN, L. R. *Language and experience: evidence from the blind child*. Cambridge: Harvard University Press, 1985.

LI, P. Meaning in acquisition: semantic structure, lexical organization, and crosslinguistic variation. In: GATHERCOLE, V. *Routes to language: studies in honor of Melissa Bowerman*. Nova Iorque: Psychology Press, 2010. p. 257–283.

LOCKE, J. *An essay concerning human understanding*. Philadelphia: Kay & Troutman, 1847.

- LUCHKINA, E.; WAXMAN, S. Semantic priming supports infants' ability to learn names of unseen objects. *PloS one*, v. 20, n. 4, p. e0321775, 2025. DOI: <https://doi.org/10.1371/journal.pone.0321775>
- LUPYAN, G.; WINTER, B. Language is more abstract than you think, or, why aren't languages more iconic? *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 373, n. 1752, p. 20170137, 2018. DOI: 10.1098/rstb.2017.0137.
- MALCORRA, B.; RIBEIRO, M.; JENSEN, L.; GOMES, G.; MELETTI, T.; MOTA, N. B. Speech connectedness predicts reading performance three months in advance: a longitudinal experiment. *npj Science of Learning*, v. 9, n. 1, p. 35, 2024. DOI: <https://doi.org/10.1038/s41539-024-00248-4>.
- MARCHMAN, V. A.; FERNALD, A. Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, v. 11, n. 3, p. F9–F16, 2008. DOI: <https://doi.org/10.1111/j.1467-7687.2008.00671.x>.
- MARCONDES, D. *Textos básicos de linguagem: de Platão a Foucault*. Zahar, 2010.
- MARGOLIS, E.; LAURENCE, S. *Concepts: core readings*. Cambridge: The MIT Press, 1999.
- MARIS, E.; OOSTENVELD, R. Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, v. 164, n. 1, p. 177–190, 2007. DOI: <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- MESTRES-MISSÉ, A. *Neural correlates of word learning and meaning acquisition*. 2007. Tese (Doutorado) – Universitat de Barcelona, 2007.
- MESTRES-MISSÉ, A.; MÜNTE, T. F.; RODRIGUEZ-FORNELLS, A. Functional neuroanatomy of contextual acquisition of concrete and abstract words. *Journal of Cognitive Neuroscience*, v. 21, n. 11, p. 2154–2171, 2009. DOI: <https://doi.org/10.1162/jocn.2008.21171>.
- MESTRES-MISSÉ, A.; MÜNTE, T. F.; RODRIGUEZ-FORNELLS, A. Mapping concrete and abstract meanings to new words using verbal contexts. *Second Language Research*, v. 30, n. 2, p. 191–223, 2014. DOI: <https://doi.org/10.1177/0267658313512668>.
- MIKELEIZ, M. I.; KONSTENIUS, A.; EZQUERRO, J. Inferential nets underlying words. *International Journal of Research and Reviews in Applied Sciences*, v. 8, n. 3, 2011. Disponível em: https://www.researchgate.net/publication/302029305_INFERENTIAL_NETS_UNDERLYING_WORDS.
- MIT NEWS. *Serious science at home: new project brings research to families and families to research*. Massachusetts Institute of Technology, 9 jun. 2020. Disponível em: <https://news.mit.edu/2020/serious-science-home-new-project-brings-research-families-and-fa>

milies-research-0609. Acesso em: 25 out. 2025.

MKRTYCHIAN, N.; GNEDYKH, D.; BLAGOVECHTCHENSKI, E.; TSVETOVA, D.; KOSTROMINA, S.; SHTYROV, Y. Contextual acquisition of concrete and abstract words: behavioural and electrophysiological evidence. *Brain Sciences*, Basel, v. 11, n. 7, p. 898, 2021. DOI: <https://doi.org/10.3390/brainsci11070898>.

MOTA, N. B.; COPELLI, M.; RIBEIRO, S. Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. *NPJ Schizophrenia*, v. 3, 2017. DOI: <https://doi.org/10.1038/s41537-017-0016-2>.

MURAKI, E. J.; SIDHU, D. M.; PEXMAN, P. M. Mapping semantic space: property norms and semantic richness. *Cognitive Processing*, [S.l.], v. 21, p. 637–649, 2020. DOI: <https://doi.org/10.1007/s10339-019-00933-y>.

MURPHY, E. Acquiring the impossible: developmental stages of copredication. *Frontiers in Psychology*, [S.l.], v. 8, art. 1072, 2017. DOI: 10.3389/fpsyg.2017.01072.

NELSON, C. A. The Developing Brain. In: SHONKOFF, J. P.; PHILLIPS, D. A. (org.). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington (DC): National Academies Press, 2000. p. 117–160. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK225562/>. Acesso em: 25 out. 2025.

PACE, Amy *et al.* Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, Palo Alto, v. 3, p. 285–308, 2017. DOI: <https://doi.org/10.1146/annurev-linguistics-011516-034226>.

PAIVIO, A. *Mental representations: a dual coding approach*. Oxford: Oxford University Press, 1990.

PATTERSON, K.; NESTOR, P. J.; ROGERS, T. T. Where do you know what you know? The representation of semantic knowledge in the human brain. **Nature reviews neuroscience**, v. 8, n. 12, p. 976–987, 2007.

PATTERSON, K; RALPH, M. A. L. The hub-and-spoke hypothesis of semantic memory. In: HICKOK, Gregory; SMALL, Steven L. (org.). *Neurobiology of Language*. Academic Press, 2016. p. 765–775. DOI: <https://doi.org/10.1016/B978-0-12-407794-2.00061-4>.

PEACOCKE, C. *A study of concepts*. Cambridge, MA: MIT Press, 1992.

PEXMAN, P. M. How does meaning come to mind? Four broad principles of semantic processing. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, [S.l.], v. 74, n. 4, p. 275–283, 2020. DOI: <https://doi.org/10.1037/cep0000235>.

PEXMAN, P. M.; DIVEICA, V.; BINNEY, R. J. *Social semantics: the organization and grounding of abstract concepts*. Philosophical Transactions of the Royal Society B: Biological Sciences, v. 378, n. 1870, p. 20210363, 2022. DOI: <https://doi.org/10.1098/rstb.2021.0363>.

PEYRE, H.; CHARKALUK, M. L.; FORHAN, A.; HEUDE, B.; RAMUS, F.; EDEN MOTHER-CHILD COHORT STUDY GROUP. Do developmental milestones at 4, 8, 12 and 24 months predict IQ at 5-6 years old? Results of the EDEN mother-child cohort. *European Journal of Paediatric Neurology*, London, v. 21, n. 2, p. 272–279, 2017. DOI: <https://doi.org/10.1016/j.ejpn.2016.11.001>.

PHILIPPE, M. D. *Introdução à filosofia de Aristóteles*. São Paulo: Paulus, 2002.

PHILOFSKY, A.; FIDLER, D. J.; HEPBURN, S. Pragmatic language profiles of school-age children with autism spectrum disorders and Williams syndrome. *American Journal of Speech-Language Pathology*, v. 16, n. 4, p. 368–380, 2007. Disponível em: [https://doi.org/10.1044/1058-0360\(2007/040](https://doi.org/10.1044/1058-0360(2007/040).

PICCOLO, L. R.; MERZ, E. C.; HE, X.; SOWELL, E. R.; NOBLE, K. G.; PEDIATRIC IMAGING, NEUROCOGNITION, GENETICS STUDY. Age-related differences in cortical thickness vary by socioeconomic status. *PLoS ONE*, v. 11, n. 9, e0162511, 2016a. DOI: 10.1371/journal.pone.0162511. Disponível em: <https://doi.org/10.1371/journal.pone.0162511>.

PICCOLO, L. D. R.; ARTECHE, A. X.; FONSECA, R. P.; GRASSI-OLIVEIRA, R.; SALLES, J. F. Influence of family socioeconomic status on IQ, language, memory and executive functions of Brazilian children. *Psicologia: Reflexão e Crítica*, v. 29, 2016b. DOI: <https://doi.org/10.1186/s41155-016-0029-8>.

Pulvermüller, F. Semantic embodiment, disembodiment or misembodiment? In search of meaning in modules and neuron circuits. *Brain and Language*, v. 127, n. 1, p. 86–103, 2013. DOI: <https://doi.org/10.1016/j.bandl.2013.05.015>.

Pulvermüller, F. *How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics*. Trends in Cognitive Sciences, v. 17, n. 9, p. 458–470, 2013. DOI: <https://doi.org/10.1016/j.tics.2013.06.004>

Pulvermüller, F. The case of CAUSE: neurobiological mechanisms for grounding an abstract concept. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 373, p. 20170129, 2018. DOI: <http://doi.org/10.1098/rstb.2017.0129>.

RABAGLIATI, H.; FERGUSON, B.; LEW-WILLIAMS, C. The profile of abstract rule learning in infancy: meta-analytic and experimental evidence. *Developmental Science*, v. 22, n. 1, e12704, 2019. DOI: 10.1111/desc.12704.

RABOVSKY, M.; MCRAE, K. Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, v. 132, n. 1, p. 68–89, 2014. DOI: <https://doi.org/10.1016/j.cognition.2014.03.010>.

RADFORD, I. C. Temporal Pattern Association and Child Language Acquisition. 1997. Tese (Doutorado) – The University of Manchester, United Kingdom, 1997. ProQuest Dissertations & Theses, 11004182.

RATHORE, M. A.; ARMSTRONG-CARTER, E.; SIYAL, S.; YOUSAFZAI, A. K.; OBRADOVIĆ, J. Pakistani preschoolers' number of older siblings and cognitive skills:

Moderations by home stimulation and gender. *Journal of Family Psychology*, v. 37, n. 1, p. 132, 2023. DOI: <https://doi.org/10.1037/fam0001018>.

RECCHIA, G.; JONES, M. N. The semantic richness of abstract concepts. *Frontiers in Human Neuroscience*, v. 6, p. 315, 2012. DOI: [10.3389/fnhum.2012.00315](https://doi.org/10.3389/fnhum.2012.00315).

REILLY, J.; PEELLE, J. E.; GARCIA, A.; CRUTCH, S. J. Linking somatic and symbolic representation in semantic memory: the dynamic multilevel reactivation framework. *Psychonomic Bulletin & Review*, [S. l.], v. 23, p. 1002–1014, 2016. DOI: <https://doi.org/10.3758/s13423-015-0824-5>.

REUTER, T.; MAZZEI, C.; LEW-WILLIAMS, C.; EMBERSON, L. Infants' lexical comprehension and lexical anticipation abilities are closely linked in early language development. *Infancy*, v. 28, n. 3, p. 532–549, 2023. DOI: <https://doi.org/10.1111/inf.12534>.

REY, G. Concepts. In: GUTTENPLAN, S. (ed.). *A companion to the philosophy of mind*. Cambridge, MA: Blackwell, 1994. p. 185–193.

RÖDER, B.; STOCK, O.; BIEN, S.; NEVILLE, H.; RÖSLER, F.. Speech processing activates visual cortex in congenitally blind humans. *European Journal of Neuroscience*, [S. l.], v. 16, n. 5, p. 930–936, set. 2002. DOI: <https://doi.org/10.1046/j.1460-9568.2002.02147.x>.

ROMEO, R. R. *et al.* Beyond the 30-million-word gap: children's conversational exposure is associated with language-related brain function. *Psychological Science*, Thousand Oaks, v. 29, n. 5, p. 700–710, maio 2018. DOI: <https://doi.org/10.1177/0956797617742725>.

SABSEVITZ, D. S.; MEDLER, D. A.; SEIDENBERG, M.; BINDER, J. R. Modulation of the semantic system by word imageability. *NeuroImage*, [S. l.], v. 27, n. 1, p. 188–200, 2005. DOI: <https://doi.org/10.1016/j.neuroimage.2005.04.012>.

SAFFRAN, J. R. Words in a sea of sounds: the output of infant statistical learning. *Cognition*, v. 81, n. 2, p. 149–169, 2001. DOI: [https://doi.org/10.1016/S0010-0277\(01\)00132-9](https://doi.org/10.1016/S0010-0277(01)00132-9).

SAFFRAN, J. R.; ASLIN, R. N.; NEWPORT, E. L. Statistical learning by 8-month-old infants. *Science*, v. 274, n. 5294, p. 1926–1928, 13 dez. 1996. DOI: <https://doi.org/10.1126/science.274.5294.1926>.

SAKREIDA, K.; SCOROLLI, C.; MENZ, M. M.; HEIM, S.; BORGHI, A. M.; BINKOFSKI, F. Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition. *Frontiers in Human Neuroscience*, [S. l.], v. 7, p. 125, 2013. DOI: <https://doi.org/10.3389/fnhum.2013.00125>.

SAPIR, E. *A Linguagem: Introdução ao estudo da Fala*. 2. ed. Tradução: Joaquim Mattoso Câmara Jr. São Paulo: Perspectiva, 1980.

SARSOUR, K.; SHERIDAN, M.; JUTTE, D.; NURU-JETER, A.; HINSHAW, S.; BOYCE, W. T. Family socioeconomic status and child executive functions: the roles of language, home environment, and single parenthood. *Journal of the International Neuropsychological Society*,

v. 17, n. 1, p. 120–132, 2011. DOI: 10.1017/S1355617710001335. Disponível em: <https://doi.org/10.1017/S1355617710001335>.

SCLIAR-CABRAL, L. *Introdução à Psicolinguística*. São Paulo: Ática, 1991.

SCOTT, K.; SCHULZ, L. Lookit (Part 1): A New Online Platform for Developmental Research. *Open Mind*, v. 1, n. 1, p. 4–14, 2017. DOI: https://doi.org/10.1162/OPMI_a_00002.

SCOTT, K.; CHU, J.; SCHULZ, L. Lookit (Part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind: Discoveries in Cognitive Science*, v. 1, n. 1, p. 15–29, 2017. DOI: https://doi.org/10.1162/OPMI_a_00001.

SEGALOWITZ, S. J.; CHEVALIER, H. Event-Related Potential (ERP) research in Neurolinguistics: part I: techniques and applications to lexical access. In: STEMMER, B.; WHITAKER, H. A. (ed.). *Handbook of Neurolinguistics*. Academic Press, 1998. p. 95-109. DOI: <https://doi.org/10.1016/B978-012666055-5/50009-5>.

SHADY, M. E. *Infants' sensitivity to function morphemes*. 1996. Dissertação de doutorado não publicada – State University of New York at Buffalo, Buffalo, NY, 1996.

SHAPIRO, K.; CARAMAZZA, A. The representation of grammatical categories in the brain. *Trends in Cognitive Sciences*, v. 7, n. 5, p. 201–206, 2003. DOI: [https://doi.org/10.1016/S1364-6613\(03\)00060-3](https://doi.org/10.1016/S1364-6613(03)00060-3).

SHAPIRO, K. A.; MOO, L. R.; CARAMAZZA, A. Cortical signatures of noun and verb production. *Proceedings of the National Academy of Sciences*, v. 103, n. 5, p. 1644–1649, 2006. DOI: <https://doi.org/10.1073/pnas.0504142103>.

SHEIKH, N. A.; TITONE, D. The embodiment of emotional words in a second language: an eye-movement study. *Cognition and Emotion*, v. 30, n. 3, p. 488–500, 2016. DOI: <https://doi.org/10.1080/02699931.2015.1018144>.

SHI, R.; WERKER, J. F.; MORGAN, J. L. Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, v. 72, n. 2, p. B11–B21, 1999. DOI: [https://doi.org/10.1016/S0010-0277\(99\)00047-5](https://doi.org/10.1016/S0010-0277(99)00047-5).

SHI, R.; WERKER, J. F. Six-month-old infants' preference for lexical words. *Psychological Science*, v. 12, n. 1, p. 70–75, 2001. DOI: <https://doi.org/10.1111/1467-9280.00312>.

SIEW, C. S. Q. Feature distinctiveness effects in language acquisition and lexical processing: insights from megastudies. *Cognitive Processing*, v. 21, n. 4, p. 669–685, 2020. DOI: 10.1007/s10339-019-00947-6.

SIMONSEN, H. G.; LIND, M. *An image is worth a thousand sounds?: on imageability and phonological neighborhood density effects in speech processing*. 2012. Dissertação (Mestrado) – University of Oslo, 2012.

SINGH, L.; CRISTIA, A.; KARASIK, L. B.; RAJENDRA, S. J.; OAKES, L. M. Diversity and representation in infant research: barriers and bridges toward a globalized science of

infant development. *Infancy*, v. 28, n. 4, p. 708–737, 2023. DOI: <https://doi.org/10.1111/infa.12545>.

SINGH, L.; YEUNG, W. J. J.; CHENG, Q.; HENG, E. Y. T. The home literacy environment mediates effects of socio-economic status on infant vocabulary development. *Developmental Science*, v. 26, n. 4, e13349, 2023. DOI: 10.1111/desc.13349. Disponível em: <https://doi.org/10.1111/desc.13349>.

SLIM, M. S.; KANDEL, M.; YACOVONE, A.; SNEDEKER, J. Webcams as windows to the mind? A direct comparison between in-lab and web-based eye-tracking methods. *Open Mind: Discoveries in Cognitive Science*, v. 8, p. 1369–1424, 2024. DOI: https://doi.org/10.1162/opmi_a_00171.

SNEDEKER, J.; GLEITMAN, L. Why it is hard to label our concepts. In: HALL, G.; WAXMAN, S. (ed.). *Weaving a Lexicon*. Cambridge, MA: MIT Press, 2004.

SPERRY, D. E.; SPERRY, L. L.; MILLER, P. J. Reexamining the verbal environments of children from different socioeconomic backgrounds. *Child Development*, Hoboken, v. 90, n. 4, p. 1303–1318, jul. 2019. DOI: <https://doi.org/10.1111/cdev.13072>.

SYVERSEN, G. *Investigating early semantic priming: Event-Related Potentials as a window into the organization of word meaning in children at age 24 months*. 2007. Dissertação (Mestrado em Psicologia) – University of Oslo, Oslo, 2007.

TANENHAUS, M. K.; SPIVEY-KNOWLTON, M. J.; EBERHARD, K. M.; SEDIVY, J. C. Integration of visual and linguistic information in spoken language comprehension. *Science*, v. 268, n. 5217, p. 1632–1634, 1995. DOI: <https://doi.org/10.1126/science.7777863>

TEIXEIRA, E. R. A adaptação dos inventários MacArthur de Desenvolvimento Comunicativo (CDI's) para o português brasileiro. In: *Anais do II Congresso Nacional da ABRALIN*, p. 479–487, 2000.

TENENBAUM, E. J. *et al.* Remote Infant Studies of Early Learning (RISE): scalable online replications of key findings in infant cognitive development. *Developmental Psychology*, Washington, v. 61, n. 1, p. 151–167, jan. 2025. DOI: <https://doi.org/10.1037/dev0001849>.

THILL, S.; TWOMEY, K. E. What's on the inside counts: a grounded account of concept acquisition and development. *Frontiers in Psychology*, v. 7, p. 402, 2016. DOI: <https://doi.org/10.3389/fpsyg.2016.00402>.

TINCOFF, R.; JUSCZYK, P. W. Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, v. 10, n. 2, p. 172–175, 1999. DOI: <https://doi.org/10.1111/1467-9280.00127>.

TOLENTINO, L. C.; TOKOWICZ, N. Are pumpkins better than heaven? An ERP investigation of order effects in the concrete-word advantage. *Brain and Language*, [S.l.], v. 110, n. 1, p. 12–22, 2009. DOI: <https://doi.org/10.1016/j.bandl.2009.01.001>.

TSAI, P. S.; YU, B. H. Y.; LEE, C. Y.; TZENG, O. J. L.; HUNG, D. L.; WU, D. H. An event-related potential study of the concreteness effect between Chinese nouns and verbs. *Brain Research*, [S.l.], v. 1253, p. 149–160, 2009. DOI: <https://doi.org/10.1016/j.brainres.2008.10.080>.

TUCKER-DROB, E. M.; BATES, T. C. Large cross-national differences in gene $\times$ socioeconomic status interaction on intelligence. *Psychological Science*, v. 27, n. 2, p. 138–149, 2016. DOI: <https://doi.org/10.1177/0956797615612727>.

TURKHEIMER, E.; HALEY, A.; WALDRON, M.; D'ONOFRIO, B.; GOTTESMAN, I. Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, v. 14, n. 6, p. 623–628, 2003. DOI: https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x.

URSACHE, A.; NOBLE, K. G.; PEDIATRIC IMAGING, NEUROCOGNITION AND GENETICS STUDY. Socioeconomic status, white matter, and executive function in children. *Brain and Behavior*, v. 6, n. 10, e00531, 2016. DOI: 10.1002/brb3.531. Disponível em: <https://doi.org/10.1002/brb3.531>.

VAN HEUVEN, W. J. B.; MANDERA, P.; KEULEERS, E.; BRYSSBAERT, M. Subtlex-UK: a new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, v. 67, n. 6, p. 1176–1190, jun. 2014. DOI: <https://doi.org/10.1080/17470218.2013.850521>.

VAN HEUGTEN, M.; CHRISTOPHE, A. Infants' acquisition of grammatical gender dependencies. *Infancy*, v. 20, n. 6, p. 675–683, 2015. DOI: 10.1111/inf.12094.

VAN HEUGTEN, M.; SHI, R. French-learning toddlers use gender information on determiners during word recognition. *Developmental Science*, v. 12, n. 3, p. 419–425, 2009. DOI: 10.1111/j.1467-7687.2008.00788.x.

VAN PETTEN, C.; LUKA, B. J. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, v. 97, n. 3, p. 279–293, 2006. DOI: <https://doi.org/10.1016/j.bandl.2005.11.003>.

VEALE, T.; KEANE, M. T. Conceptual scaffolding: a spatially founded meaning representation for metaphor comprehension. *Computational Intelligence*, v. 8, n. 3, p. 494–519, 1992. DOI: <https://doi.org/10.1111/j.1467-8640.1992.tb00377.x>.

VIANA, F. L.; CADIME, I.; LIMA, R. MacArthur-Bates Communicative Development Inventory (MB-CDI). Versão em Português Europeu, 2018. Disponível em: <https://mb-cdi.stanford.edu/adaptations.html>.

VIGLIOCCO, G.; VINSON, D. P. Semantic representation. In: GASKELL, M. G. (ed.). *The Oxford handbook of psycholinguistics*. Oxford: Oxford University Press, 2007. p. 195–215.

WAXMAN, S. R.; BOOTH, A. E. Principles that are invoked in the acquisition of words, but not facts. *Cognition*, Amsterdam, v. 77, n. 2, p. B33–B43, 2000. DOI: [https://doi.org/10.1016/S0010-0277\(00\)00103-7](https://doi.org/10.1016/S0010-0277(00)00103-7).

WAXMAN, S. R. Tudo tinha um nome e de cada nome nascia um novo pensamento: vínculo entre a aprendizagem de palavras e organização conceitual no início da aquisição lexical. In: CORRÊA, L. M. S. (org.). *Aquisição da linguagem e problemas do desenvolvimento linguístico*. Rio de Janeiro: Editora da PUC-RJ, 2006. p. 128–173.

WERKER, J.; HENSCH, T. *Critical periods in speech perception: new directions*. Annual Review of Psychology, v. 66, n. 1, p. 173–196, 2015. DOI: <https://doi.org/10.1146/annurev-psych-010814-015104>.

WEST, W. C.; HOLCOMB, P. J. Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognitive Neuroscience*, v. 12, n. 6, p. 1024–1037, 2000. DOI: <https://doi.org/10.1162/08989290051137558>.

WILLITS, J. A.; WOJCIK, E. H.; SEIDENBERG, M. S.; SAFFRAN, J. R. *Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming*. Infancy, v. 18, n. 6, p. 1053–1075, 2013. DOI: <https://doi.org/10.1111/inf.12026>.

WOJCIK, E. H.; SAFFRAN, J. R. The ontogeny of lexical networks: toddlers encode the relationships among referents when learning novel words. *Psychological Science*, Thousand Oaks, v. 24, n. 10, p. 1898–1905, 2013. DOI: <https://doi.org/10.1177/0956797613478198>.

WORDBANK. *Wordbank: an open repository for developmental vocabulary data*. Disponível em: <https://wordbank.stanford.edu/>.

WUNDT, W. M. *An introduction to psychology*. Londres: G. Allen, Limited, 1912.

ZDRAZILOVA, L. *Understanding the meanings of abstract concepts: keeping it grounded and real*. 2017. Tese (Doutorado) – University of Calgary, Calgary, Canadá, 2017. DOI: <https://doi.org/10.11575/PRISM/28571>.

ZDRAZILOVA, L.; SIDHU, D. M.; PEXMAN, P. M. Communicating abstract meaning: concepts revealed in words and gestures. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 373, n. 1752, p. 20170138, 2018. DOI: 10.1098/rstb.2017.0138.

ZINOVIEFF, F. M. *Interaction of lexical-semantic and imagery representations*. 2000. Tese (Doutorado) – Bangor University, United Kingdom, 2000.

ZENG, N.; AYYUB, M.; SUN, H.; WEN, X.; XIANG, P.; GAO, Z. Effects of physical activity on motor skills and cognitive development in early childhood: a systematic review. *BioMed Research International*, v. 2017, e2760716, 2017. DOI: 10.1155/2017/2760716. Disponível em: <https://doi.org/10.1155/2017/2760716>

APÊNDICE A — Palavras com percentuais de produção a partir de 50% e classificação em relação à concretude do questionário - "Sua criança sabe essas palavras?" e do - "Essa palavra é concreta ou abstrata?"

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
		Compreende	Compreende e diz	Não Compreende e não diz		
2	Avó	4	97	0	0,96	Concreto
3	Bola	5	96	0	0,95	Concreto
4	Água	5	96	0	0,95	Concreto
5	Pé	5	96	0	0,95	Concreto
6	Papai	9	96	2	0,95	Concreto
7	Banana	5	95	1	0,94	Concreto
8	Pão	5	95	1	0,94	Concreto
9	Xixi	5	95	1	0,94	Concreto
10	Mão	7	94	0	0,93	Concreto
11	Neném	4	94	3	0,93	Concreto
12	Ovo	7	93	1	0,92	Concreto
13	Cocô	6	93	2	0,92	Concreto
14	Bumbum	7	93	1	0,92	Concreto
15	Cabeça	8	93	0	0,92	Concreto
16	Gato	8	92	1	0,91	Concreta
17	Vaca	6	92	3	0,91	Concreta
18	Bolo	5	92	4	0,91	Concreto
19	Uva	3	92	6	0,91	Concreto
20	Nariz	8	92	1	0,91	Concreto
21	Olho	9	92	0	0,91	Concreto
22	Avô	5	92	4	0,91	Concreto
23	O próprio nome	8	92	1	0,91	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
24	Carro	8	92	1	0,91	Concreto
25	Pato	5	91	5	0,9	Concreta
26	Peixe (animal)	8	91	2	0,9	Concreta
27	Sapo	7	91	3	0,9	Concreta
28	Dente	9	91	1	0,9	Concreto
29	Barriga	10	91	0	0,9	Concreto
30	Borboleta	9	90	2	0,89	Concreta
31	Cachorro	8	90	3	0,89	Concreta
32	Arroz	7	90	4	0,89	Concreto
33	Batata	4	90	7	0,89	Concreto
34	Maçã	5	90	6	0,89	Concreto
35	Suco	7	90	4	0,89	Concreto
36	Cabelo	10	90	1	0,89	Concreto
37	Sapato	8	90	3	0,89	Concreto
38	Cavalo	9	89	3	0,88	Concreta
39	Galinha (animal)	8	89	4	0,88	Concreta
40	Macaco	9	89	3	0,88	Concreta
41	Bicho	8	89	4	0,88	Concreta
42	Balão	9	89	3	0,88	Concreto
43	Casa	10	89	2	0,88	Concreta
44	Chuva	10	89	2	0,88	Concreto
45	Lua	8	89	4	0,88	Concreto
46	Dedo	10	89	2	0,88	Concreto
47	Orelha	8	89	4	0,88	Concreto
48	Avião	8	89	4	0,88	Concreto
49	Dodói	9	88	4	0,87	Abstrato
50	Leite	4	88	9	0,87	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
51	Colher	11	88	2	0,87	Concreto
52	Árvore	7	88	6	0,87	Concreto
53	Flor	9	88	4	0,87	Concreto
54	Elefante	9	87	5	0,86	Concreta
55	Passarinho	12	87	2	0,86	Concreta
56	Porco	9	87	5	0,86	Concreta
57	Feijão	8	87	6	0,86	Concreto
58	Pipoca	6	87	8	0,86	Concreto
59	Braço	10	87	4	0,86	Concreto
60	Tia	5	87	9	0,86	Concreto
61	Meia	11	87	3	0,86	Concreto
62	Aranha	9	86	6	0,85	Concreta
63	Tartaruga	8	86	7	0,85	Concreta
64	Urso	9	86	6	0,85	Concreta
65	Boi	7	86	8	0,85	Concreta
66	Quente	8	86	7	0,85	Abstrato
67	Livro	10	86	5	0,85	Concreto
68	Sol	9	86	6	0,85	Concreto
69	Mãe	7	86	8	0,85	Concreto
70	Blusa	13	86	2	0,85	Concreto
71	Fralda	15	86	0	0,85	Concreto
72	Sujo	11	85	5	0,84	Abstrato
73	Carne	5	85	11	0,84	Concreto
74	Macarrão	10	85	6	0,84	Concreto
75	Cadeira	14	85	2	0,84	Concreto
76	Cama	14	85	2	0,84	Concreto
77	Bolsa (sacola)	9	85	7	0,84	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
78	Copo	12	85	4	0,84	Concreto
79	Céu	13	85	3	0,84	Concreto
80	Estrela	9	85	7	0,84	Concreto
81	Bolsa (de sair)	10	85	6	0,84	Concreto
82	Calça	13	85	3	0,84	Concreto
83	Roupa	15	85	1	0,84	Concreto
84	Moto	8	85	8	0,84	Concreto
85	Coelho	11	84	6	0,83	Concreta
86	Jacaré	8	84	9	0,83	Concreta
87	Biscoito	8	84	9	0,83	Concreto
88	Laranja (fruta)	5	84	12	0,83	Concreto
89	Escova de dente	13	84	4	0,83	Concreto
90	Lixo	13	84	4	0,83	Concreto
91	Sabão/sabonete	9	84	8	0,83	Concreto
92	Língua	13	84	4	0,83	Concreto
93	Perna	11	84	6	0,83	Concreto
94	Pai	7	84	10	0,83	Concreto
95	Tio	5	84	12	0,83	Concreto
96	Caminhão	11	84	6	0,83	Concreto
97	Formiga	12	83	6	0,82	Concreta
98	Leão	12	83	6	0,82	Concreta
99	Cenoura	8	83	10	0,82	Concreto
100	Papel	8	83	10	0,82	Concreto
101	Bochecha	11	83	7	0,82	Concreto
102	Chinelo	13	83	5	0,82	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
103	Óculos	12	83	6	0,82	Concreto
104	Tubarão	14	82	5	0,81	Concreta
105	Brinquedo	13	82	6	0,81	Concreto
106	Comida	9	82	10	0,81	Concreto
107	Chave	15	82	4	0,81	Concreto
108	Bicicleta	11	82	8	0,81	Concreto
109	Girafa	10	81	10	0,8	Concreta
110	Caneta	14	81	6	0,8	Concreto
111	Chocolate	3	81	17	0,8	Concreto
112	Queijo	7	81	13	0,8	Concreto
113	Porta	10	81	10	0,8	Concreto
114	Telefone	10	81	10	0,8	Concreto
115	Joelho	15	81	5	0,8	Concreto
116	Unha	9	81	11	0,8	Concreto
117	Abelha	11	80	10	0,79	Concreta
118	Rato	11	80	10	0,79	Concreta
119	Morango	15	80	6	0,79	Concreto
120	Prato	10	80	11	0,79	Concreto
121	Remédio	8	80	13	0,79	Concreto
122	Menino	4	80	17	0,79	Concreto
123	Baleia	7	79	15	0,78	Concreta
124	Molhado	9	79	13	0,78	Concreto
125	Boneca	16	79	6	0,78	Concreto
126	Melancia	8	79	14	0,78	Concreto
127	Rua	10	79	12	0,78	Concreto
128	Mesa	14	79	8	0,78	Concreto
129	Quarto	16	79	6	0,78	Concreto
130	Garfo	9	79	13	0,78	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
131	Vassoura	11	79	11	0,78	Concreto
132	Menina	5	79	17	0,78	Concreto
133	Chapéu	12	79	10	0,78	Concreto
134	Acordado	14	78	9	0,77	Abstrato
135	Azul	11	78	12	0,77	Concreto
136	Escuro	6	78	17	0,77	Abstrato
137	Café	7	78	16	0,77	Concreto
138	Milho	8	78	15	0,77	Concreto
139	Frio	13	78	10	0,77	Abstrato
140	Escola	9	78	14	0,77	Concreto
141	Banheiro	16	78	7	0,77	Concreto
142	Sofá	14	78	9	0,77	Concreto
143	Faca	9	78	14	0,77	Concreto
144	Areia	10	78	13	0,77	Concreto
145	Balanço	14	78	9	0,77	Concreto
146	Tigre	14	77	10	0,76	Concreta
147	Bonito	7	77	17	0,76	Abstrato
148	Festa	7	77	17	0,76	Concreto
149	Caixa	11	77	13	0,76	Concreto
150	Piscina	12	77	12	0,76	Concreto
151	Casaco	11	17	13	0,76	Concreto
152	Tênis	15	77	9	0,76	Concreto
153	Ônibus	11	77	13	0,76	Concreto
154	Pizza	4	76	21	0,75	Concreto
155	Janela	15	76	10	0,75	Concreto
156	Escorregador	15	76	10	0,75	Concreto
157	Pedra	11	76	14	0,75	Concreto
158	Peito	11	76	14	0,75	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
159	Amigo	9	76	16	0,75	Concreto
160	Cobra	11	75	15	0,74	Concreta
161	Mosquito	12	75	14	0,74	Concreta
162	Medo	13	75	13	0,74	Abstrato
163	Cozinha	19	75	7	0,74	Concreto
164	Geladeira	16	75	10	0,74	Concreto
165	Sala	14	75	12	0,74	Concreto
166	Balde	14	75	12	0,74	Concreto
167	Amarelo	10	74	17	0,73	Concreto
168	Quebrado	7	74	20	0,73	Concreto
169	Lápis	16	74	11	0,73	Concreto
170	Cobertor/Coberta	10	74	10	0,73	Concreto
171	Rosto	14	74	13	0,73	Concreto
172	Amiga	11	74	16	0,73	Concreto
173	Pijama	8	74	19	0,73	Concreto
174	Verde	10	73	18	0,72	Concreto
175	Vermelho	10	73	18	0,72	Concreto
176	Nuvem	13	73	15	0,72	Concreto
177	Camisa	14	73	14	0,72	Concreto
178	Mamão	12	72	17	0,71	Concreto
179	Peixe (comida)	9	72	20	0,71	Concreto
180	Chuveiro	20	72	9	0,71	Concreto
181	Escada	18	72	11	0,71	Concreto
182	Tv	18	72	11	0,71	Concreto
183	Short	12	72	17	0,71	Concreto
184	Grande	5	71	25	0,7	Abstrato
185	Sorvete	5	71	25	0,7	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
186	Parque	9	71	21	0,7	Concreto
187	Pente	11	71	19	0,7	Concreto
188	Buraco	13	71	17	0,7	Concreto
189	Barco	12	71	18	0,7	Concreto
190	Caranguejo	11	70	20	0,69	Concreta
191	Galo	15	70	16	0,69	Concreta
192	Laranja (cor)	11	70	20	0,69	Concreto
193	Pesado	6	70	25	0,69	Abstrato
194	Massa	9	70	22	0,69	Concreto
195	Fome	18	70	13	0,69	Abstrato
196	Praia	5	70	26	0,69	Concreto
197	Fogão	16	70	15	0,69	Concreto
198	Travesseiro	15	70	16	0,69	Concreto
199	Costas	12	70	19	0,69	Concreto
200	Trem	7	70	24	0,69	Concreto
201	Garrafa	16	69	16	0,68	Concreto
202	Branco	13	68	20	0,67	Concreto
203	Lindo	11	68	22	0,67	Abstrato
204	Preto	9	68	24	0,67	Concreto
205	Trabalho	10	68	23	0,67	Concreto
206	Lobo	12	67	22	0,66	Concreta
207	Barata	14	67	20	0,66	Concreta
208	Limpo	10	67	24	0,66	Abstrato
209	Feliz	21	67	13	0,66	Abstrato
210	Pia	14	67	20	0,66	Concreto
211	Retrato/foto	14	67	20	0,66	Concreto
212	Ventilador	13	67	21	0,66	Concreto
213	Zebra	16	66	19	0,65	Concreta

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
214	Iogurte	4	66	31	0,65	Concreto
215	Calor	21	66	14	0,65	Abstrato
216	Mercado	12	66	23	0,65	Concreto
217	Relógio	15	66	20	0,65	Concreto
218	Planta (decoração)	13	66	22	0,65	Concreto
219	Bota	13	66	22	0,65	Concreto
220	Guarda-chuva	17	66	18	0,65	Concreto
221	Triste	19	65	17	0,64	Abstrato
222	Escova de cabelo	16	65	20	0,64	Concreto
223	Boné	13	65	23	0,64	Concreto
224	Pequeno	11	64	26	0,63	Abstrato
225	Banheira	18	64	19	0,63	Concreto
226	Almofada	14	64	23	0,63	Concreto
227	Planta (vegetação)	15	64	22	0,63	Concreto
228	Sacola	10	64	27	0,63	Concreto
229	Tesoura	17	64	20	0,63	Concreto
230	Ouvido	18	64	19	0,63	Concreto
231	Sandália	12	64	25	0,63	Concreto
232	Doce	5	63	33	0,62	Concreto
233	Piru/Pipiu (pênis)	10	63	28	0,62	Concreto
234	Hipopótamo	17	62	22	0,61	Concreta
235	Onça	20	62	19	0,61	Concreta
236	Picolé	4	62	35	0,61	Concreto
237	Mamadeira	13	62	26	0,61	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
238	Bico/Chupeta	11	62	28	0,61	Concreto
239	Pinguim	15	61	25	0,6	Concreta
240	Forte	10	61	30	0,6	Abstrato
241	Médico	18	61	22	0,6	Concreto
242	Banco	16	61	24	0,6	Concreto
243	Queixo	15	61	25	0,6	Concreto
244	Armário	20	60	21	0,59	Concreto
245	Pulseira	8	60	33	0,59	Concreto
246	Vestido	15	60	26	0,59	Concreto
247	Duro	3	59	39	0,58	Abstrato
248	Rápido	14	59	28	0,58	Abstrato
249	Gelo	8	59	34	0,58	Concreto
250	Manteiga	14	59	28	0,58	Concreto
251	Computador	19	59	23	0,58	Concreto
252	Ombro	19	59	23	0,58	Concreto
253	Cueca	10	59	32	0,58	Concreto
254	Lagartixa	17	58	26	0,57	Concreta
255	Bom	6	58	37	0,57	Abstrato
256	Cola	9	58	34	0,57	Concreto
257	Sopa	12	58	31	0,57	Concreto
258	Gaveta	20	58	23	0,57	Concreto
259	Elevador	15	58	28	0,57	Concreto
260	Perereca/Pepeca (vagina)	14	58	29	0,57	Concreto
261	Calcinha	11	58	32	0,57	Concreto
262	Galinha (comida)	6	58	37	0,57	Concreto
263	Alto	11	57	33	0,56	Abstrato
264	Cansado	11	57	33	0,56	Abstrato

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
265	Rasgado	6	57	38	0,56	Abstrato/Concreto =
266	Praça/pracinha	15	57	29	0,56	Concreto
267	Dinda	6	57	38	0,56	Concreto
268	Carrinho de bebê	20	57	24	0,56	Concreto
269	Roxo	9	56	36	0,55	Concreto
270	Vazio	16	56	29	0,55	Abstrato
271	Bravo	19	56	26	0,55	Abstrato
272	Professor/a	7	56	38	0,55	Concreto
273	Shopping	6	56	39	0,55	Concreto
274	Helicóptero	13	56	32	0,55	Concreto
275	Devagar	17	55	29	0,54	Abstrato
276	Grama	20	55	26	0,54	Concreto
277	Pá	10	55	36	0,54	Concreto
278	Camiseta	17	55	29	0,54	Concreto
279	Seco	12	54	35	0,53	Abstrato
280	Berço	20	54	27	0,53	Concreto
281	Lixeira	21	54	26	0,53	Concreto
282	Cinto	11	54	36	0,53	Concreto
283	Dragão	3	54	44	0,53	Concreto
284	Pirulito	5	53	43	0,52	Concreto
285	Palhaço	8	53	40	0,52	Concreto
286	Polícia	8	53	40	0,52	Concreto
287	Cheio	15	52	34	0,51	Abstrato
288	Prima	11	52	38	0,51	Concreto
289	Babador	13	52	36	0,51	Concreto
290	Homem-aranha	5	52	44	0,51	Concreto

N de palavras	Palavras	N de respostas totais (N= 110) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
291	Novo	6	50	45	0,50	Abstrato
292	Bala	6	51	44	0,50	Concreto
293	Torrada	15	51	35	0,50	Concreto
294	Floresta	8	50	43	0,50	Concreto
295	Hospital	11	50	40	0,50	Concreto
296	Torneira	22	50	29	0,50	Concreto
297	Cara	13	51	37	0,50	Concreto
298	Dindo	8	50	43	0,50	Concreto
299	Primo	11	50	40	0,50	Concreto
300	Botão	16	50	35	0,50	Concreto

Fonte: elaboração própria (2024).

APÊNDICE B — Palavras com percentuais de produção a partir de 50% e classificação em relação à concretude do questionário - *Quais palavras sua criança já aprendeu?* e do - *Essa palavra é concreta ou abstrata?*

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
		Compreende	Compreende e diz	Não Compreende e não diz		
1	Cachorro	5	51	0	1,00	Concreta
2	Gato	5	51	0	1,00	Concreta
3	Passarinho	7	49	0	1,00	Concreta
4	Pé	4	52	0	1,00	Concreta
5	Abraço	9	47	0	1,00	Concreta
6	Água	2	54	0	1,00	Concreta
7	Banana	1	55	0	1,00	Concreta
8	Biscoito	3	53	0	1,00	Concreta
9	Mamãe	1	54	1	0,98	Concreta
10	Bola	2	53	1	0,98	Concreta
11	Batata	2	53	1	0,98	Concreta
12	Pão	2	53	1	0,98	Concreta
13	Neném	2	53	1	0,98	Concreta
14	Cocô	3	52	1	0,98	Concreta
15	Mão	3	52	1	0,98	Concreta
16	Nariz	3	52	1	0,98	Concreta
17	Avó	3	52	1	0,98	Concreta
18	Carro	3	52	1	0,98	Concreta
19	Peixe	5	50	1	0,98	Concreta
20	Boca	6	49	1	0,98	Concreta
21	Beijo	6	49	1	0,98	Concreta
22	Casa	6	49	1	0,98	Concreta

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
23	Colher	7	48	1	0,98	Concreta
24	Brinquedo	8	47	1	0,98	Concreta
25	Braço	8	47	1	0,98	Concreta
26	Copo	9	46	1	0,98	Concreta
27	Papai	2	52	2	0,96	Concreta
28	Uva	3	51	2	0,96	Concreta
29	Barriga	4	50	2	0,96	Concreta
30	Arroz	4	50	2	0,96	Concreta
31	Leão	5	49	2	0,96	Concreta
32	Cabeça	5	49	2	0,96	Concreta
33	Olho	5	49	2	0,96	Concreta
34	Xixi	5	49	2	0,96	Concreta
35	Céu	5	49	2	0,96	Concreta
36	Vaca	6	48	2	0,96	Concreta
37	Fralda	6	48	2	0,96	Concreta
38	Bumbum	6	48	2	0,96	Concreta
39	Dente	6	48	2	0,96	Concreta
40	Orelha	6	48	2	0,96	Concreta
41	Laranja	6	48	2	0,96	Concreta
42	Chuva	6	48	2	0,96	Concreta
43	Galinha	7	47	2	0,96	Concreta
44	Boneca	7	47	2	0,96	Concreta
45	Cabelo	7	47	2	0,96	Concreta
46	Aranha	8	46	2	0,96	Concreta
47	Formiga	8	46	2	0,96	Concreta
48	Cadeira	8	46	2	0,96	Concreta
49	Escova de dente	9	45	2	0,96	Concreta
50	Banheiro	10	44	2	0,96	Concreta

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
51	Avô	1	52	3	0,95	Concreta
52	Feijão	3	50	3	0,94	Concreta
53	Piscina	3	50	3	0,94	Concreta
54	O próprio nome	4	49	3	0,94	Concreta
55	Dodói	4	49	3	0,94	Abstrata
56	Cavalo	6	47	3	0,94	Concreta
57	Dedo	6	47	3	0,94	Concreta
58	Maçã	6	47	3	0,94	Concreta
59	Livro	7	46	3	0,94	Concreta
60	Blusa	7	46	3	0,94	Concreta
61	Flor	7	46	3	0,94	Concreta
62	Bolsa	9	44	3	0,94	Concreta
63	Chave	11	42	3	0,93	Concreta
64	Quarto	12	41	3	0,93	Concreta
65	Tio	1	51	4	0,93	Concreta
66	Ovo	2	50	4	0,93	Concreta
67	Macaco	3	49	4	0,92	Concreta
68	Bolo	3	49	4	0,92	Concreta
69	Suco	3	49	4	0,92	Concreta
70	Árvore	4	48	4	0,92	Concreta
71	Balão	5	47	4	0,92	Concreta
72	Sapato	5	47	4	0,92	Concreta
73	Lua	5	47	4	0,92	Concreta
74	Lixo	5	47	4	0,92	Concreta
75	Borboleta	6	46	4	0,92	Concreta
76	Sapo	6	46	4	0,92	Concreta
77	Bolsa	6	46	4	0,92	Concreta
78	Meia	6	46	4	0,92	Concreta
79	Comida	6	46	4	0,92	Concreta

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
80	Porta	7	45	4	0,92	Concreta
81	Roupa	8	44	4	0,92	Concreta
82	Cama	8	44	4	0,92	Concreta
83	Sofá	10	42	4	0,91	Concreta
84	Tv	11	41	4	0,91	Concreta
85	Cozinha	12	40	4	0,91	Concreta
86	Geladeira	12	40	4	0,91	Concreta
87	Pato	3	48	5	0,91	Concreta
88	Porco	3	48	5	0,91	Concreta
89	Papel	3	48	5	0,91	Concreta
90	Elefante	4	47	5	0,90	Concreta
91	Bicho	5	46	5	0,90	Concreta
92	Calça	5	46	5	0,90	Concreta
93	Óculos	5	46	5	0,90	Concreta
94	Mosquito	6	45	5	0,90	Concreta
95	Chinelo	6	45	5	0,90	Concreta
96	Machucado	6	45	5	0,90	Concreta
97	Jacaré	7	44	5	0,90	Concreta
98	Língua	7	44	5	0,90	Concreta
99	Perna	7	44	5	0,90	Concreta
100	Sabão/sabonete	7	44	5	0,90	Concreta
101	Bicicleta	7	44	5	0,90	Concreta
102	Coelho	8	43	5	0,90	Concreta
103	Rua	8	43	5	0,90	Concreta
104	Bochecha	9	42	5	0,89	Concreta
105	Azul	0	50	6	0,89	Concreta
106	Choro	10	41	5	0,89	Concreta
107	Cobertor/Cober ta	10	41	5	0,89	Concreta

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
108	Telefone	10	41	5	0,89	Concreta
109	Tia	1	49	6	0,89	Concreta
110	Caneta	11	40	5	0,89	Concreta
111	Macarrão	2	48	6	0,89	Concreta
112	Boi	3	47	6	0,89	Concreta
113	Chuveiro	12	39	5	0,89	Concreta
114	Mãe	4	46	6	0,88	Concreta
115	Pai	4	46	6	0,88	Concreta
116	Sala	13	38	5	0,88	Concreta
117	Estrela	5	45	6	0,88	Concreta
118	Moto	5	45	6	0,88	Concreta
119	Tartaruga	6	44	6	0,88	Concreta
120	Unha	6	44	6	0,88	Concreta
121	Remédio	6	44	6	0,88	Concreta
122	Garrafa	7	43	6	0,88	Concreta
123	Vassoura	8	42	6	0,88	Concreta
124	Sorriso	9	41	6	0,87	Concreta
125	Prato	9	41	6	0,87	Concreta
126	Mesa	9	41	6	0,87	Concreta
127	Tênis	10	40	6	0,87	Concreta
128	Caminhão	4	45	7	0,87	Concreta
129	Camisa	12	38	6	0,86	Concreta
130	Carinho	12	38	6	0,86	Abstrata
131	Chapéu	5	44	7	0,86	Concreta
132	Rosto	5	44	7	0,86	Concreta
133	Pipoca	5	44	7	0,86	Concreta
134	Peito	6	43	7	0,86	Concreta
135	Cenoura	6	43	7	0,86	Concreta
136	Avião	6	43	7	0,86	Concreta

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
137	Abelha	7	42	7	0,86	Concreta
138	Baleia	7	42	7	0,86	Concreta
139	Dor	7	42	7	0,86	Abstrata
140	Acordado	7	42	7	0,86	Abstrata
141	Ônibus	7	42	7	0,86	Concreta
142	Carne	1	47	8	0,85	Concreta
143	Tubarão	4	44	8	0,85	Concreta
144	Urso	4	44	8	0,85	Concreta
145	Pizza	4	44	8	0,85	Concreta
146	Girafa	5	43	8	0,84	Concreta
147	Sujo	5	43	8	0,84	Abstrata
148	Garfo	6	42	8	0,84	Concreta
149	Quente	6	42	8	0,84	Abstrata
150	Janela	6	42	8	0,84	Concreta
151	Fogão	13	36	7	0,84	Concreta
152	Planta	7	41	8	0,84	Concreta
153	Caixa	7	41	8	0,84	Concreta
154	Amarelo	1	46	9	0,84	Concreta
155	Café	2	45	9	0,83	Concreta
156	Chocolate	3	44	9	0,83	Concreta
157	Joelho	9	39	8	0,83	Concreta
158	Escorregador	10	38	8	0,83	Concreta
159	Barata	6	41	9	0,82	Concreta
160	Casaco	6	41	9	0,82	Concreta
161	Morango	6	41	9	0,82	Concreta
162	Cobra	12	36	8	0,82	Concreta
163	Short	7	40	9	0,82	Concreta
164	Travesseiro	7	40	9	0,82	Concreta

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
165	Lápis	8	39	9	0,81	Concreta
166	Pente	8	39	9	0,81	Concreta
167	Escada	8	39	9	0,81	Concreta
168	Frio	3	43	10	0,81	Abstrata
169	Queijo	4	42	10	0,81	Concreta
170	Escova de cabelo	10	37	9	0,80	Concreta
171	Calor	5	41	10	0,80	Abstrata
172	Verde	0	45	11	0,80	Concreta
173	Praça/pracinha	6	40	10	0,80	Concreta
174	Faca	6	40	10	0,80	Concreta
175	Gaveta	16	32	8	0,80	Concreta
176	Balanço	7	39	10	0,80	Concreta
177	Mamadeira	7	39	10	0,80	Concreta
178	Banheira	12	35	9	0,80	Concreta
179	Praia	4	41	11	0,79	Concreta
180	Molhado	4	41	11	0,79	Concreta
181	Laranja	0	44	12	0,79	Concreta
182	Rato	5	40	11	0,78	Concreta
183	Pedra	5	40	11	0,78	Concreta
184	Vermelho	1	43	12	0,78	Concreta
185	Galo	6	39	11	0,78	Concreta
186	Sorvete	6	39	11	0,78	Concreta
187	Balde	11	35	10	0,78	Concreta
188	Branco	2	42	12	0,78	Concreta
189	Bico/Chupeta	7	38	11	0,78	Concreta
190	Planta	7	38	11	0,78	Concreta
191	Sujeira	7	38	11	0,78	Concreta
192	Medo	3	41	12	0,77	Abstrata

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
193	Melancia	3	41	12	0,77	Concreta
194	Parque	3	41	12	0,77	Concreta
195	Pia	12	34	10	0,77	Concreta
196	Tigre	8	37	11	0,77	Concreta
197	Fome	4	40	12	0,77	Abstrata
198	Iogurte	4	40	12	0,77	Concreta
199	Retrato/foto	4	40	12	0,77	Concreta
200	Buraco	5	39	12	0,76	Concreta
201	Escuro	5	39	12	0,76	Abstrata
202	Leite	2	41	13	0,76	Concreta
203	Onça	7	37	12	0,76	Concreta
204	Vestido	7	37	12	0,76	Concreta
205	Sandália	7	37	12	0,76	Concreta
206	Costas	7	37	12	0,76	Concreta
207	Zebra	8	36	12	0,75	Concreta
208	Picolé	4	39	13	0,75	Concreta
209	Quebrado	4	39	13	0,75	Concreta
210	Barco	8	36	12	0,75	Concreta
211	Berço	8	36	12	0,75	Concreta
212	Dinda	1	41	14	0,75	Concreta
213	Amor	5	38	13	0,75	Abstrata
214	Milho	5	38	13	0,75	Concreta
215	Nuvem	5	38	13	0,75	Concreta
216	Fogo	2	40	14	0,74	Concreta
217	Lindo	6	37	13	0,74	Abstrata
218	Menina	3	39	14	0,74	Concreta
219	Menino	3	39	14	0,74	Concreta
220	Sacola	7	36	13	0,73	Concreta
221	Mercado	4	38	14	0,73	Concreta

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
222	Trem	4	38	14	0,73	Concreta
223	Ouvindo	9	34	13	0,72	Concreta
224	Amiguinho	9	34	13	0,72	Concreta
225	Preto	2	39	15	0,72	Concreta
226	Torneira	13	31	12	0,72	Concreta
227	Pijama	6	36	14	0,72	Concreta
228	Festa	6	36	14	0,72	Abstrata
229	Relógio	6	36	14	0,72	Concreta
230	Armário	17	28	11	0,72	Concreta
231	Escola	3	38	15	0,72	Concreta
232	Bonito	3	38	15	0,72	Abstrata
233	Lixeira	7	35	14	0,71	Concreta
234	Tesoura	7	35	14	0,71	Concreta
235	Bota	4	37	15	0,71	Concreta
236	Grande	4	37	15	0,71	Abstrata
237	Boné	8	34	14	0,71	Concreta
238	Piru/Pipiu (pênis)	8	34	14	0,71	Concreta
239	Amiguinha	8	34	14	0,71	Concreta
240	Pequeno	5	36	15	0,71	Concreta
241	Areia	7	34	15	0,69	Concreta
242	Trabalho	4	36	16	0,69	Concreta
243	Dindo	1	38	17	0,69	Concreta
244	Caranguejo	14	29	13	0,69	Concreta
245	Peixe	2	37	17	0,69	Concreta
246	Guarda-chuva	3	36	17	0,68	Concreta
247	Roxo	3	36	17	0,68	Concreta
248	Carrinho de bebê	13	29	14	0,67	Concreta

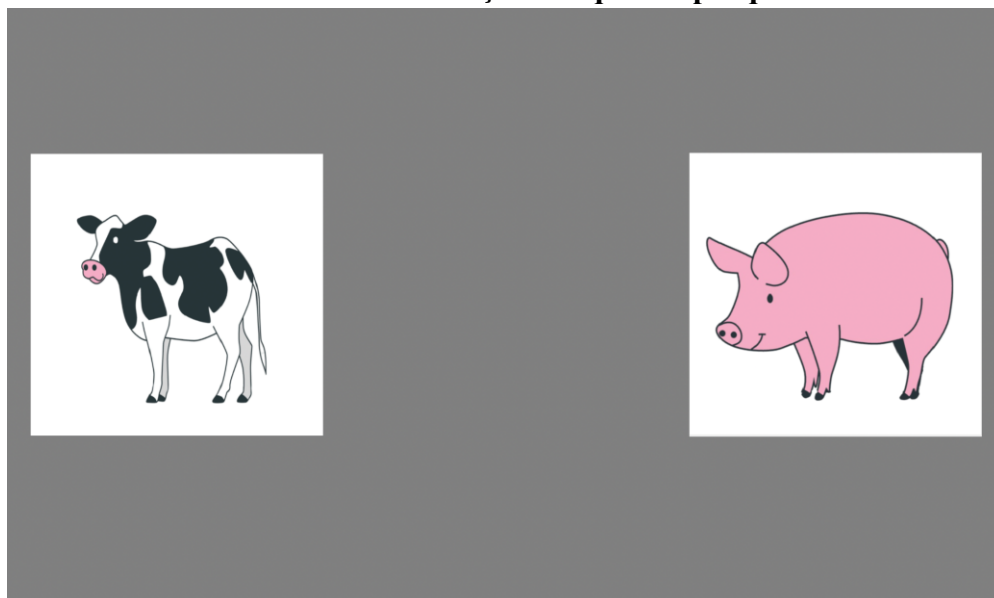
N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
249	Helicóptero	10	31	15	0,67	Concreta
250	Queixo	7	33	16	0,67	Concreta
251	Gelo	2	36	18	0,67	Concreta
252	Amiga	8	32	16	0,67	Concreta
253	Amigo	8	32	16	0,67	Concreta
254	Perereca/Pepeca (vagina)	3	35	18	0,66	Concreta
255	Massa	6	33	17	0,66	Concreta
256	Lobo	4	34	18	0,65	Concreta
257	Ombro	4	34	18	0,65	Concreta
258	Pesado	4	34	18	0,65	Abstrata
259	Shopping	7	32	17	0,65	Concreta
260	Gramma	13	28	15	0,65	Concreta
261	Doce	5	33	18	0,65	Concreta
262	Galinha	5	33	18	0,65	Concreta
263	Bom	5	33	18	0,65	Abstrata
264	Calma	8	31	17	0,65	Abstrata
265	Médico	8	31	17	0,65	Concreta
266	Mamão	3	34	19	0,64	Concreta
267	Pulseira	4	33	19	0,63	Concreta
268	Bala	7	31	18	0,63	Concreta
269	Saia	5	32	19	0,63	Concreta
270	Pinguim	8	30	18	0,63	Concreta
271	Manteiga	3	33	20	0,62	Concreta
272	Botão	6	31	19	0,62	Concreta
273	Seco	6	31	19	0,62	Concreta
274	Alto	4	32	20	0,62	Concreta
275	Limpo	4	32	20	0,62	Abstrata
276	Pirulito	7	30	19	0,61	Concreta

N de palavras	Palavras	N de respostas totais (N=56) e seus níveis de compreensão			Tx. proporção de respostas - CD	Classificação - Concretude (adultos)
277	Primo	7	30	19	0,61	Concreta
278	Alegria	10	28	18	0,61	Abstrata
279	Lagartixa	5	30	21	0,59	Concreta
280	Sopa	5	30	21	0,59	Concreta
281	Polícia	3	31	22	0,58	Concreta
282	Cola	6	29	21	0,58	Concreta
283	Vazio	4	30	22	0,58	Abstrata
284	Duro	2	31	23	0,57	Abstrata
285	Prima	6	28	22	0,56	Concreta
286	Pá	4	29	23	0,56	Concreta
287	Rápido	5	28	23	0,55	Abstrata
288	Calcinha	4	28	24	0,54	Concreta

Fonte: elaboração própria (2024).

APÊNDICE C — Imagens dos trials experimentais

Trial 02 da familiarização - vaquinha-porquinho



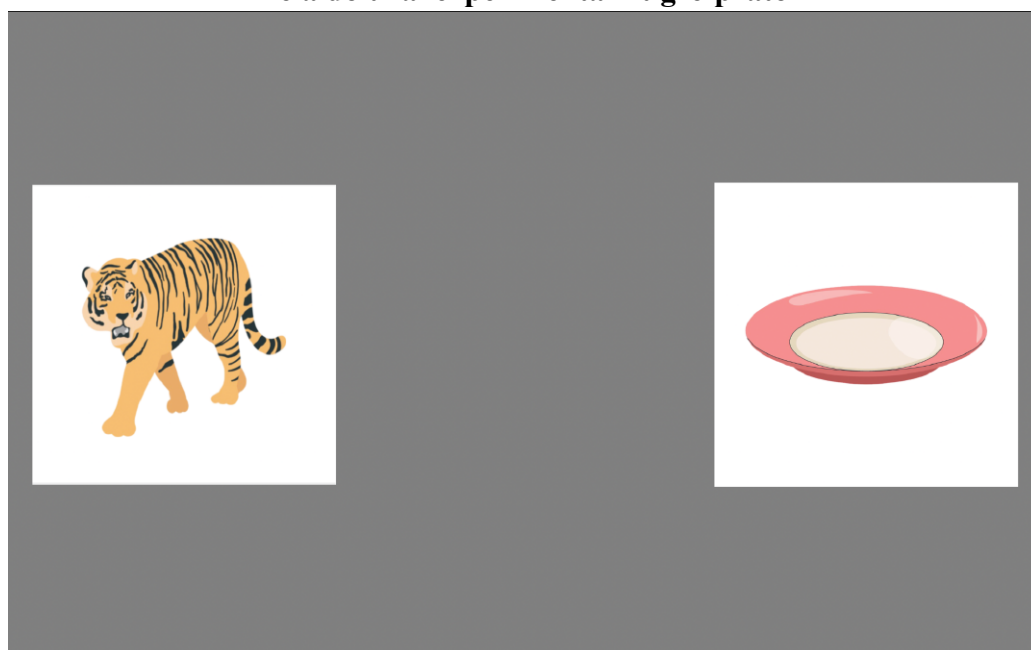
Fonte: elaboração própria (2024).

Tela do trial experimental - fogão-balão



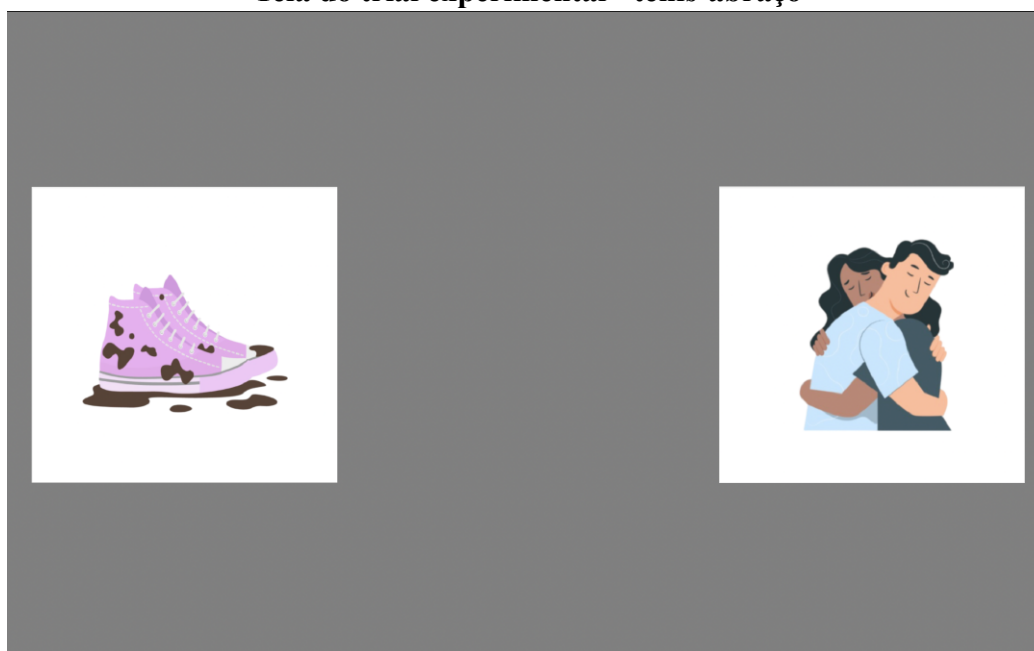
Fonte: elaboração própria (2024).

Tela do trial experimental - tigre-prato



Fonte: elaboração própria (2024).

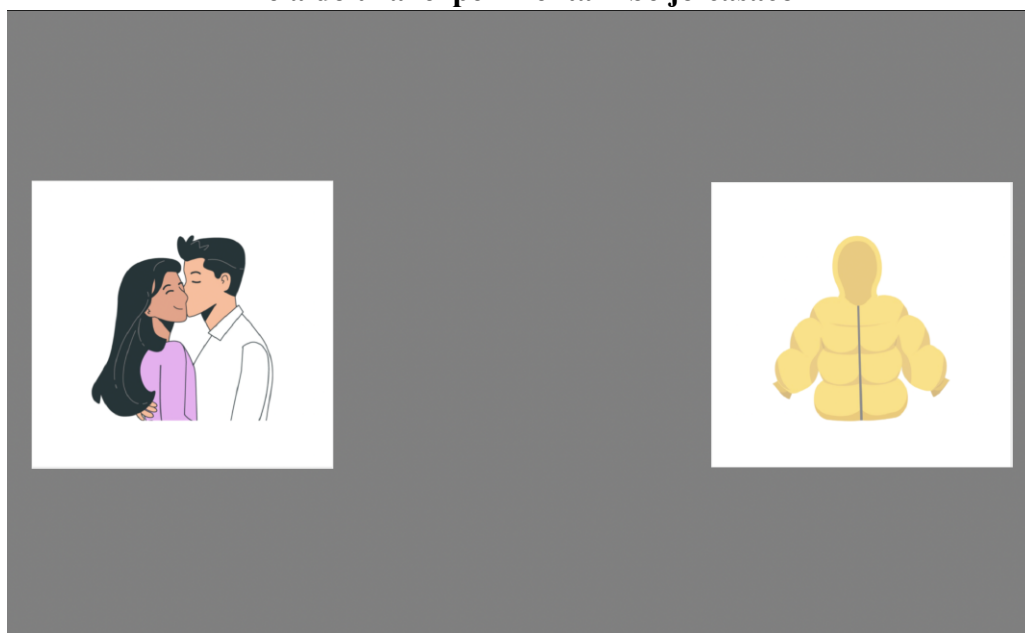
Tela do trial experimental - tênis-abraço



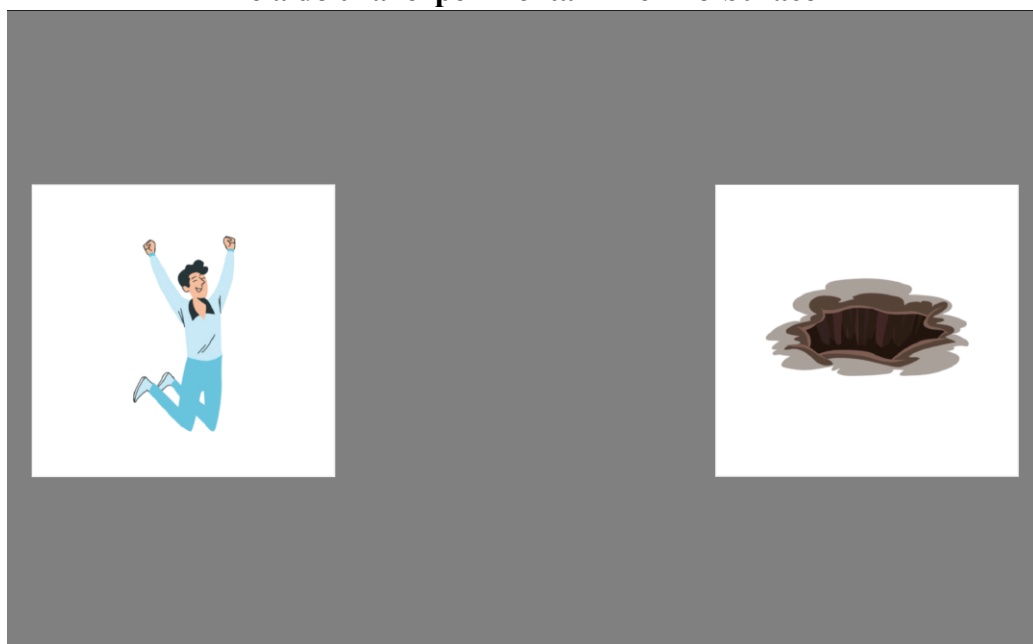
Fonte: elaboração própria (2024).

Tela do trial experimental - pijama-pizza

Fonte: elaboração própria (2024).

Tela do trial experimental - beijo-casaco

Fonte: elaboração própria (2024).

Tela do trial experimental - menino-buraco

Fonte: elaboração própria (2024).

APÊNDICE D — Randomização da apresentação dos estímulos experimentais

	L1	L2	L3	L4	L5	L6	L7	L8
Trial Fogão - Balão	E - fogão	D - quente	E - balão	D - festa	D- fogão	E - quente	D - balão	E -festa
Trial Pizza - Pijama	D - fome	E - pizza	D - sono	E - pijama	E - fome	D - pizza	E - sono	D - pijama
Trial Casaco - Beijo	E - casaco	D -amor	E - beijo	D - frio	D - casaco	E - amor	D - beijo	E -frio
Trial Tigre - Prato	D - medo	E - tigre	D - vazio	E - prato	E - medo	D - tigre	E - vazio	D - prato
Trial Tênis - Abraço	E - tênis	D - sujo	E- abraço	D - carinho	D - tênis	E- sujo	D - abraço	E- carinho
Trial Buraco - Menino	D - escuro	E - buraco	D - alegria	E - menino	E - escuro	D - buraco	E - alegria	D - menino

Fonte: elaboração própria (2024).
Cor Rosa - palavra concreta
(E) - imagem apresentada à esquerda
Azul - palavra abstrata
(D) - imagem apresentada à direita

APÊNDICE E — Frases do testes psicolinguísticos (adulto e criança)

Lista de itens	Concreto	Abstrato
Balão-fogão	"Olha! Fogão! Viu! Fogão!" "Olha! Balão! Viu! Balão!"	"Olha! Quente! Viu! Quente" "Olha! Festa! Viu! Festa!"
Prato-tigre	"Olha! Prato! Viu! Prato!" "Olha! Tigre! Viu! Tigre!"	"Olha! Vazio! Viu! Vazio!" "Olha! Medo! Viu! Medo!"
Tênis-abraço	"Olha! Tênis! Viu! Tênis!" "Olha! Abraço! Viu! Abraço!"	"Olha! Sujo! Viu! Sujo!" "Olha! Carinho! Viu! Carinho!"
Pijama-pizza	"Olha! Pijama! Viu! Pijama!" "Olha! Pizza! Viu! Pizza!"	"Olha! Sono! Viu! Sono!" "Olha! Fome! Viu! Fome!"
Beijo-casaco	"Olha! Beijo! Viu! Beijo!" "Olha! Casaco! Viu! Casaco!"	"Olha! Amor! Viu! Amor!" "Olha! Frio! Viu! Frio!"
Buraco-menino	"Olha! Menino! Viu! Menino!" "Olha! Buraco! Viu! Buraco!"	"Olha! Alegria! Viu! Alegria!" "Olha! Escuro! Viu! Escuro!"

Fonte: elaboração própria (2024).

APÊNDICE F — Análise do onset/duração das palavras nas sentenças dos testes psicolinguísticos (adulto e criança)

audio_file_name	condition	target	Olha_onset	Olha_duration	Target_1_onset	Target_1_duration	Viu_onset	Viu_duration	Target_2_onset	Target_2_duration
abraco-tenis_concrete_abraço.wav	concrete	abraço	111ms	903ms	1512ms	1295ms	3863ms	435ms	4837ms	867ms
abraco-tenis_concrete_tenis.wav	concrete	tenis	97ms	714ms	1417ms	896ms	3311ms	446ms	4361ms	740ms
abraco-tenis_abstract_sujo.wav	abstract	sujo	99ms	590ms	1241ms	773ms	3019ms	495ms	4031ms	971ms
abraco-tenis_abstract_carinho.wav	abstract	carinho	100ms	797ms	1386ms	1139ms	3520ms	467ms	4429ms	1087ms
balao-fogao_concrete_balao.wav	concrete	balao	95ms	635ms	1476ms	677ms	3166ms	474ms	4283ms	619ms
balao-fogao_concrete_fogao.wav	concrete	fogao	100ms	621ms	1190ms	763ms	3149ms	479ms	4170ms	744ms
balao-fogao_abstract_festa.wav	abstract	festa	112ms	736ms	1146ms	1063ms	3219ms	449ms	4039ms	1001ms
balao-fogao_abstract_quente.wav	abstract	quente	95ms	674ms	1317ms	790ms	3107ms	500ms	4149ms	861ms
beijo-casaco_concrete_beijo.wav	concrete	beijo	99ms	790ms	1393ms	945ms	3463ms	441ms	4405ms	883ms
beijo-casaco_concrete_casaco.wav	concrete	casaco	86ms	583ms	1157ms	970ms	3161ms	474ms	4099ms	967ms
beijo-casaco_abstract_amor.wav	abstract	amor	201ms	669ms	1453ms	602ms	3230ms	437ms	4126ms	620ms
beijo-casaco_abstract_frio.wav	abstract	frio	98ms	861ms	1158ms	1100ms	3281ms	505ms	4047ms	989ms
buraco-menino_concrete_buraco.wav	concrete	buraco	95ms	609ms	1190ms	864ms	3067ms	474ms	4034ms	986ms

buraco-menino_concrete_menino.wav	concrete	menino	187ms	855ms	1528ms	1160ms	3188ms	466ms	4113ms	1035ms
buraco-menino_abstract_escuro.wav	abstract	escuro	101ms	654ms	1165ms	1031ms	3166ms	431ms	4021ms	1066ms
buraco-menino_abstract_alegria.wav	abstract	alegria	9ms	745ms	1108ms	1083ms	3191ms	433ms	3838ms	1332ms
pizza-pijama_concrete_pizza.wav	concrete	pizza	19ms	768ms	1501ms	647ms	3141ms	486ms	4149ms	674ms
pizza-pijama_concrete_pijama.wav	concrete	pijama	178ms	868ms	1522ms	1116ms	3586ms	421ms	4480ms	829ms
pizza-pijama_abstract_fome.wav	abstract	fome	9ms	592ms	1390ms	753ms	3155ms	501ms	4161ms	823ms
pizza-pijama_abstract_sono.wav	abstract	sono	9ms	878ms	1315ms	1077ms	3423ms	389ms	4139ms	917ms
tigre-prato_concrete_tigre.wav	concrete	tigre	9ms	685ms	1320ms	871ms	3185ms	460ms	4142ms	923ms
tigre-prato_concrete_prato.wav	concrete	prato	9ms	794ms	1426ms	580ms	3301ms	332ms	4151ms	816ms
tigre-prato_abstract_medo.wav	abstract	medo	19ms	635ms	1358ms	727ms	3128ms	511ms	4124ms	878ms
tigre-prato_abstract_vazio.wav	abstract	vazio	8ms	644ms	1195ms	1021ms	3202ms	450ms	4071ms	995ms

Fonte: elaboração própria (2025).

APÊNDICE G — Parecer favorável do Comitê de Ética



CER U-Paris Cité (Comité d'Éthique de la Recherche)

Président : Christophe Blaison (à partir de septembre 2024)
Présidente par intérim Jacqueline Fagard : (jacqueline.fagard@u-paris.fr)
Vice-présidente : Anne-Caroline Fiévet (anne-caroline.fievet@u-paris.fr)

N° 2024-62-de
CARVALHO-
BOGO

PROTOCOLE : How do toddlers interpret concrete and abstract words?

Noms du/des chercheur(s) : Alex de Carvalho, Olivia Bogo

Email pour la correspondance : alex.de-carvalho@u-paris.fr; oliviabogo@letras.ufjf.br

Labo / service : LaPsyDÉ: Laboratoire de Psychologie du développement et de l'éducation de l'enfant; Université Paris Cité

Évalué à la séance du 02/07/2024

AVIS : Favorable

L'avis du CER U-Paris Cité n'exonère pas des formalités réglementaires. A cet égard, il vous appartient notamment, si vous traitez des données se rapportant à un individu directement ou indirectement identifiable, de vous conformer au règlement européen sur la protection des données (RGPD) en vigueur depuis 2018. Pour cela, vous devez solliciter les conseils du/de la DPD (délégué.e. à la protection des données) de votre établissement.

Les investigateurs faisant appel aux services du CER U-Paris Cité s'engagent à lui signaler tout événement non anticipé survenant en cours d'étude. Ces éléments seront utilisés aux fins d'amélioration des futurs services et conseils que le CER U-Paris Cité pourrait donner.

N° IRB : 00012024-62

Jacqueline Fagard

email : cer_parisdescartes@services.cnrs.fr ; **secrétaire :** David Brami (david.brami@parisdescartes.fr)

APÊNDICE H — Termo de Consentimento Livre e Esclarecido

Termo de Consentimento Livre e Esclarecido para participar da pesquisa: Como crianças brasileiras interpretam palavras concretas e abstratas?

Pesquisadores liderados por Alex de Carvalho em Universidade de Paris Cité estão conduzindo este estudo, "Como crianças brasileiras interpretam palavras concretas e abstratas?", no Lookit.

Objetivo

No período em que a criança está adquirindo a sua língua materna, estabelecer a relação entre uma palavra e o item que ela nomeia no mundo pode ser uma tarefa difícil, principalmente as palavras mais abstratas, por exemplo "delicioso", "fofo", que não são tão simples de referenciar. Neste estudo, queremos entender como crianças brasileiras, entre 24 e 30 meses de idade, compreendem palavras conceitualmente diferentes, como "bolo" e "gostoso".

Procedimentos

Este estudo leva cerca de 10 minutos para ser concluído. Seu filho irá ver duas imagens pareadas - lado a lado - e escutará um áudio que descreverá uma dessas figuras. Por exemplo, aparecerá a imagem de um cachorro e de um gato e a criança ouvirá o áudio "olha! gato! viu! gato!". Você, responsável, não deve interagir com a sua criança ou apontar para a tela para não intervir no teste.

Participação

Você e sua criança são livres para escolher se querem participar deste estudo. Se você e sua criança decidirem participar, não há problema em parar a qualquer momento durante a sessão. Por favor, pause ou interrompa a sessão se sua criança ficar muito agitada ou não quiser participar! Se este for um estudo com múltiplas sessões, não há problema em não concluir todas as sessões.

Remuneração

Não há nenhum tipo de retribuição financeira ou benefício direto a você ou a sua criança pela participação, mas esperamos que vocês gostem de contribuir com esta pesquisa.

Coleta de dados e gravação de webcam

Durante a sessão, você e sua criança serão gravados pela webcam e pelo microfone do seu computador. As gravações de vídeo e outros dados inseridos são enviados com segurança para a plataforma Lookit e armazenados indefinidamente. No final da sessão, você deverá escolher um nível de privacidade para as gravações da sua webcam. Nesse momento, você terá a opção de retirar seus dados de vídeo. Você pode ver suas gravações anteriores no Lookit a qualquer momento.

Os dados são armazenados de forma segura nos servidores do Lookit e por pesquisadores. No entanto, existe sempre um pequeno risco de que dados transmitidos pela Internet possam ser interceptados ou de que a segurança de dados armazenados possa ser comprometida.

Nenhum vídeo será publicado ou compartilhado, a menos que você permita isso ao selecionar um nível de privacidade. Se não recebermos uma gravação de

consentimento para esta sessão (o vídeo que você fará à direita) e não pudermos verificar se você concordou em participar, nenhum outro vídeo da sua sessão será visualizado.

Uso de dados pelos pesquisadores do estudo

O grupo de pesquisa liderado por Alex de Carvalho em Universidade de Paris Cité terá acesso a vídeos e outros dados coletados durante esta sessão. Também teremos acesso ao perfil da sua conta, ao questionário demográfico e ao perfil da criança que está participando, incluindo alterações que você fizer no futuro em qualquer uma dessas informações. Poderemos comparar as respostas da sua criança em relação às respostas anteriores a este ou a outros estudos realizados pelo nosso grupo, às respostas dos irmãos a este ou a outros estudos realizados pelo nosso grupo, ou às respostas de questionários demográficos.

Nós avaliaremos para onde o seu filho olhou e por quanto tempo ele olhou para a imagem do gato, após escutar a palavra "gato". Ao final da tarefa, você poderá decidir o nível de privacidade dos seus vídeos.

Uso de dados pelo Lookit

Como este estudo está sendo realizado na plataforma Lookit, os pesquisadores que trabalham no projeto Lookit no MIT também terão acesso aos dados coletados durante esta sessão, além dos dados da sua conta, perfis de crianças e respostas de questionários demográficos. Esses dados podem ser usados pelo Lookit para detectar e corrigir problemas técnicos ou identificar novos recursos que possam ser úteis. Os dados podem ser usados, também, para fornecer suporte aos pesquisadores do estudo, avaliar a qualidade dos dados (por exemplo, até que ponto um observador consegue dizer para qual direção as crianças estão olhando), avaliar o sucesso do site em alcançar uma população diversificada, e caracterizar o envolvimento familiar (como, por exemplo, observar quais aspectos de um estudo tornam os pais mais interessados em voltar mais tarde).

Publicação de resultados

Os resultados da pesquisa poderão ser apresentados em eventos científicos ou publicados em revistas científicas. Os dados brutos (como, por exemplo, o tempo do olhar para a esquerda e para a direita da tela) podem ser publicados quando não permitem a identificação das crianças. **Nunca publicamos as datas de nascimento ou os nomes das crianças, assim como não publicamos os dados demográficos juntamente com o vídeo da sua criança.**

US Patriot Act Disclosure

Lookit é uma organização dos EUA e todas as informações coletadas do site são armazenadas em servidores baseados nos EUA. Portanto, suas gravações de vídeo estão sujeitas às leis dos EUA, como o US Patriot Act. Esta lei permite que as autoridades tenham acesso aos registros dos provedores de serviços de Internet. Se você optar por participar deste estudo, você entende que sua gravação de vídeo será armazenada e acessada nos EUA. A política de segurança e privacidade do Lookit pode ser encontrada no seguinte link: <https://lookit.mit.edu/privacy/>.

Direitos dos participantes da pesquisa

Você não está renunciando a quaisquer reivindicações, direitos ou recursos legais devido à sua participação neste estudo de pesquisa. Se você sentir que foi tratado injustamente ou tiver dúvidas sobre seus direitos como sujeito de pesquisa, entre em contato com Comité d'Éthique de la Recherche, Université Paris Cité, pelo email: cer_parisdescartes@services.cnrs.fr

Informação de contato do pesquisador

Este estudo é conduzido por Alex de Carvalho em Universidade de Paris Cité. Se você ou sua criança tiverem alguma dúvida ou preocupação sobre este estudo, ou no caso muito improvável de uma lesão relacionada ao estudo, entre em contato com Alex de Carvalho - alex.de-carvalho@parisdescartes.fr.
Se você ou sua criança tiverem alguma dúvida ou preocupação sobre a plataforma Lookit, entre em contato com a equipe do Lookit em lookit@mit.edu ou 617 324 4859.

Eu, _____, responsável legal do aluno(a) _____, estou ciente e autorizo a participação dele(a) na pesquisa “Como crianças brasileiras interpretam palavras concretas e abstratas?”.

Petrópolis - RJ, ____ de _____ de 2025.

Assinatura:

APÊNDICE I — Material de divulgação do teste

UFRJ e Universidade de Paris convidam
crianças de 24 a 30 meses de idade a
participar da pesquisa:

"Como crianças brasileiras interpretam
substantivos concretos e abstratos?"



Você já pensou como bebês aprendem o
significado das palavras da nossa língua?

Palavras como 'comida' ou 'colher' são fáceis para os
bebês. Mas e as palavras como 'gostoso' ou
'delicioso'? Elas não referenciam algo visível no mundo.
Será que são mais difíceis de aprender?

Queremos descobrir com esta pesquisa!

Para participar, clique no link abaixo.



[Como crianças brasileiras interpretam
palavras concretas e abstratas?](#)

Você precisará fazer a pesquisa no seu computador.

Para saber mais, fale com **Olívia**, a pesquisadora responsável.
Aponte a câmera do seu celular para o QR Code:

