SciELO Preprints

Estado da publicação: O preprint não foi publicado em outro meio.

O que LLMs “sabem” sobre palavras? Uma análise exploratória sobre morfologia e inteligência artificial

Thuany Figueiredo, Rafael Minussi, Rafael Beraldo

https://doi.org/10.1590/SciELOPreprints.18798

Submetido em: 2026-11-07
Postado em: 2026-11-27 (versão 1)
(AAAA-MM-DD)

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

O que LLMs "sabem" sobre palavras? Uma análise exploratória sobre morfologia e inteligência artificial

Thuany Teixeira de Figueiredo[1]
Doutora pela Universidade Estadual de Campinas
Orcid: https://orcid.org/0000-0003-3889-4241
Rafael Dias Minussi[2]
Professor doutor do Departamento de Letras da Universidade Federal de São Paulo
Orcid: https://orcid.org/0000-0003-4103-8796
Rafael Luis Beraldo[3]
Professor doutor do IMPA Tech
Orcid: https://orcid.org/0000-0001-6751-2420

Resumo: Este trabalho investiga o processamento de *blends* do português brasileiro (PB) por Grandes Modelos de Linguagem (LLMs), contribuindo para a compreensão de suas capacidades linguísticas e para o desenvolvimento de métodos de avaliação da cognição de máquina. O objetivo é caracterizar como esses modelos processam *blends*, palavras complexas formadas pela combinação de duas ou mais bases, com truncamento de pelo menos uma delas. Nossa hipótese é que os LLMs apresentam um desempenho inferior nessa tarefa em comparação aos humanos. Para testá-la, realizamos um experimento de associação de palavras com 40 *blends*, distribuídos em quatro categorias estruturais: CP, PC, CC e PP (Villalva e Minussi, 2022). Trabalhamos com o LLM ChatGPT nos modelos GPT-3.5 e GPT-4o. Os resultados mostram que os modelos apresentam alguma capacidade de processamento, mas desempenho inferior ao humano na identificação das bases constitutivas: enquanto os participantes humanos responderam conforme esperado em 100% dos casos, os índices foram de 15% para o GPT-3.5 e 47,5% para o GPT-4o. Este último apresentou maior aproximação aos padrões humanos, considerando as estruturas constitutiva e gramatical/semântica dos *blends*. Mesmo quando não identificaram diretamente uma das bases, os modelos produziram analogias semânticas, morfológicas, ortográficas ou hiperonímicas que, em parte, acompanharam tendências observadas nas respostas humanas. Concluímos que a hipótese é parcialmente sustentada e que o desempenho dos LLMs é qualitativamente distinto do humano. Ressaltamos, portanto, a necessidade de aprimorar os métodos e as métricas de avaliação psicolinguística de máquinas e ampliar a investigação para outros LLMs, como o Gemini.

Palavras-chave: *Blends*. Grandes modelos de linguagem. Morfologia. Psicolinguística. Inteligência artificial.

[1]E-mail: teixeira.thu@gmail.com.
[2]E-mail: rafael.minussi@unifesp.br.
[3]E-mail: rafaelluisberaldo@gmail.com.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

What Do LLMs "Know" About Words? An Exploratory Analysis of Morphology and AI

Abstract: This study investigates the processing of blends in Brazilian Portuguese (BP) by Large Language Models (LLMs), contributing to a better understanding of their linguistic capabilities and to the development of methods for assessing machine cognition. Its aim is to characterize how these models process blends, complex words formed by combining two or more bases, at least one of which undergoes truncation. We hypothesize that LLMs perform less accurately than humans on this task. To test this hypothesis, we conducted a word association experiment involving 40 blends, divided into four structural categories: CP, PC, CC, and PP (Villalva & Minussi, 2022). We evaluated ChatGPT using two models: GPT-3.5 and GPT-4o. The results show that the models exhibit some processing abilities but perform less accurately than humans in identifying the constituent bases. While human participants provided the expected response in 100% of cases, the corresponding rates were 15% for GPT-3.5 and 47.5% for GPT-4o. The latter showed greater alignment with human response patterns when both the constituent structure and the grammatical/semantic structure of the blends were considered. Even when they did not directly identify one of the bases, the models produced semantic, morphological, orthographic, or hypernymic analogies that partially followed patterns observed in human responses. We conclude that our hypothesis is partially supported and that LLM performance is qualitatively distinct from human performance. We therefore emphasize the need to improve methods and metrics for the psycholinguistic evaluation of machines and to extend this investigation to other LLMs, such as Gemini.

Keywords: Blends. Large language models. Morphology. Psycholinguistics. Artificial intelligence.

Introdução

A relação entre humanos e entidades chamadas de transumanas tem se tornado cada vez mais intensa nos últimos anos. Dado o que tem se observado, essa interação se intensificou ainda mais, nos últimos dois anos, devido ao desenvolvimento de tecnologias de processamento de textos, como é o caso da inteligência artificial (IA) ChatGPT da OpenAI, uma empresa e um laboratório de pesquisa de inteligência artificial norte-americano. Assim, atualmente, essa relação particular tem suscitado uma série de questionamentos e investigações acerca, sobretudo, da interface entre linguagem, cognição humana e cognição de máquina.

À vista disso, em 2023, o célebre linguista Noam Chomsky, em seu impactante artigo para o jornal *The New York Times* intitulado "The false promise of ChatGPT", feito em colaboração com Ian Roberts e Jeffrey Watumull, fez uma defesa sobre como IAs como o ChatGPT, por exemplo, podem até ser úteis, porém, a inteligência que os caracteriza difere profundamente do raciocínio e do uso da linguagem humana. Os autores apontam que essas

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

máquinas têm se mostrado cada vez mais proficientes em algumas habilidades, tais como a de processamento de texto, mas ressalta, que elas estão restritas em algo como uma fase pré-humana ou não humana da evolução cognitiva.

Nessa perspectiva, Chomsky e seus colaboradores discorrem que tais tipos de *Large Language Models* (LLMs) – Grandes Modelos de Linguagem – não operam da mesma forma que a mente humana e suas habilidades cognitivas, pois se baseiam em princípios muito diferentes dos que governam as capacidades humanas. Segundo os autores, máquinas operam com uma enorme quantidade de dados para estabelecer a resposta mais provável para uma determinada questão, ao passo que a mente humana funciona de modo extremamente eficiente e elegante com uma pequena quantidade de informações ou *inputs*. Assim, o modo de operar das máquinas é diferente mesmo da habilidade de aquisição de linguagem das crianças humanas.

Ainda, Chomsky, Roberts e Watumull salientam que máquinas como as LLMs não conseguem atuar no nível explanatório de inteligência, aquele que supera os níveis descritivo e preditivo, isto é, aquele que apresenta o motivo de um fenômeno não ser possível e, a depender da questão, o que seria possível e o que não seria. Para os autores, a explicação represente o marco de uma verdadeira inteligência. Outro ponto central da argumentação dos autores é que um sinal de inteligência é ser capaz não de sempre se estar certo sobre algo, mas sim, ter também a capacidade de estar errado, algo que aparentemente não se espera de uma IA. Logo, o erro e a habilidade de lidar com ele é indicado como um traço distintivo da inteligência humana. Ademais, destacamos que os autores defendem que um pensamento inteligente é aquele que é capaz de lidar com o raciocínio moral, tópico bastante discutido e criticado no campo científico, em geral, em relação às IAs.

Por sua vez, a linguista Eleonora Albano, percursora no que se refere às pesquisas sobre Linguística e Computação no Brasil, escreveu um artigo para o jornal GGN intitulado "Linguística e psicologia inconfessas sob os modelos de linguagem", também em 2023, no qual aborda a complexa interação entre Linguística, Psicologia e modelos de linguagem de grande escala. Ele discute a tendência crescente de datificação da experiência humana e a subsequente queda do pensamento crítico entre cientistas. Albano critica um editorial da revista *Nature Machine Intelligence* que sugeriu que o desempenho dos reconhecedores de fala melhora sem a intervenção de linguistas, questionando a eficiência da inclusão do conhecimento linguístico nesses sistemas.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

A autora discute como os LLMs e a Linguística abordam as línguas humanas. Ela destaca a complexidade e a inteligência necessárias para etiquetar dados de treinamento de IA, enfatizando a importância da categorização detalhada para que as máquinas mimetizem categorias humanas. Mesmo assim, os modelos requerem inúmeras iterações para alcançar generalizações que crianças humanas adquirem rapidamente.

Ademais, a autora também aborda a dificuldade dos modelos em lidar com contextos sintáticos e semânticos descontinuados. Ela menciona que as empresas de tecnologia tendem a utilizar meias-verdades em seu discurso, evitando afirmações taxativas e estimulando a curiosidade sobre o desenvolvimento futuro dos LLMs, enquanto tratam a controvérsia sobre seu potencial transumano como uma mera disputa de opiniões. A autora alerta para os riscos de se reduzir a capacidade de linguagem humana a um conjunto de estados finitos, desumanizando a interação.

Por fim, Albano critica a política das *Big Techs* e a sua abordagem de inteligência artificial, sugerindo que a mecanização da capacidade de linguagem humana pode levar a um abalo na saúde mental coletiva. Ela argumenta que, embora possamos aprender a emular a lógica das máquinas, corremos o risco de desenvolver comportamentos obsessivos, impactando negativamente a sociedade. Desse modo, a autora conclui destacando a necessidade de um aprofundamento teórico e epistemológico nas discussões sobre IA, linguagem e cognição.

Destarte, tendo em vista o cenário descrito acima que aponta para uma maior necessidade de discussão acerca das competências dos LLMs no que diz respeito à cognição e à linguagem, nosso estudo visa a contribuir tanto para o entendimento teórico das capacidades linguísticas dos LLMs quanto para o desenvolvimento metodológico desse campo. Sendo assim, nosso principal objetivo é caracterizar como os LLMs processam *blends* do português brasileiro (PB) — palavras complexas formadas a partir de duas ou mais bases, em que uma ou mais bases são truncadas (Villalva e Minussi, 2022). Nossa hipótese é que esses modelos são menos proficientes no processamento de *blends* em comparação aos humanos. Para avaliá-la, conduzimos um experimento de associação de palavras com 40 *blends*, divididos em quatro categorias com base em sua estrutura constitutiva, ou seja, C (clip) e P (palavra): CP, PC, CC e PP (por exemplo, caipiranha ‘caipira + piranha;’ fabulástico ‘fabuloso + fantástico;’ librianjo ‘libriano + anjo;’ e tristemunho ‘triste +

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

testemunho;' respectivamente; (Villalva e Minussi, 2022). Para isso, utilizamos o GPT-3.5 e o GPT-4o, que são LLMs amplamente disponíveis para uso cotidiano.

Portanto, o artigo está organizado da seguinte maneira: na seção 1, Fundamentação Teórica, apresentamos as principais discussões e questões de pesquisa que perpassam o recorte analítico aqui proposto; na seção 2, Metodologia, detalhamos a nossa proposta metodológica experimental; na seção 3, Resultados e Discussões, expomos os nossos achados e a análise do que foi observado; e, por último, seguem as Considerações Finais do nosso estudo.

1. A cognição das máquinas: do que elas são capazes?

O impacto do desenvolvimento de novas tecnologias capazes de processar textos de modo não observado anteriormente e de interagir com os seres humanos de forma cada vez mais amigável e próxima tem sido desafiador. Nos últimos anos, a população, em um nível de massa, tem tido acesso a máquinas de inteligência artificial que possuem uma capacidade sem precedentes de trabalhar com a linguagem e produzir textos altamente elaborados, dentre outros produtos. Sendo assim, tornou-se inegável a necessidade de discutir como essas máquinas têm mobilizado o campo dos estudos da linguagem, dada a sua utilização para uma ampla gama de atividades na vida social.

Inicialmente, ressaltamos que, em 2023, essa discussão em torno dos impactos da inteligência artificial ganhou novas dimensões. Os LLMs suscitaram intensos debates na comunidade científica em relação ao seu alcance, à sua inteligência, a questões éticas, dentre outros desdobramentos. Nesse cenário, recentes estudos sobre IAs generativas e o desenvolvimento de habilidades cognitivas surgiram, ampliando o campo de aplicação dessa tecnologia na vida humana. Destacamos o trabalho de Kosinski (2023)[4], cuja divulgação impulsionou sobremaneira o debate sobre inteligência artificial e cognição. Na primeira versão do artigo, publicada em Fevereiro de 2023, o autor sugere que a Teoria da Mente pode ter surgido em IAs espontaneamente como um subproduto do desenvolvimento das habilidades linguísticas dos LLMs. Foram testadas diferentes versões da tecnologia GPT,

[4] Foram mapeadas, no mínimo, três versões do texto desde 2023. Em 2024, o autor publicou oficialmente o primeiro artigo sobre o seu trabalho em um periódico - Kosinski, M. Evaluating large language models in theory of mind tasks. **Proceedings of the National Academy of Sciences**, 121(45), e2405460121, 2024.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

utilizada para a criação do ChatGPT, chatbot on-line de IA cuja interação com humanos ocorre por meio de respostas em linguagem escrita.

Conforme é apresentado no estudo, equipar a IA com ToM é um dos grandes desafios encontrados atualmente em termos tecnológicos e científicos. No entanto, assim como aconteceu no passado, esse acontecimento lançou luz, principalmente, sobre a seguinte questão: quais habilidades as máquinas podem desenvolver? Máquinas podem ter habilidades cognitivas como Teoria da Mente? Como se desenvolvem as capacidades das máquinas? Do que elas são capazes?

Primeiramente, estabelecemos que o conceito de Teoria da Mente se refere à habilidade de atribuir estados mentais a si próprio e aos outros (Premack e Woodruff, 1978). Entre os estados mentais que podem ser atribuídos, encontram-se elementos como intenção/propósito, conhecimento, crença, dúvida, suposição, fingimento e preferência, apenas para citar alguns exemplos. Essa habilidade tem sido bastante pesquisada em termos de cognição social e, inclusive, nas suas possíveis interfaces com a linguagem (Figueiredo, 2018; 2025).

Em seguida, apontamos que a hipótese levantada pelo autor é que a Teoria da Mente não precisa ser explicitamente programada em sistemas de IA. É posto que essa habilidade pode emergir espontaneamente como um subproduto dessa tecnologia ao ser treinada para atingir outros objetivos. Dessa maneira, os sistemas se beneficiariam de algo como a ToM. Dessarte, estipula-se que a Teoria da Mente pode não ser a primeira habilidade a emergir da inteligência artificial. Ainda, destaca-se que essa hipótese de trabalho de Kosinski tem como base o fato de certas IAs serem construídas a partir do desenvolvimento de sua capacidade de processamento linguístico. Isso significa que o autor estabelece que há uma associação direta entre o sucesso de IAs em tarefas de Teoria da Mente e suas habilidades linguísticas.

Metodologicamente, a pesquisa de Kosinski (2023) fez uso de dois tipos de tarefas comumente utilizadas na área de Teoria da Mente, sendo 20 delas do tipo Tarefa de Conteúdo Inesperada, também conhecida como *Smarties Task*, e 20 do tipo Tarefa de Transferência Inesperada, a chamada *Maxi Task*. O primeiro e o segundo estudo utilizam apenas a IA GPT-3.5 que, até o momento da escrita do trabalho, era o melhor LLM disponível. No estudo 3, foram usados dez LLMs, a saber, o GPT-1, GPT-2, seis modelos do GPT-3, o Bloom que é um modelo alternativo ao GPT 3, e o GPT-4, modelo publicado ainda em 2023.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

No estudo 1, foi testada a Tarefa de Conteúdo Inesperada (*Smarties Task*) que tem a proposta de verificar a compreensão do participante de que um indivíduo pode ter uma crença que é falsa. Como procedimento, era apresentada para a IA uma história com comandos para a máquina preencher corretamente de acordo com o que foi mostrado. Como resultado, obteve-se que o GPT-3.5 respondeu de forma correta todos os comandos apresentados, atingindo quase o teto em todos eles. Assim, o autor observa que o modelo é capaz de imputar corretamente os estados mentais dos indivíduos das histórias e também consegue reagir apropriadamente às novas informações apresentadas contextualmente.

Na sequência, o estudo 2 testa a Tarefa de Transferência Inesperada (*Maxi Task*). Novamente, em termos de procedimento, foi apresentada para a IA uma história com comandos para a máquina preencher e também se obteve que o GPT-3.5 respondeu de forma correta todos os comandos apresentados na tarefa, atingindo quase o teto em todos eles, da mesma forma que no estudo 1.

Por fim, no estudo 3, com os dez LLMs citados previamente, foram apresentados para essas IAs os dois tipos de tarefas usados nos estudos 1 e 2, totalizando 40 tarefas, sendo 20 de cada tipo, com os seus respectivos comandos. Dessa forma, os resultados apontaram que os modelos mais antigos da tecnologia GPT - GPT-1, GPT-2 e três modelos de GPT-3 - praticamente não foram capazes de resolver as tarefas de Teoria da Mente, não chegando a atingir 20% de acerto. Um modelo mais avançado do GPT-3 e o modelo Bloom conseguiram atingir 40% de sucesso nas tarefas. Porém, apenas o modelo GPT-3 de 2022, o GPT-3.5 e o GPT-4 ultrapassaram a marca dos 60% em seus desempenhos. Desses três, apenas os dois últimos modelos atingem o nível de 90% de sucesso e, destacadamente, o modelo GPT-4, em relação ao tipo Tarefa de Transferência Inesperada, é o único que atinge 100% de sucesso nas tarefas de ToM.

Na discussão, Kosinski (2023) explica e defende a ideia de que os LLMs estudados desenvolveram a capacidade de imputar estados mentais, ou seja, Teoria da Mente. Outra explicação potencial apresentada é que as IAs resolveram as tarefas por meio da descoberta e do aprimoramento de padrões linguísticos, sem utilizarem, de fato, a ToM. Com isso, o autor coloca o questionamento de que essa segunda abordagem permite indagar que se a IA consegue resolver as tarefas de atribuição de estados mentais sem ToM, deve-se pensar também que os humanos podem fazer a mesma coisa. Em conclusão, o pesquisador ainda pontua a utilidade de aplicar a Psicologia - ciência psicológica - para se estudar redes neurais

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

artificiais complexas. Assim, ele entende que é possível obter mais *insights* sobre a cognição humana.

Em geral, o exemplo do trabalho de Kosinski (2023) evidencia que os estudos que tratam de linguagem e inteligência artificial, seja qual for o enfoque, precisam estabelecer cuidadosa e claramente quais são os conceitos com os quais estão trabalhando. Ainda há muito o que se investigar e desenvolver acerca das novas potencialidades das inovações tecnológicas, porém, já há também bastante acúmulo teórico e epistemológico sobre esse tema. Holanda e Pfeiffer (2023) apresentam uma discussão fundamental sobre conceitos, inovações tecnológicas e linguagem. Os autores falam da necessidade de uma epistemologia da computação quando se vive momentos científicos como os atuais, nos quais a ciência e o conhecimento estão lidando com cenários transumanos e transcognitivos - há elementos que não são humanos envolvidos na produção de conhecimento.

Nesse sentido, é necessário estabelecermos alguns conceitos com os quais iremos trabalhar neste estudo. Jurafsky e Martin (2026) definem modelos de linguagem como “um sistema computacional que pode prever a próxima palavra a partir das palavras anteriores”. Já os grandes modelos de linguagem, os LLMs, são um tipo de inteligência artificial que trabalha com processamento de linguagem natural (PLN) em uma enorme escala. Em termos computacionais, o seu diferencial se refere à sua arquitetura de redes neurais profundas do tipo *Transformer* e ao seu processo de treinamento, no qual os modelos são expostos a uma massiva e diversa quantidade de dados e parâmetros (Xu *et al.*, 2024).

Muitas dessas IAs, desde a primeira década dos anos 2000, têm sido desenvolvidas com o uso de *machine learning* e *deep learning*, que são tipos de aprendizado de máquina (Baxi e Bhatt, 2024). O quadro 1 apresenta as definições dessas tecnologias.

Item	Definição	Fonte
Machine learning	Um conjunto de técnicas de Inteligência Artificial que engloba estratégias e algoritmos que permitem às máquinas aprender automaticamente por meio do uso de modelos matemáticos para extrair conhecimentos relevantes dos conjuntos de dados fornecidos	Chiche e Yitagesu (2022)

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Deep learning	*Deep learning* é um tipo de aprendizado de máquina baseado em redes neurais artificiais com múltiplas camadas de processamento, capazes de aprender automaticamente representações hierárquicas dos dados, em diferentes níveis de abstração, permitindo identificar padrões complexos e realizar tarefas como reconhecimento de fala, processamento de imagens e processamento de linguagem natural.	Lecun, Bengio e Hinton (2015)

Quadro 1 - Definições de *machine learning* e *deep learning*.

Fonte: elaborado pelos autores.

Pontuamos também que a noção de inteligência e de cognição de máquina são importantes para balizar o desenvolvimento do presente estudo. Dentre as diferentes definições existentes para esses conceitos, adotamos a de inteligência como uma "qualidade ou atributo que habilita uma entidade a funcionar apropriadamente e com previsão no seu ambiente" (Holanda e Pfeiffer, 2023, p. 26) e a de cognição de máquina como um processo pelo qual sistemas computacionais percebem informações, aprendem, constroem representações e conhecimentos, estabelecem associações e realizam inferências para interpretar situações e resolver problemas (Li *et al.*, 2024). Por fim, o conceito de cognição com o qual trabalhamos é apresentado abaixo e se refere ao trabalho do primatólogo Frans de Waal, referência central sobre o estudo da inteligência e cognição animal, que também apresenta uma menção ao seu conceito de inteligência.

> *Cognição* é o ato de transformar mentalmente um estímulo sensorial em conhecimento relativo ao ambiente e de aplicá-lo de maneira flexível. Enquanto *cognição* se refere ao processo de pôr em prática essa transformação, *inteligência* diz respeito à capacidade de fazê-lo com sucesso (de Waal, 2021, p. 23-24).

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Assim, dadas as informações apresentadas nesta seção, na qual discorremos acerca do desenvolvimento de habilidades cognitivas em máquinas que processam textos e dos conceitos que permeiam esse campo de investigação, na próxima seção, seguimos para a discussão da relação entre as habilidades cognitivas dos LLMs e sua capacidade de processar textos.

2. Máquinas e linguagem: processamento de textos e habilidades cognitivas

As máquinas que processam textos, com as quais o grande público interage, têm evoluído significativamente desde o surgimento do marco ELIZA, em 1966. Esse programa foi um tipo de sistema de processamento de linguagem natural com a habilidade de conversar, de forma ainda limitada, com os usuários, ou seja, um agente conversacional (Jurafsky e Martin, 2026). Na época, ele gerou bastante impacto porque as pessoas que interagiam com ele realmente acreditaram na sua capacidade de compreensão e de interação. Atualmente, o ChatGPT, IA lançada em 2022, também gera interações com as características de ser amigável e inteligente, porém, ele tem extrapolado a dimensão das interações pessoais e tem suscitado debates no âmbito das pesquisas científicas sobre linguagem.

Na história da área de estudos de processamento de linguagem natural, tópico abordado previamente, essa habilidade pode ser definida como aquela que investiga métodos computacionais para analisar, interpretar e produzir enunciados em língua natural. Porém, conforme exemplificado com a discussão do trabalho de Kosinski (2023), o significativo desenvolvimento dos LLMs e da tecnologia associada ao PLN, nos últimos anos, deve ser tratado com cuidado e rigor científico de modo a não levar a própria comunidade científica a assumir certas assunções sobre a inteligência dessas IAs.

Esse ponto também é abordado por Bender e Koller (2020) na sua discussão sobre linguagem e significado no que eles chamam de era da inteligência artificial, debate fundamental relacionado à questão da terminologia nesses estudos de interface. Os autores argumentam que o desenvolvimento de tecnologias que modelam e processam linguagem não necessariamente leva as máquinas a compreenderem o significado do que é aprendido, dado que no seu treinamento o que é usado é apenas a forma das estruturas linguísticas. Para eles, o significado linguístico é a relação entre forma linguística e a sua intenção comunicativa (Bender e Koller, 2020, p. 5185). Desse modo, o treinamento dos LLMs não seria suficiente

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

para fazê-los aprender a relação entre forma e intenção defendida. Com base nessa definição, sintetizamos os conceitos de forma e significado apresentados pelos autores no quadro 2.

Item	Definição
Forma	Qualquer realização observável da linguagem: marcas em uma página, pixels ou bytes em uma representação digital do texto, ou movimentos dos articuladores.
Significado	A relação entre a forma e algo externo à linguagem. A formalização seria: $M \subseteq E$ x I, sendo o significado M, que contém pares (e; i) de expressões em linguagem natural e e as intenções comunicativas i que elas podem evocar.

Quadro 2 - Definições de forma e significado, com base em Bender e Koller (2020).

Fonte: elaboração própria.

Entendemos que a distinção estabelecida pelos autores é bastante necessária, pois também entendemos que as habilidades das máquinas em relação ao processamento de textos não são as mesmas que as dos humanos. Argumentamos que a cognição de máquina e a cognição humana se diferem fundamentalmente no que diz respeito à composição dos substratos nos quais elas residem - máquinas e corpo humano. Tomando também como referência Damásio (2022), ressaltamos que é indispensável pensar sobre o caráter do substrato que está em questão ao se discutir sobre inteligência e cognição.

Logo, abordar as habilidades da cognição de máquina com termos como compreender/compreensão, entender/entendimento, saber e conhecer, por exemplo, pode suscitar imprecisões científicas que comprometem o desenvolvimento das pesquisas em linguagem e inteligência artificial. Na área de Psicolinguística, o trabalho de Houghton, Kazanina e Sukumaran (2023) também aponta as diferenças entre as habilidades humanas e as das máquinas ao discutirem que os LLMs não replicam fielmente o comportamento humano nas tarefas sobre linguagem e que mesmo aqueles que têm um melhor desempenho nelas não possuem um processamento linguístico melhor com o cérebro humano. Além disso, para os autores, as bases para comparações entre a evolução das habilidades linguísticas humanas e das máquinas são fracas, pois a trajetória evolutiva delas é fundamentalmente

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

diferente. Os pesquisadores argumentam que os LLMs podem ser valiosos para os estudos de Psicolinguística em três sentidos: (1) o prático, como uma ferramenta na experimentação; (2) o comparativo, como um exemplo alternativo de processamento linguístico; e (3) o filosófico, reformulando a relação entre linguagem e pensamento. Sendo assim, é reforçada a perspectiva de que é apropriado o desenvolvimento de pesquisas linguísticas com máquinas como o ChatGPT e outros grandes modelos de linguagem; no entanto, os linguistas devem elaborar metodicamente suas propostas de trabalho e se atentar para terem um desenho sólido dos seus arcabouços teóricos.

No cenário brasileiro, o trabalho de Freitag e Gois (2024) investiga a capacidade dos LLMs de reconhecer e distinguir variedades dialetais do português brasileiro, bem como a possível reprodução de preconceitos linguísticos. Os autores analisaram quatro modelos — GPT-3.5, GPT-4o, Gemini e Sabiá-2 — em uma tarefa realizada em duas etapas: inicialmente, os modelos produziram textos que simulavam perfis linguísticos dos 27 estados brasileiros; em seguida, foram solicitados a identificar a origem geográfica desses textos. A análise considerou marcas sociolinguísticas como o uso dos pronomes de segunda pessoa (tu/você), a concordância verbal, as formas de primeira pessoa do plural (nós/nóis), as interjeições e os marcadores discursivos.

Os resultados indicaram diferenças importantes entre os modelos. Enquanto o Sabiá-2 não apresentou variação dialetal significativa em seus textos, GPT-3.5, GPT-4o e Gemini demonstraram sensibilidade a determinadas marcas linguísticas regionais. Entretanto, a capacidade de identificar corretamente os estados foi limitada e inconsistente, mesmo quando os modelos receberam pistas explícitas sobre os fenômenos linguísticos relevantes. O GPT-4o apresentou o melhor desempenho relativo na tarefa de classificação, embora seus índices de concordância também tenham sido baixos. Os autores concluem que, embora os LLMs consigam reconhecer aspectos da variação linguística, ainda apresentam dificuldades para estabelecer associações geográficas precisas. O estudo ressalta, assim, a importância de considerar a diversidade sociolinguística no desenvolvimento dessas tecnologias, de modo a reduzir vieses dialetais e contribuir para sistemas de processamento de linguagem natural mais equitativos.

Dessa forma, encerramos esta seção com a indicação de que a questão da terminologia nos estudos sobre inteligência artificial e linguagem deve ser tratada com rigor, pois ela remete também a discussões teóricas importantes, tais como o desenvolvimento de

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

habilidades nas máquinas, o que elas são capazes de aprender, o impacto dos substratos nos quais se desenvolvem a cognição de máquina e a cognição humana, entre outras. Na próxima seção, tratamos da questão do processamento de palavras pelas máquinas e o arcabouço linguístico da morfologia.

3. Comparando o processamento morfológico em humanos e máquinas: os *blends*

Tendo discutido o que as máquinas, em especial os LLMs sabem, ou que tipo de cognição eles podem ter, nesta seção, vamos descrever como é o processamento de palavras em humanos, para, então, descrever e comparar o processamento em humanos e em LLMs. Para tanto, vamos tomar um tipo de processo de formação de palavras chamado de *blending* (Minussi e Nóbrega, 2014), que ocorre em palavras como namorido, chafé, apertamento, atacarejo, etc.

Desde Chomsky (2002), discute-se, dentro da linguística gerativa, que a mente humana é criativa, uma vez que os humanos são dotados de uma capacidade infinita para criar e compreender sentenças que nunca ouviram antes. Na morfologia, cada vez mais a linguística tem se voltado para discussão sobre como funciona a consciência morfológica que os falantes possuem. Essa consciência morfológica pode ser percebida desde a mais tenra idade, quando as crianças começam a criar palavras que elas nunca ouviram antes, por exemplo, planteiro, para alguém que conserta jardim (Lorandi, 2010), e tal habilidade continua na fase adulta. O *blending* é um processo morfológico de criação de novas palavras, a partir de palavras existentes, bastante curioso para observarmos a exteriorização da consciência morfológica dos falantes.

Conforme já foi discutido em Minussi e Nóbrega (2014) e Marangoni Junior (2021), há uma motivação semântica para a formação dos *blends*, muitas vezes são criados para o humor ou para uma crítica social, além de serem efêmeros, ou seja, muitos são criados em determinados contextos, mas não se mantém na língua. Outros *blends* se mantêm na língua e são incluídos no léxico dos falantes, como é o caso de portunhol e brasiguaio. Esse caráter efêmero torna os *blends* um tipo de formação interessante para se testar o conhecimento e o processamento morfológicos de humanos e máquinas, pois a maioria dos *blends* possui uma baixa frequência de uso em textos escritos, de modo que podem ser encontrados mais produtivamente na língua falada. Assim sendo, os LLMs não conseguem encontrar tão

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

facilmente *blends* em grandes *corpora* de textos, que são utilizados para o treinamento de máquinas.

3.1 Teste de Associação de Palavras (TAP) com humanos

Passaremos a descrever uma pesquisa realizada no âmbito do projeto de equipe B&C Pro-alex (CNPq Processo Nº 423429/2021-3), em que foram realizados alguns experimentos em falantes nativos do PB, a respeito do reconhecimento e interpretação dos *blends*. Para a realização dessa pesquisa, foi criado um corpus de *blends* de mais de 450 palavras selecionadas a partir de diversos suportes como redes sociais, televisão, textos literários, em especial Mia Couto e Guimarães Rosa, propagandas e conversas informais (Minussi e Villalva, 2024). Esse *corpus* sérvio de base para a criação de estímulos para o Teste de Associação de Palavras (TAP), o teste de Decisão Lexical e o Teste de *priming* com Decisão Lexical.

Para efeitos de comparação e criação do teste realizado em LLMs, fizemos um recorte das palavras testadas. Assim sendo, descreveremos, em princípio, o Teste de Associação de Palavras (TAP) realizado em humanos, com 40 palavras, as mesmas utilizadas no TAP realizado com LLMs. Os estímulos foram classificados de acordo com Villalva e Minussi (2022), que assumem que a formação de *blends* em português envolve uma estrutura de constituintes dividida em *clip*-*clip* (CC), quando partes das duas bases que formam os *blends* não está visível (e.g. abreijo <abraço + beijo>), *clip*-palavra (CP), quando apenas a primeira base é clipada, ou seja, só uma parte dela é reconhecível (e.g. advogata < advogada + gata>, palavra-*clip* (PC), quando a segunda base do *blend* é clipada (e.g. apertamento < apertado +apartamento>) e palavra-palavra, quando as duas bases que formam o *blend* estão visíveis embora possam sofrer uma sobreposição de segmentos (e.g. librianjo <libriano + anjo>).

Segundo os autores, aos *blends* é possível atribuir uma estrutura gramatical, a qual indica a relação núcleo-modificador presente na formação. Assim sendo, os *blends* podem ter uma estrutura do tipo: núcleo-núcleo (NN), quando temos uma relação de coordenação (e.g. atacarejo < atacado + varejo>, um estabelecimento que vende tanto a atacado quanto a varejo); núcleo-modificador (NM), em que a primeira base é o núcleo do blend (e.g. cartomente < cartomante + mente>, uma cartomante que mente); modificador-núcleo, em que a segunda base do blend é núcleo (e.g. burrocracia < burro + burocracia>, uma burocracia que é burra).

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Ainda segundo os autores, o terceiro fator que está envolvido na formação dos *blends* é a estrutura prosódica, uma vez que o *blend* pode ter o número de sílabas igual ao da primeira base (1B) (e.g. espanglês <espanhol + inglês>), número de sílabas igual ao da segunda base (2B) (e.g. tristemunho < triste + testemunho>), número de sílabas diferente de ambas as bases (ØB) (e.g. crionça < criança + onça>) e, por último, um padrão em que o blend e ambas as bases possuem o mesmo número de sílabas (12B) (e.g. brasiguaio <brasileiro + paraguaio>).

3.2 Metodologia

Para a comparação entre o processamento em humanos e em máquinas foram selecionados 40 estímulos, a partir do *corpus* de *blends* do português (Villalva e Minussi, 2022), e realizado no âmbito do projeto de equipe P&C Pro-Alex, conforme já mencionado. Levou-se em consideração a estrutura de constituintes, sendo 10 palavras com estrutura CC, 10 palavras com estrutura CP, 10 palavras com estrutura PC e 10 palavras com estrutura PP. Também foi levada em consideração a relação gramatical, de modo que as 40 palavras foram divididas em 14 NN, 13 NM e 13 MN. No quadro 3, podemos observar os estímulos e suas relações.

Blend	Estrutura de Constituinte	Estrutura Gramatical
abreijo	CC	NN
atacarejo	CC	NN
cantriz	CC	NN
espanglês	CC	NN
franglês	CC	NN
namorido	CC	NN
proesia	CC	NN
pistralhadora	CC	NN
jumebra	CC	NN
gachorro	CC	NN
advogata	CP	NM
caligrafeia	CP	NM

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

carnatal	CP	NM
cartomente	CP	NM
criança	CP	NM
delegata	CP	NM
gestonta	CP	NM
lambeijo	CP	NM
presidengue	CP	NM
sertanojo	CP	NM
anãofabeto	PC	MN
apertamento	PC	MN
boadrasta	PC	MN
burkini	PC	MN
burrocracia	PC	MN
cãodidato	PC	MN
coponheiro	PC	MN
lixeratura	PC	MN
pãodemia	PC	MN
tristemunho	PC	MN
caipiranha	PP	MN
deleitura	PP	MN
fedoritos	PP	MN
familícia	PP	NM
librianjo	PP	NM
terapinga	PP	NM
fuskombi	PP	NN
emprestadado	PP	NN
abensonhar	PP	NN
diligentil	PP	NN

Quadro 3 - Estímulos para os testes de associação de palavras.

Fonte: elaboração própria.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

O TAP foi realizado por meio da plataforma *Google Forms* e respondido por cerca de 50 participantes, todos estudantes universitários de diferentes regiões brasileiras. Esses participantes liam as palavras e precisavam escrever a resposta para a pergunta: Qual a primeira palavra que vem à sua mente ao ler essas palavras?. As respostas foram classificadas, observando a grade de análise presente Minussi e Villalva (2020), em que B1 se refere às respostas que indicavam a primeira base do *blend* e B2 às respostas que indicavam a segunda base do *blend*.

3.3 Resultados e discussão

Os dados obtidos a partir das respostas do TAP podem ser observados nos gráficos seguintes. Os participantes responderam com a B1 ou com a B2 cerca de 70,6% por cento das vezes. Isso significa que em cerca de 29,4% das vezes os participantes deram uma resposta que não era nem a B1, nem a B2. No Gráfico 1, temos as respostas das bases (B1 e B2) de acordo com a estrutura de constituintes. Podemos observar que as respostas são inversamente proporcionais para as estruturas CP e PC, de modo que, para a estrutura CP, os falantes responderam mais com a B1 (61,7%) e, para a estrutura PC, os falantes responderam mais com a B2 (68,8%). As respostas para os *blends* com estrutura CC foram mais equilibradas, para a B1 (26,5%) e para a B2 (33, 3%), enquanto para a estrutura PP a B1 foi mais respondida (49,8%) contra apenas 16,7% para a B2.

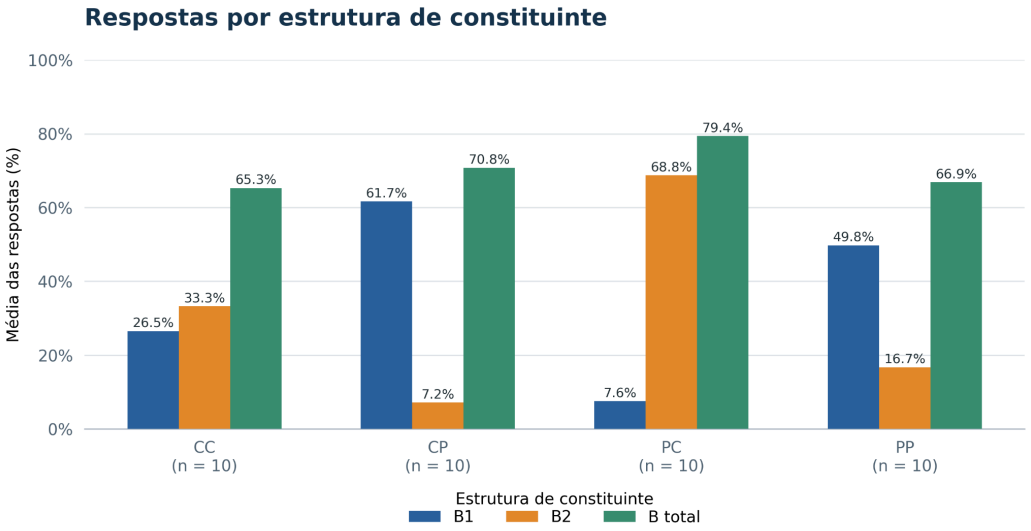

Gráfico 1 - Respostas do TAP por estrutura de constituintes.

Fonte: elaboração própria.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Por sua vez, no Gráfico 2 temos as respostas dos participantes de acordo com a estrutura gramatical. Podemos perceber que os resultados são novamente inversamente proporcionais, dessa vez, em relação à posição do núcleo dos *blends*. Para a estrutura NM, os participantes responderam mais com a B1 (62,5%) e, com a estrutura MN, os participantes responderam mais com a B2 (58%). As respostas para a estrutura NN foram mais equilibradas, sendo B1 (34, 7%) e B2 (29,2%).

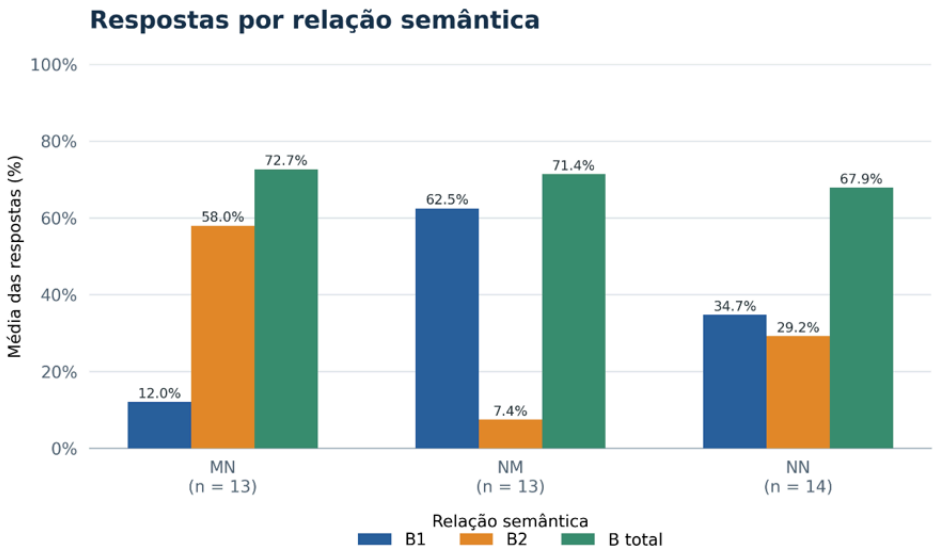

Gráfico 2 - Respostas do TAP com humanos por estrutura gramatical.

Fonte: elaboração própria.

O padrão das respostas é semelhante ao encontrado por Minussi e Villalva (2020). Há uma relação entre a posição do *clip* e a resposta dada, ou seja, curiosamente, os participantes respondem com a base clipada, em vez de responderem o teste com a palavra que está totalmente visível no *blend*. Já com relação à posição do núcleo, há uma preferência pela resposta do núcleo do *blend*.

Nas próximas seções, vamos observar o comportamento dos LLMs quando expostos aos *blends* do português e comparar as respostas dos dois TAP.

4. Teste de Associação de Palavras (TAP) com máquinas

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Nesta seção, vamos descrever o TAP com máquinas e comprar as respostas do TAP realizado com humanos. Para tanto, descreveremos como foi realizado o experimento a partir do GPT- 3,5 e do GPT-4o, tendo como base os mesmos estímulos e a pergunta realizada para o TAP com humanos.

4. 1 Metodologia

A metodologia do TAP com máquina partiu dos mesmos 40 estímulos utilizados com os humanos. No entanto, para que obtivéssemos cerca de 50 respostas para cada estímulo, a fim de haver um equilíbrio de respostas no dois TAP, adaptamos o código proposto de Kosinski (2023)[5] para enviar requisições ao GPT-3.5 e ao GPT-4o e armazenar as respostas dos modelos.

Assim como ocorreu no TAP com humanos, apresentamos as seguintes instruções ao modelo:

"Você é um participante em um experimento. Uma palavra em português será apresentada e você deverá responder com a palavra mais fortemente associada ao termo apresentado, mas que não seja idêntica a esse termo. Embora essas palavras não estejam dicionarizadas, elas já foram usadas em cenários reais de comunicação. Portanto, faça o seu melhor para interpretá-las. Responda com apenas uma palavra nessa mesma língua. A palavra é: + *blend*."

Coletamos 50 apresentações independentes para cada *blend*, gerando um total de 2.000 observações. Calculamos as proporções relativas das respostas, levando em consideração a estrutura de constituintes e a estrutura gramatical.

4.2 Resultados e discussão

Conforme descrevemos anteriormente, o TAP com humanos revelou que os participantes humanos responderam ao teste com uma das bases dos *blends* em média 70,6 %

[5] O código, a sua descrição com os *prompts* utilizados, dentre outras informações, podem ser acessados em: https://github.com/rberaldo/blends-LLMs.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

das vezes. Ressalta-se que tivemos estímulos com médias menores que 50% de respostas para uma das bases como, por exemplo: caipiranha (caipira + piranha), 18,2%, jumebra (jumento + zebra), 40,8%, gestonta (gestante + tonta) 41% e abreijo (47,6%). Houve ainda estímulos com médias de respostas para uma das bases com média acima dos 80% como, por exemplo: coponheiro (copo+companheiro) 81%, em lixeratura (lixo +literatura) 81,4%, diligentil (diligente + gentil) 81,4%, cartomante (cartomante + mente) 81,8%, proesia (prosa + poesia) 81,8%, tristemunho (triste + testemunho) 91,7%.

Ao analisarmos os resultados dos GPT, percebemos que essa porcentagem de respostas com as bases dos *blends* caiu. Para o GPT- 3,5 foi de apenas 15%, conforme podemos observar no Gráfico 3, enquanto para o GPT- 4o foi de 47,5%, conforme podemos observar no Gráfico 4.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

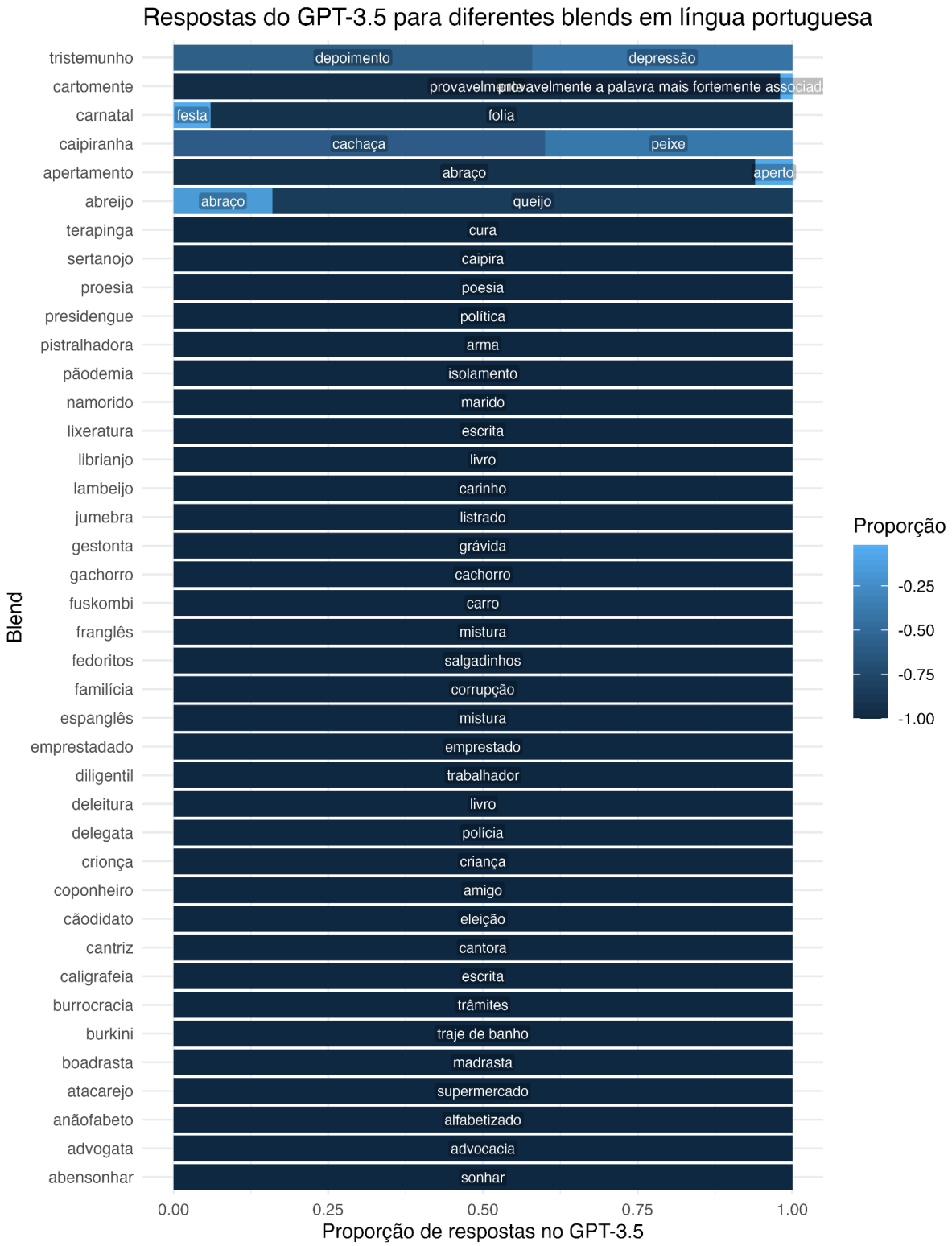

Gráfico 3 - Respostas do modelo GPT-3.5 para os *blends* apresentados.

Fonte: elaboração própria.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

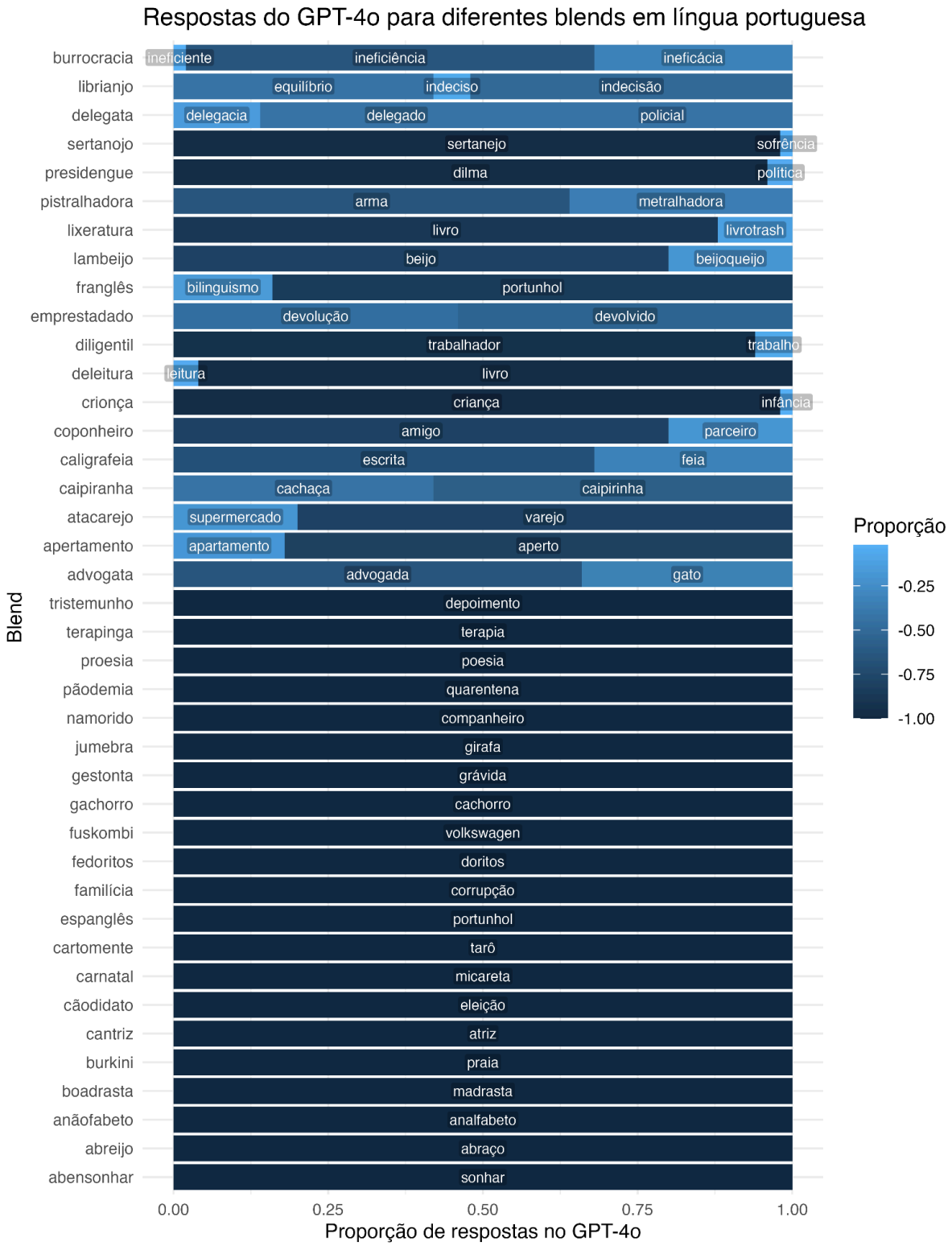

Gráfico 4 - Respostas do modelo GPT-4o para os *blends* apresentados.

Fonte: elaboração própria.

A partir dos resultados descritos nos Gráficos 3 e 4, pudemos realizar a comparação com o TAP com humanos, utilizando, para tanto, apenas os estímulos, cujas respostas eram

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

uma das bases dos *blends*. Assim sendo, o Gráfico 5, que descreve as respostas por estrutura de constituintes, mostra que os LLMs seguiram a mesma tendência de resposta dos humanos. Isto é, os LLMs responderam mais a B1 para estrutura CP e mais a B2 para a estrutura PC. Contudo, embora essa tendência possa ser observada, as porcentagens das respostas dos LLMs são mais baixas em relação às respostas dos humanos.

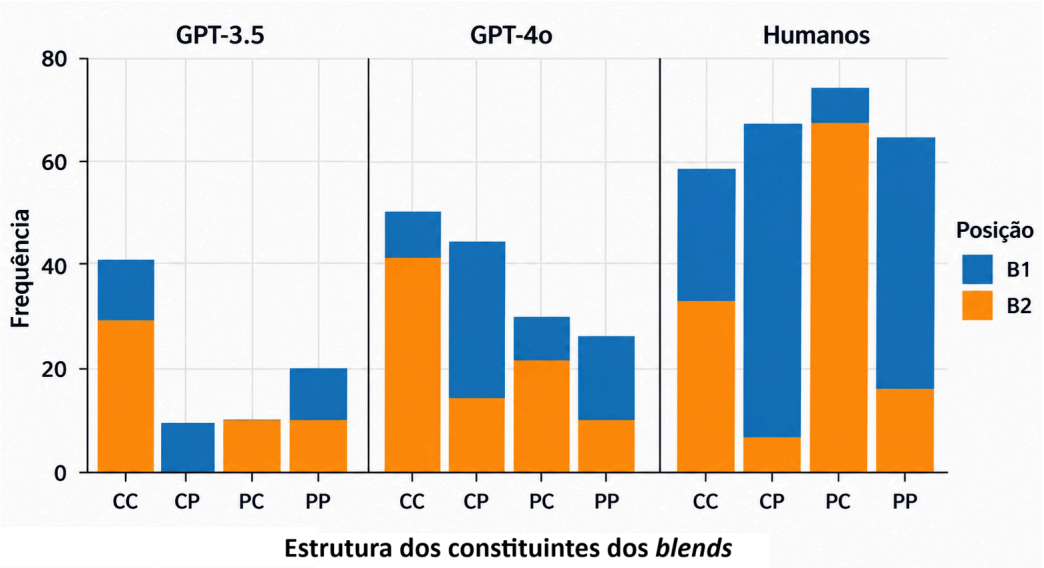

Gráfico 5 - Respostas (%) da B1 e B2 por respondente em relação à estrutura dos constituintes do *blend*.

Fonte: elaboração própria

Em relação à estrutura gramatical, também observamos nos LLMs a tendência de respostas dos humanos. O GPT- 3,5 é categórico, na estrutura NM, só há respostas com a B1 e, na estrutura MN, só há respostas com a B2. Para o GPT- 4o a mesma tendência, embora não categórica, é clara, assim como nos humanos, ressaltando, novamente, que as respostas dos humanos têm uma porcentagem mais alta que a dos LLMs estudados.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

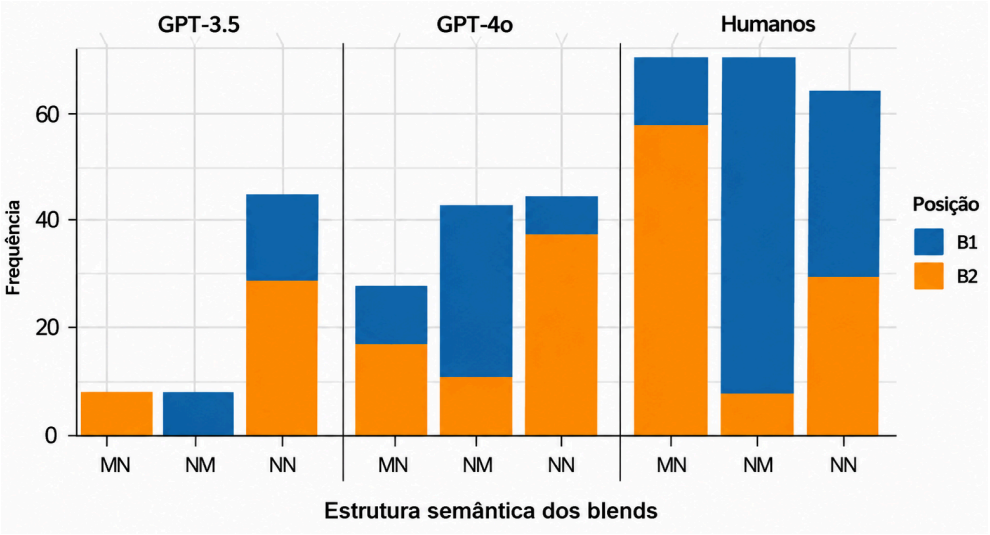

Gráfico 6 - Respostas (%) na base 1 ou 2 por respondente e estrutura semântica do *blend*.

Fonte: elaboração própria.

Para não descartamos a respostas que não mencionaram nem a B1 nem a B2, realizados uma análise qualitativa dessas respostas levando em consideração relações semânticas (metonímia, hiperonímia, sinonímia, campo semântico e definição), relação morfológica (raiz ou radical) e relação ortográfica, quando a palavra tinha um conjunto de letras parecidas, mas nenhuma relação semântica ou morfológica. A partir dessa análise, podemos perceber, curiosamente, as mesmas tendências, que podem ser observadas no Gráfico 7, o mapa de calor. Para a estrutura CP as relações estavam ligadas à B1, enquanto para a estrutura PC as relações estavam ligadas à B2. No que tange a estrutura CC, houve relações ligadas tanto a B1 quanto a B2, e, para as relações observadas para a estrutura PP foram mais dispersas.

	CC	CP	PC	PP
B1	0	8	0	4
B2	0	0	7	2
B1 e B2	6	1	0	2

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Gráfico 5 - Mapa de calor das respostas alternativas por relação com a base e estrutura sintática.

Fonte: elaboração própria.

De modo geral, observamos os seguintes resultados no TAP com máquinas: os modelos GPT, especialmente o GPT-3.5, apresentam menor consistência e maior variação aleatória em comparação aos humanos. Ademais, os humanos exibem tendências estruturais claras na escolha da base predominante, o que pode refletir uma sensibilidade semântica e cognitiva mais refinada ao tipo de constituintes dos *blends*. Por fim, a discrepância entre humanos e LLMs sugere que, apesar dos avanços, os LLMs ainda não reproduzem integralmente o padrão humano no processamento de palavras do tipo *blend*.

Além disso, os resultados revelam que os modelos GPT ainda não captam plenamente a lógica semântica subjacente ao comportamento humano na seleção da base predominante nos *blends*. Enquanto os humanos seguem padrões estáveis e previsíveis, os LLMs — sobretudo o GPT-3.5 — apresentam variações que sugerem um tipo de processamento mais superficial ou estatístico, em vez de conceitual. Isso reforça a ideia de que, apesar dos avanços, a cognição de máquina das IAs ainda apresentam dificuldades em avançar em um tipo de processamento de texto que leve em consideração informações essencialmente linguísticas e não apenas aquelas mais formais, o que se assemelha à discussão de Bender e Koller (2020) sobre LLMs e aprendizado de significado.

5. Considerações finais

O presente estudo investigou as capacidades de processamento morfológico de Grandes Modelos de Linguagem, com foco na interpretação de *blends* do português brasileiro. Para isso, foi realizado um experimento de associação de palavras com 40 *blends*, classificados em quatro categorias conforme sua estrutura constitutiva (CP, PC, CC e PP). O experimento foi aplicado a participantes humanos e aos modelos GPT-3.5 e GPT-4o, cujas respostas foram analisadas considerando a identificação das bases dos *blends* e as relações semânticas, morfológicas, ortográficas e gramaticais estabelecidas. Os resultados foram comparados para investigar semelhanças e diferenças entre o processamento humano e o processamento realizado pelos modelos. A análise mostrou que os LLMs apresentam alguma

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

capacidade de interpretar essas palavras complexas, mas seu desempenho é menos consistente e preciso do que o humano, evidenciando diferenças qualitativas no processamento linguístico.

Em síntese, os resultados deste estudo indicam que embora o LLM ChatGPT seja capaz de processar *blends* do português brasileiro e estabelecer relações entre suas bases constitutivas, seu desempenho apresenta diferenças significativas em relação ao processamento humano, especialmente no que diz respeito à precisão e à consistência das respostas. Essa constatação nos leva a uma questão mais ampla sobre a natureza da cognição de máquina: em que medida esses modelos conseguem mobilizar conhecimentos que ultrapassam as regularidades formais e estatísticas da língua, incorporando também informações de natureza pragmática e aspectos relacionados à intencionalidade dos falantes? Considerando que uma característica fundamental da cognição humana é a capacidade de estabelecer interfaces entre diferentes domínios do conhecimento — como sintaxe, semântica, pragmática e cognição social —, torna-se relevante investigar se, e em que condições, os sistemas de inteligência artificial conseguem estabelecer relações semelhantes. Assim, mais do que verificar se as máquinas são capazes de processar determinados fenômenos linguísticos, cabe compreender a natureza das operações envolvidas nesse processamento e em que medida elas se aproximam ou se distinguem dos mecanismos cognitivos humanos.

Por fim, nosso estudo reforça a necessidade de um aprofundamento teórico e epistemológico nas discussões sobre IA, linguagem e cognição — não apenas no campo da Linguística, mas também no âmbito mais amplo das Ciências Cognitivas. Acreditamos que, para que esse debate avance de forma adequada no contexto de uma sociedade composta por humanos e transumanos, é imprescindível que pesquisas futuras se concentrem na modelagem de sistemas competentes em cognição social e processamento de linguagem. Somente assim, será possível compreender com maior clareza as possibilidades de desenvolvimento cognitivo em inteligências artificiais.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Referências

ALBANO, E. A linguística e a psicologia inconfessas sob os modelos de linguagem. Jornal CGN, 25 dez. 2023. Ciência e tecnologia. Disponível em: https://jornalggn.com.br/ciencia/linguistica-e-psicologia-inconfessas-sob-os-modelos-de-linguagem/. Acesso em: mar. 2024.

BAXI, Jatayu; BHATT, Brijesh. Recent advancements in computational morphology: A comprehensive survey. arXiv preprint arXiv:2406.05424, 2024.

BENDER, E. M.; KOLLER, A. Climbing towards NLU: On meaning, form, and understanding in the age of data. In: Proceedings of the 58th annual meeting of the association for computational linguistics, p. 5185-5198, 2020.

CHICHE, A.; YITAGESU, B. Part of speech tagging: a systematic review of deep learning and machine learning approaches. Journal of Big Data, v. 9, n. 1, p. 10, 2022.

CHOMSKY, N. Syntactic structures. 2. ed. Berlin: Mouton de Gruyter, 2002.

CHOMSKY, N; ROBERTS, I; WATUMULL, J. The false promise of ChatGPT. The New York Times, 2023. Disponível em: https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html . Acesso em: 03 mar 2024.

DAMÁSIO, A. Sentir e saber: As origens da consciência. São Paulo: Companhia das Letras, 2022.

DE WAAL, F. Somos inteligentes o bastante para saber quão inteligentes são os animais? Rio de Janeiro: Zahar, 2021.

FIGUEIREDO, T. T. Linguagem e Teoria da Mente de segunda ordem: investigando os estados mentais de ignorância e crença falsa. Dissertação (Mestrado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas (UNICAMP), Campinas, 2018.

FIGUEIREDO, T. T.. Entre linguagem e emoções: uma investigação sobre estados emocionais no português brasileiro. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas (UNICAMP), Campinas, 2025.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

FREITAG, R. M. K.; DE GOIS, T. S. Performance in a dialectal profiling task of LLMs for varieties of Brazilian Portuguese. arXiv preprint arXiv:2410.10991, 2024.

HOLANDA, G. M.; PFEIFFER, C. R. C. Sentimento da inteligência artificial: novas tecnologias, antigos conceitos. Campinas, SP: Pontes Editores, 2023.

HOUGHTON, C.; KAZANINA, N.; SUKUMARAN, P. Beyond the limitations of any imaginable mechanism: large language models and psycholinguistics. arXiv preprint arXiv:2303.00077, 28 fev. 2023. Disponível em: http://arxiv.org/abs/2303.00077.

JURAFSKY, D.; MARTIN, J. H. Speech and language processing. London: Pearson, 2026.

KOSINSKI, M. Theory of mind may have spontaneously emerged in large language models. arXiv preprint arXiv:2302.02083, 2023. Acesso em: mar. 2023.

KOSINSKI, M. Evaluating large language models in theory of mind tasks. Proceedings of the National Academy of Sciences, v. 121, n. 45, p. e2405460121, 2024.

LANGLEY, C.; CÎRSTEA, B. I.; CUZZOLIN, F.; SAHAKIAN, B. J. Theory of Mind in Humans and in Machines. Frontiers in Artificial Intelligence, v. 5, p. 917565, 2022.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. Nature, v. 521, n. 7553, p. 436-444, 2015.

LI, D.; YIN, J.; ZHANG, T.; HAN, W.; BAO, H. The four most basic elements in machine cognition. Data Intelligence, v. 6, n. 2, p. 297-319, 2024.

LORANDI, Aline. Formas morfológicas variantes na aquisição da morfologia: evidências da sensibilidade da criança à gramática da língua. Letrônica, Porto Alegre, v. 3, n. 1, p. 81-97, 2010.

MAN, K.; DAMASIO, A. Homeostasis and soft robotics in the design of feeling machines. Nature Machine Intelligence, v. 1, n. 10, p. 446-452, 2019.

MAO, Y.; LIU, S.; ZHAO, P.; NI, Q.; LIN, X.; HE, L. A Review on Machine Theory of Mind. arXiv preprint arXiv:2303.11594, 2023. Acesso em: jan. 2024.

MARANGONI JUNIOR, C. E. A blenditividade na formação de palavras: a derivação dos blends na interface entre morfologia, fonologia e pragmática. Dissertação (Mestrado em

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo (USP), São Paulo, 2021.

MARCHETTI, A.; DI DIO, C.; CANGELOSI, A.; MANZI, F.; MASSARO, D. Developing ChatGPT's Theory of Mind. Frontiers in Robotics and AI, v. 10, p. 1189525, 2023.

MINUSSI, R. D.; NÓBREGA, V. A. A interface sintaxe-pragmática na formação de palavras: avaliando os pontos de acesso da Enciclopédia na arquitetura da gramática. Veredas, v. 18, n. 1, p. 161-184, 2014.

MINUSSI, R. D.; VILLALVA, A. Reconhecimento e acesso lexical dos blends em Português Europeu e Português Brasileiro. Todas as Letras, v. 22, n. 1, p. 1-14, 2020.

MINUSSI, R. D.; VILLALVA, A. Reconhecimento e acesso lexical de blends em PB. Apresentação realizada no XV CELSUL (Círculo de Estudos Linguísticos do Sul), 2022.

MINUSSI, R. D.; VILLALVA, A. Descrição de um corpus para os blends do português: classificação e padrões estruturais. In: RASIA, Gesualda; NEGRI, Lígia; RODRIGUES, Patrícia (orgs.). CELSUL 25 anos: práticas linguageiras e gramaticais. Campinas, SP: Mercado de Letras, p. 187-208, 2024.

MITCHELL, J. et al. Do LSTMs know about Principle C? Conference on Cognitive Computational Neuroscience, p. 188-191, 2019. Disponível em: https://ccneuro.org/2019/Papers/ViewPapers.asp?PaperNum=1241.

MITCHELL, M. How do we know how smart AI systems are? Science, v. 381, n. 6654, p. eadj5957, 2023.

MURPHY, E. et al. Fundamental principles of linguistic structure are not represented by ChatGPT. Biolinguistics, v. 19, p. 1-55, 2025.

PREMACK, D.; WOODRUFF, G. Does the chimpanzee have a theory of mind. Behavioural and Brain Science, v. 1, p. 515-526, 1978.

ROCHA, M.; DA SILVA, H. H.; MORALES, A. S.; SARKADI, S.; PANISSON, A. R. Applying theory of mind to multi-agent systems: a systematic review. In: Brazilian Conference on Intelligent Systems, Belo Horizonte, MG. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2023. p. 367-381.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

STRACHAN, J. W. A.; ALBERGO, D.; BORGHINI, G.; PANSARDI, O.; SCALITI, E.; GUPTA, S.; SAXENA, K.; RUFO, A.; PANZERI, S.; MANZI, G.; GRAZIANO, M. S. A.; BECCHIO, C. Testing theory of mind in large language models and humans. Nature Human Behaviour, v. 8, p. 1285-1295, 2024.

ULLMAN, T. Large language models fail on trivial alterations to theory-of-mind tasks. arXiv preprint arXiv:2302.08399, 2023. Acesso em: dez. 2023.

VERMA, M.; BHAMBRI, S.; KAMBHAMPATI, S. Theory of Mind abilities of Large Language Models in Human-Robot Interaction: An Illusion? Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, p. 36-45, 2024.

VILLALVA, A.; MINUSSI, R. D. Description and analysis of a Portuguese blend corpus. Corpus, n. 23. Disponível em: https://doi.org/10.4000/corpus.6436. 2022.

XU, Hanyi et al. Large language models for education: a survey. arXiv preprint arXiv:2405.13001, 2024.

ZHOU, P.; MADAAN, A.; POTHARAJU, S. P.; GUPTA, A.; MCKEE, K. R.; HOLTZMAN, A.; PUJARA, J.; REN, X.; MISHRA, S.; NEMATZADEH, A.; UPADHYAY, S.; FARUQUI, M. How far are large language models from agents with theory-of-mind? arXiv preprint arXiv:2310.03051, 2023. Acesso em: dez. 2023.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Contribuição de autoria

Thuany Teixeira de Figueiredo: Conceitualização; Investigação; Metodologia; Análise formal; Redação – rascunho original; Redação – revisão e edição.

Rafael Dias Minussi: Curadoria de dados; Análise formal; Investigação; Metodologia; Redação – rascunho original.

Rafael Luis Beraldo: Curadoria de dados; Investigação; Metodologia; Software.

Conflito de interesses

Os autores declaram que não há conflitos de interesses financeiros, pessoais, acadêmicos ou institucionais que possam ter influenciado a realização da pesquisa, a análise e interpretação dos resultados ou a elaboração deste manuscrito.

Declaração de disponibilidade de dados

Os dados e materiais utilizados na realização desta pesquisa, incluindo as informações relativas aos estímulos e às respostas obtidas nas tarefas experimentais com os participantes humanos e os modelos de linguagem, estão disponíveis publicamente no repositório GitHub, por meio do endereço: https://github.com/rberaldo/blends-LLMs.

Declaração de uso de inteligência artificial

Os autores declaram que utilizaram ferramentas de inteligência artificial, especificamente modelos de linguagem de grande escala (LLMs), como parte integrante do procedimento metodológico e experimental desta pesquisa, em consonância com os objetivos propostos. O uso dessas ferramentas, incluindo a elaboração e a aplicação de estímulos e a coleta e análise das respostas geradas pelos modelos, foi explicitamente descrito no corpo do artigo. Ressalta-se, portanto, que a utilização da inteligência artificial constituiu objeto de investigação do estudo, tendo sido realizada de forma intencional e alinhada à proposta experimental.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/SciELOPreprints.18798

Declaração de Aprovação do Comitê de Ética

O estudo experimental envolvendo participantes humanos, descrito na seção de Metodologia deste artigo, foi submetido e devidamente aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de São Paulo (UNIFESP). A aprovação foi concedida no âmbito do projeto de pesquisa intitulado "Word Processing – The Interplay of Morphological Complexity and Lexical Access (Processamento de palavras – a relação entre complexidade morfológica e acesso lexical)", registrado sob o Certificado de Apresentação de Apreciação Ética (CAAE) nº 38097520.2.0000.5505.

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença Creative Commons CC-BY.
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.