Revista EaD &
tecnologias digitais na educação

Processamento de Sentenças Ambíguas e a Preferência pela Primeira Menção: um estudo off-line entre falantes do Vale do Jiquiriçá

Karolaine Figuerêdo de Souza (UFRB)
https://orcid.org/0009-0009-4974-2789
karolfigueredo@aluno.ufrb.edu.br

René Alain Santana de Almeida (UFRB)
https://orcid.org/0000-0002-9288-0740
renealain@ufrb.edu.br

Resumo: O presente artigo objetiva analisar como o fator estrutural da primeira menção (first-mention) contribui para o processamento de sentenças ambíguas por falantes nativos do português brasileiro residentes no Vale do Jiquiriçá, região do estado da Bahia. Para tanto, o paradigma experimental utilizou a metodologia off-line, que captura respostas após o processamento completo dos estímulos e permitiu uma análise detalhada das estratégias cognitivas. Baseado em estudos anteriores de autores como Cozijn et al. (2011) e Kaiser (2011), explorou-se a tendência de interpretar sentenças ambíguas com o pronome referindo-se ao primeiro protagonista mencionado na oração. Foram aplicados experimentos com frases como "Vinicius e Leonardo amaram a festa. Ele dançou radiante", para verificar as interpretações feitas pelos participantes em relação ao pronome anafórico. O questionário utilizado apresentou uma escala tipo Likert de três pontos para medir a concordância dos participantes em cada interpretação e incluiu, além das sentenças experimentais, sentenças verificadoras de atenção. Um teste piloto foi realizado previamente para ajustes metodológicos. Os resultados mostraram uma tendência significativa dos participantes em optar pelo primeiro protagonista mencionado, confirmando a hipótese inicial. Entretanto, houve casos em que o tamanho do nome dos protagonistas pareceu influenciar a decisão, com preferência por nomes mais curtos, em algumas situações. Esses achados reforçam a hipótese de que falantes de português brasileiro residentes no Vale do Jiquiriçá optam pelo primeiro protagonista mencionado ao processarem sentenças ambíguas.

Palavras-chave: Processamento de Sentenças. Ambiguidade. Método off-line.

Abstract: This paper aims to analyze how the first-mention structural factor contributes to the processing of ambiguous sentences by native speakers of Brazilian Portuguese living in the Jiquiriçá Valley, a region in the state of Bahia. To this end, the experimental paradigm used the off-line methodology, which captures responses after the stimuli have been fully processed and allows for a detailed analysis of cognitive strategies. Based on previous studies by authors such as Cozijn et al. (2011) and Kaiser (2011), the tendency to interpret ambiguous sentences with the pronoun referring to the first protagonist mentioned in the sentence was explored. Experiments were carried out with sentences such as "Vinicius and Leonardo loved the party. He danced radiantly", to check the interpretations made by the participants in relation to the anaphoric pronoun. The questionnaire used a three-point Likert scale to measure the participants' agreement with each interpretation and included, in addition to the experimental sentences, attention verifier sentences. A pilot test was carried out beforehand for methodological adjustments. The results showed a significant tendency for participants to opt for the first protagonist mentioned, confirming the initial hypothesis. However, there were cases in which the length of the protagonists' names seemed to influence the decision, with a preference for shorter names in some situations. These findings reinforce the hypothesis that Brazilian Portuguese speakers living in the Jiquiriçá Valley opt for the first mentioned protagonist when processing ambiguous sentences.

Keywords*:* Sentence processing. Ambiguity. Off-line method.

1 INTRODUÇÃO

A presente pesquisa integra o projeto de pesquisa intitulado "*Fatores estruturais e prosódicos no processamento linguístico de sentenças"*, focando de modo mais específico no impacto do fator estrutural da primeira menção no processamento de sentenças ambíguas. Além disso, ressaltamos que, o processamento linguístico de sentenças ambíguas vem sendo objeto de estudo em diversas línguas como, por exemplo, o estudo sobre a resolução de pronomes ambíguos realizado por Jarvikivi et al (2005) para o finlandês; o estudo sobre o processamento de pronomes realizado por Cozijn et al (2011), para o holandês; também existem estudos realizados por Kaiser (2011) que buscam identificar os efeitos da subjetividade, da pronominalização e do foco contrastivo na interpretação de pronomes no discurso subsequente com falantes nativos do inglês.

Além disso, estudos também foram realizados no Português Brasileiro (PB) por autores como Almeida (2017), que estudou a influência da prosódia da fala na resolução de ambiguidade sintática, através de um estudo de processamento de sentença; Fonseca (2008), com um estudo sobre processamento de sentenças ambíguas do tipo "SN1-V-SN2- Atributo" do PB e; Myamoto (2005), ao estudar orações relativas ambíguas e a homogeneidade do processamento de sentenças. Além dessas pesquisas, o PB é contemplado também em estudo realizado por Maia (2013), dedicado a investigar diferenças na aferição *on-line* e *off-line* de construções que são estruturalmente ambíguas em português. O estudo envolveu os princípios da Aposição Mínima (*Minimal Attachment*) e da Aposição Local (*Late Closure*). Leitão (2005) também pesquisou sobre temática semelhante para o PB, ao objetivar mostrar que os pronomes lexicais e os objetos nulos são capazes de reativar seus respectivos antecedentes, buscando comprovar também que em PB os pronomes lexicais são processados mais rapidamente do que nomes repetidos.

Para realização do presente trabalho, partimos da ideia de ambiguidade enquanto um dos diversos fatores que são inerentes à língua. Almeida, Oliveira Jr e Cozijn (2021) ressaltam que estudos sobre a ambiguidade de sentenças são significativos porque fornecem entendimento sobre as decisões do *parser*[1], proporcionando melhor compreensão sobre o processamento de frases.

Estudos sobre processamento de sentenças podem ser realizados a partir de duas metodologias: *on-line* ou *off-line*. A primeira refere-se a captura de dados provenientes da reação dos participantes em tempo real, durante a leitura ou escuta dos estímulos, podendo ser usado nessa metodologia o rastreador ocular[2], por exemplo. Ao contrário, na metodologia *off-line*, os dados analisados são referentes às reações dos participantes após os estímulos, quando já houve o processamento e consequentemente a assimilação de todos os níveis linguísticos.

Na metodologia *off-line* as reações dos participantes são registradas após o processamento linguístico ter sido concluído, ou seja, quando os participantes já processaram todas as informações dos estímulos e formaram uma resposta. No presente estudo, utilizamos a abordagem *off-line* de processamento da linguagem por se tratar de uma metodologia que permite uma avaliação detalhada das estratégias de processamento, bem como das decisões dos participantes.

Nesse sentido, este estudo pode contribuir para o avanço da área de pesquisa sobre processamento da linguagem humana, fornecendo *insights* significativos sobre os processos cognitivos que são subjacentes à compreensão da linguagem. Além disso, foi escolhido realizar o estudo com falantes nativos do português brasileiro que residem no Vale do Jiquiriçá-Bahia, uma região que possui carência de estudos na área.

Partindo do pressuposto de que existem fatores estruturais e semânticos que compõem o processo de formação e compreensão da linguagem (Kaiser, 2013; Almeida, Oliveira Jr e Cozijn, 2021), o presente estudo objetiva analisar como o fator estrutural da primeira menção (*first-mention*) contribui no processamento de sentenças ambíguas por falantes do PB residentes no Vale do Jiquiriçá-Bahia.

Baseados em estudos prévios (Jarvikivi et al, 2005, para o filandês; Kaiser, 2011, para o inglês; Cozijn et al, 2011, para o holandês; Hartshorne et al, 2014, para o inglês ; Almeida, 2017, para o PB; Maia, 2013, para o PB; Leitão, 2005, para o PB), a nossa hipótese é que para desambiguar a sentença, o participante opte preferencialmente pelo primeiro protagonista mencionado. Logo, em sentenças do tipo: “Maria e Laura comeram o bolo. Ela passou mal.” o participante concluiria, diante dessa ambiguidade sintática[3], que quem passou mal foi Maria e não Laura, uma vez que Maria foi a primeira mencionada na oração.

[1]O termo *parser* tem origem no latim e se refere a procedimentos mentais que estabelecem a estrutura de uma frase, estruturando os itens lexicais de forma hierárquica, não tendo acesso às informações semânticas ou pragmáticas (Maia e Finger, 2005)

[2]Ferramenta de monitoramento da posição dos olhos humanos enquanto se está envolvido em tarefas que requerem a visualização de estímulos visuais, como textos escritos, por exemplo (Forster, 2017).

[3] A ambiguidade sintática ocorre quando pode haver mais de uma maneira de combinar os elementos da sentença (Basso et al, 2009).

2 PROCESSAMENTO DE SENTENÇAS

O estudo de processamento de sentenças tem como objetivo principal a compreensão de como os falantes processam e entendem o significado de sentenças. De acordo com Myamoto (2005), um dos pontos básicos no que diz respeito ao estudo de processamento de sentenças é determinar se existe variação no processamento de diferentes línguas.

> Uma questão básica no processamento de sentenças é a de determinar se existe variação no processamento de línguas distintas. Mais precisamente, assuma que o processamento de sentenças envolve dois componentes básicos: a gramática de uma dada língua (e outros conhecimentos tais como normas sociais, conhecimentos gerais) e um algoritmo que usa a gramática para processar sentenças. (Myamoto, 2005, p. 72)

O que diferencia a gramática do algoritmo citados pelo autor é que, enquanto a gramática integra o conhecimento possuído pelos falantes nativos, o algoritmo busca compreender como a gramática é utilizada quando as sentenças são processadas. O estudo do processamento de sentenças busca entender se cada língua possui um algoritmo diferente, questionando se deve ou não ser padronizado da mesma forma que as gramáticas são.

> Na investigação sobre a natureza das estratégias utilizadas pelos seres humanos no processamento de sentenças, um dos maiores desafios tem sido explicar como os indivíduos (leitores ou ouvintes) resolvem questões de ambigüidade. Em outras palavras, o que se quer saber é como se dá, nas línguas individuais, a interpretação de uma sentença em situações em que há pelo menos duas alternativas possíveis e se esses mecanismos de processamento operam da mesma forma nas diferentes línguas. (Finger e Zimmer, 2005, p. 112)

Na busca por identificar como ocorre o processamento de sentenças na mente humana, um dos desafios é compreender como as questões de ambiguidade são resolvidas, assimilando os mecanismos cognitivos e linguísticos que se confrontam quando é necessário que o indivíduo decida qual interpretação realizar em sentenças ambíguas, não somente em uma língua específica, mas também em diversas línguas, a fim de entender se é um processo variável ou universal.

Para o estudo do processamento de sentenças, duas metodologias experimentais são consideradas mais adequadas: metodologia *on-line* e metodologia *off-line*. A escolha da metodologia utilizada pode influenciar nos dados obtidos, bem como na compreensão do processamento linguístico.

> [...] i) a metodologia off-line, em que as reações dos participantes são coletadas após os mesmos terem lido ou ouvido os estímulos, ou seja, no momento em que o processamento já foi concluído e já houve uma integração entre todos os níveis linguísticos; e ii) a metodologia on-line, em que as reações dos participantes são capturadas no momento em que eles estão lendo ou ouvindo os estímulos, praticamente simultâneas ao processamento. (Almeida, 2017, p. 9)

Na metodologia *off-line* os participantes são expostos a estímulos através de áudio ou até mesmo texto escrito, para somente depois que o processamento já tiver sido concluído responderem ao questionamento, quando houve uma integração entre todos os níveis linguísticos. Ao contrário, na metodologia *on-line* as reações dos participantes

são capturadas no mesmo momento em que estão recebendo os estímulos, quando ainda não houve uma integração entre todos os níveis linguísticos.

Existem alguns fatores que podem influenciar na produção e compreensão de frases durante o processamento da linguagem, dentre esses fatores, os mais estudados são os fatores estruturais das sentenças, como por exemplo: o fator da proximidade (l*ate closure*), que pode ser explicado através da Teoria do Garden Path (TGP), primeira menção (*first-mention*) e condição de sujeito (*subjecthood*).

2.1 TEORIA GARDEN PATH

A teoria *Garden Path* (TGP) refere-se a um dos modelos teóricos que busca explicar como ocorre o processamento e entendimento da linguagem na mente humana. Em 1992, Dillinger traduziu o termo para português como "Teoria do Labirinto", levando em consideração que as estratégias para o processamento de sentenças na nossa mente nem sempre são feitas de forma assertiva, em meio a diversas possibilidades de escolha.

> A metáfora do garden path ou "caminho do jardim" é basicamente semelhante a do labirinto. Trata-se de um modelo estrutural e o labirinto, à semelhança de uma frase, é uma estrutura, com várias bifurcações a serem escolhidas ao se trafegar por ele. Ao se entrar em uma sala em que há várias portas, escolhe-se uma delas, provavelmente a mais próxima e, algumas vezes, a escolha leva para fora, ao jardim, e não ao interior da estrutura, como pretendido. (Maia e Finger, 2005, p. 17)

A mente humana funciona como uma espécie de labirinto, existindo uma infinidade de possibilidades de entradas no *parser*, quando é feita uma interpretação errônea é necessário voltar e escolher um novo caminho nesse labirinto, para que a interpretação seja adequada.

> A TGP propõe como postulados fundamentais que: (1) há um processador sintático autônomo ou parser, que usa uma porção do seu conhecimento gramatical isolado do conhecimento de mundo e outras informações para a identificação inicial das relações sintagmáticas; (2) o parser confronta-se com sintagmas de aposição ambígua e compromete-se com uma estrutura única; (3) pressionado pelo sistema de memória de curto prazo, que tem um limite estreito de computação e armazenamento, o parser segue princípios psicológicos de minimalidade e de localidade na escolha desta estrutura preferencial: use o menor número possível de nós (Princípio da Aposição Mínima) e, se duas aposições mínimas existem, aponha cada nova palavra ao sintagma corrente (Princípio da Aposição Local). (Maia, 2010, p. 12)

Entre os princípios estudados dentro da teoria de *Garden Path* estão o princípio da aposição local (*Late Closure*) e o princípio da aposição mínima (*Minimal Attachment*). Tais princípios se complementam e, portanto, não devem ser vistos isoladamente. O princípio da aposição mínima sugere que o *parser* utilize o mínimo de esforço ao realizar uma interpretação imediata, ou seja, quando surge um novo termo na oração, esse novo termo deve ser ligado (de forma mental) ao sintagma que estiver sendo construído, ou seja, ao sintagma mais próximo.

O princípio da aposição local determina que, em caso de sentenças ambíguas, a preferência interpretativa é pela palavra que se encontra mais próxima da palavra críti-

ca, que gera a ambiguidade. Os estudos voltados para o fator estrutural da proximidade (*Late Closure*) abarcam o princípio da aposição local da TGP, pesquisado por autores como Frazier (1979) e Pickering e Traxler, (1998). De acordo com Almeida, Oliveira Jr e Cozijn (2021), o princípio da aposição local assevera que, em caso de sentenças ambíguas, a preferência será pelo referente mais próximo da palavra crítica e, portanto, em sentenças do tipo "O curandeiro conduziu o seringueiro na selva bastante faminto", a preferência seria por apor a palavra crítica "faminto" ao "seringueiro", uma vez que, dentre os protagonistas mencionados, é o sintagma que está mais próximo do atributo.

2.2 FATOR ESTRUTURAL DA PRIMEIRA MENÇÃO (*FIRST-MENTION*)

Contrariamente ao *Late Closure*, achados de algumas pesquisas (Cozijn et al, 2011; Almeida, 2017, por exemplo) demonstram que durante o processamento de sentenças, o primeiro protagonista mencionado é processado na mente do falante de forma mais "fácil" do que o protagonista que é mencionado em seguida. Tais estudos salientam que a primeira menção de um referente desempenha um papel importante no processamento da linguagem, principalmente na resolução de expressões referenciais em sentenças.

Em outras palavras, quando dois ou mais protagonistas são mencionados em um texto e, em seguida, utiliza-se um pronome para fazer referência a um dos protagonistas mencionados, o *parser* opta pelo caminho mais "simples", associando o pronome ao primeiro protagonista mencionado, uma vez, que ao serem expostos a uma ambiguidade estrutural em uma sentença, os falantes tendem a escolher a interpretação mais simples ou acessível, porque assim minimizam o esforço cognitivo. Uma outra teoria existente é a de que, para o falante, o mais simples e acessível é o protagonista mais próximo ao pronome, e não o mais distante.

Para Cozijn et al (2011), a estrutura sintática da sentença influencia na preferência pelo sujeito, existindo uma tendência, em diversas línguas, a associar o referente de um pronome ao sujeito mencionado na oração anterior. Ou seja, Para Cozijn et al (2011), em orações do tipo "Ana viu Lara no aniversário porque ela foi convidada.", o falante associa que o pronome "ela" está se referindo a Ana, mesmo que o pronome também possa estar fazendo referência a "Lara". A preferência por "Ana" ocorre porque ela ocupa a posição de sujeito, enquanto "Lara" ocupa a posição de objeto direto. Trata-se de uma sentença ambígua porque a estrutura sintática da oração permite mais de uma interpretação. A preferência pelo sujeito funciona como uma estratégia para resolvê-la.

> A preferência pelo sujeito é uma preferência por atribuir o referente de um pronome para o sujeito da oração anterior, mesmo na ausência de pistas adicionais para fazer referência e onde interpretações alternativas são plausíveis, como apresentado nas pesquisas de Arnold et al (2000), Kameyama (1996) e Kaiser (2011) para o inglês; Järvikivi et al (2005) para o finlandês e Cozijn et al (2011) para o holandês. Nessas pesquisas, o sujeito é o referente preferido na maioria dos casos, seja em textos escritos ou locucionados, e, portanto, reúnem evidências de que uma preferência pelo favorecimento do sujeito como referente de um pronome parece ser muito eficaz. (Almeida, 2017, p. 16)

A condição de sujeito (*subjecthood*) influencia no processamento de sentenças, assim como a primeira menção, que por vezes coincide com a preferência pelo sujeito, entretanto, enquanto a preferência pelo sujeito está relacionada com a estrutura sintática, conforme citado anteriormente, a preferência pela primeira menção tem relação com a estrutura sequencial.

> É claro que condição de sujeito e primeira menção muitas vezes coincidem, por isso é difícil separar suas influências. Entretanto, elas se originam de fontes diferentes. A condição de sujeito resulta da sintaxe, enquanto a primeira menção resulta da semântica. Em algumas línguas, como inglês e holandês, o primeiro protagonista mencionado em uma frase também é frequentemente o sujeito dessa frase; portanto, a preferência pela primeira menção geralmente acompanha a preferência pelo sujeito. (Almeida, Oliveira Jr e Cozijn, 2021, p.5)

Portanto, o fator da primeira menção influencia na forma como as sentenças são interpretadas. No caso de preferência pela primeira menção, a interpretação seria de que o pronome colocado na segunda oração faz referência ao primeiro protagonista mencionado na oração anterior. Então, se o falante realiza a leitura ou escuta a seguinte sentença: “Carla e Bruna foram comprar roupas. Ela se sentiu bonita”, a preferência seria, portanto, optar por interpretar que Carla se sentiu bonita, uma vez que foi a primeira a ser mencionada na oração.

A predileção pelo primeiro protagonista mencionado é mais facilmente observada em sentenças com sujeito composto, isto é, em estruturas gramaticais que possuem dois ou mais substantivos que atuam como sujeitos de uma sentença, os quais são ligados pelas conjunções “e” ou “ou”. A partir do momento que surge um pronome para fazer referência a um único sujeito, dá-se origem a uma ambiguidade que possibilita diferentes interpretações. Em ocorrências desse tipo no inglês, estudos realizados com falantes nativos do idioma concluíram que, em grande parte dos casos, a preferência é pelo primeiro protagonista mencionado.

Gernsbacher e Hargreaves (1988), por exemplo, desenvolveram uma pesquisa sobre a ordem em que os participantes são mencionados em frases e como isso afeta a rapidez com que as palavras são reconhecidas. Os resultados evidenciaram que, em grande parte dos casos, os falantes possuem uma tendência maior em realizar a interpretação favorável ao primeiro protagonista mencionado, para solucionarem casos de ambiguidade. Os autores concluíram que os participantes mencionados primeiro são mais acessíveis cognitivamente porque eles formam a base para a representação mental da frase, e é por meio deles que as informações seguintes são integradas na representação em desenvolvimento.

> Concluímos que a vantagem da primeira menção não surge de nenhum dos fatores estéticos que investigamos. Nos sugerem, em vez disso, que a vantagem surge porque a compreensão requer a construção de uma representação ou estrutura mental. Isto implica estabelecer uma base e mapear informações subsequentes sobre essa base. Os participantes mencionados em primeiro lugar são mais acessíveis porque é através deles que as informações subsequentes são mapeadas na representação em desenvolvimento. (Gernsbacher e Hargreaves, 1988, p. 713)

De acordo com Gernsbacher e Hargreaves (1988), o primeiro protagonista mencionado não é processado de forma mais rápida devido a fatores estéticos e por sim processos cognitivos envolvidos na compreensão de informações, uma vez que, para que o falante compreenda uma frase, é necessário construir uma representação ou estrutura mental dessa informação e, ao mencionar um participante em primeiro lugar, ele se torna a base para a construção dessa estrutura mental.

Diante do exposto, é possível inferir que, no que diz respeito ao processamento linguístico e preferência pela primeira menção, a ordem em que os protagonistas são mencionados pode interferir no processamento de sentenças ambíguas com sujeito composto, como as que nos propomos a investigar no presente estudo.

3 METODOLOGIA

Para alcançar o objetivo de analisar como o fator estrutural da primeira menção (*first-mention*) contribui no processamento de sentenças ambíguas por falantes nativos do PB residentes no Vale do Jiquiriçá, região do estado da Bahia, foram utilizadas sentenças elaboradas em sua ordem canônica, SVO, seguidas da segunda oração com a estrutura: sujeito-verbo-predicativo do sujeito. Levamos em consideração estudos realizados por autores como Cozijn et al (2011), que realizou estudos com falantes do holandês e constatou uma predileção pela primeira menção. O nosso objetivo é verificar se o mesmo acontece no português brasileiro, conforme verificado em Almeida (2017). Para isso, foi realizado um teste experimental com falantes nativos do PB, residentes no Vale do Jiquiriçá-BA.

Para a elaboração das sentenças, houve, além do cuidado de seguirmos a ordem canônica do PB, a preocupação de utilizarmos os protagonistas de uma mesma sentença com o mesmo número de sílabas em ambos, para que o “tamanho” do nome não interferisse na escolha do participante, e assim descartamos qualquer possibilidade do participante escolher o protagonista que teoricamente tivesse o nome mais fácil de memorizar devido ao menor ou maior tamanho.

Ainda no que diz respeito às sentenças experimentais, essas foram escritas utilizando dois protagonistas do mesmo sexo, ao total foram elaboradas trinta sentenças do tipo “Vinicius e Leonardo amaram a festa. Ele dançou radiante.”, para que os participantes respondessem se quem havia dançado radiante foi o primeiro protagonista mencionado (Vinicius) ou o segundo protagonista mencionado (Leonardo).

Além das sentenças experimentais, foram elaboradas também 16 sentenças verificadoras de atenção do tipo “Marli e Marcelo fizeram o almoço. Ela comeu em casa.”, essas sentenças foram elaboradas seguindo a mesma estrutura sintática das experimentais, ou seja, ordem canônica do PB: SVO para a primeira oração e a segunda tendo a estrutura sujeito-verbo-adjunto adverbial. Optamos por utilizar para cada sentença um sujeito do sexo masculino e outro do feminino, a segunda sentença foi iniciada com pronome pessoal do caso reto (ele/ela) que fazia referência a um desses dois sexos, visto que no português brasileiro os pronomes marcam gênero. Para as sentenças verificadoras de atenção não foi necessário inserir protagonistas com o mesmo número de sílabas.

A pesquisa foi realizada através da metodologia *off-line* de processamento linguístico. Para isso, foi aplicado um questionário *on-line* com falantes nativos do estado da Bahia, através de um questionário com trinta sentenças experimentais seguidas de

mesma estrutura, além de sentenças verificadoras (sentenças não ambíguas), que são importantes para verificarmos a atenção dos participantes.

As opções de resposta foram apresentadas através de escala tipo Likert, metodologia desenvolvida por Rensis Likert por volta da década de 30, permitindo que os participantes entrevistados da pesquisa manifestem seus níveis de concordância ou discordância em relação aos questionamentos realizados, podendo inclusive expressar neutralidade. A metodologia foi escolhida por levarmos em consideração que não se pode simplesmente perguntar a quem o pronome está se referindo, esperando obter resposta como sujeito A ou B ou se a sentença é ambígua ou não.

> De acordo com Schütze e Sprouse (2013), escalas tipo Likert são apropriadas para testes de julgamento perceptual dessa natureza. Chelliah (2013) salienta ainda o cuidado metodológico necessário na elaboração desse tipo de questionário que envolve ambiguidade. Ela afirma que para confirmar os significados de uma construção ambígua não se pode simplesmente perguntar se uma sentença é ambígua para obter como respostas sim ou não. (Almeida, Oliveira Jr e Cozijn, 2021, p. 11)

Infere-se que as escalas tipo Likert são as mais apropriadas para estudos de tal natureza, uma vez que permite a avaliação de atitudes, opiniões e percepções dos participantes em relação a estímulos linguísticos específicos, possibilitando além da coleta de dados, mas também contribuindo significativamente para análises estatísticas mais sólidas.

Antes de distribuirmos a pesquisa ao público alvo, realizamos um teste piloto com amigos e familiares que não residiam no Vale do Jiquiriçá, a fim de avaliarmos se os métodos e procedimentos planejados funcionariam como esperado na prática. Além disso, a aplicação de um teste piloto possibilitou que possíveis falhas fossem detectadas, o que permitiu ajustes antes de iniciarmos a coleta de dados.

Para o teste piloto elaboramos uma lista no Google Forms, onde no primeiro momento o participante teve acesso a "seção de boas-vindas" com breve explicação de como o questionário funcionária, em seguida o participante foi direcionado para a "seção de prática", nesta seção o participante teve acesso a duas questões experimentais e uma verificadora de atenção. Ao finalizar a seção e compreender a dinâmica do formulário, o participante foi direcionado para as questões experimentais, neste momento as sentenças foram intercaladas da seguinte forma: uma sentença verificadora de atenção para duas experimentais.

Ademais, intercalamos a ordem dos protagonistas no momento do participante marcar, fizemos de forma alternada, inserindo em alguns casos o segundo protagonista da frase como primeira opção para marcar. No total, vinte pessoas participaram do teste e as páginas estavam travadas, só sendo possível acessar a pergunta seguinte ao responder a pergunta da atual página que estavam. Em nenhum momento souberam que se tratava de um teste piloto ou que possuía relação com a ambiguidade do sujeito.

Ao analisarmos os dados obtidos no teste piloto foi possível perceber que houve casos de desatenção, inferimos isso ao observar que ao decorrer das respostas alguns participantes responderam de forma equivocada as sentenças verificadoras de atenção, ainda que fosse claro a quem o pronome estava fazendo referência, marcando: I) os mesmos números para ambos os protagonistas; II) foram observados uma quantidade significativa de neutralidade, onde o participante não concorda nem discorda; III) inter-

pretou de forma equivocada as sentenças verificadoras e, além disso, não prestaram atenção a ordem do protagonista mencionado na pergunta, visto que inicialmente estavam, em grande parte, seguindo uma preferência pela primeira menção mas, quando a ordem dos protagonistas era invertida no questionamento, o segundo protagonista passou a ser o preferível.

Durante a análise dos gráficos, foi possível perceber que, inicialmente não houve diferença significativa entre as respostas, mas ao decorrer do formulário foi visível diferenças expressivas. Por se tratar de um teste aplicado a amigos e familiares, foi possível receber um *feedback* dos participantes, que relataram que o teste estava muito extenso e exaustivo.

Além disso, os participantes comentaram sentir dificuldade nas colunas de resposta, relatando estar confusa, uma vez que, para o teste piloto utilizamos uma escala com cinco opções de respostas:

1. Discordo totalmente;
2. Discordo;
3. Não concordo nem discordo;
4. Concordo;
5. Concordo totalmente.

Para que não houvesse confusão ao assinalar a coluna correspondente a cada protagonista, decidiu-se que o número de colunas da escala seria reduzido, limitando a uma escala com apenas três pontos:

1. Discordo;
2. Não concordo nem discordo;
3. Concordo.

Após análise dos dados obtidos no teste piloto e *feedback* dos participantes, concluiu-se que durante a pesquisa seria menos exaustivo ao participante se o teste estivesse dividido em duas etapas, a primeira contendo 15 sentenças experimentais e 15 sentenças verificadoras de atenção e a segunda etapa do teste da mesma forma. Para isso, precisamos elaborar mais duas sentenças experimentais, visto que, duas das trinta já elaboradas estavam direcionadas a seção de treino e não fariam parte da análise de dados.

Entre as alterações realizadas, houve a troca de nomes dos protagonistas da oração “Clara e Lara foram à festa. Ela dançou entusiasmada.” para “Mari e Lara foram à festa. Ela dançou entusiasmada.” porque percebemos que as protagonistas tinham nomes muito semelhantes e optamos por evitar que isso interferisse de alguma forma na interpretação dos participantes. Dessa forma, das trinta sentenças experimentais elaboradas, as quinze primeiras ficaram no primeiro questionário e as demais, foram direcionadas para o segundo formulário. As quinze sentenças verificadoras de atenção foram as mesmas para ambos os formulários.

Para o teste experimental tivemos o cuidado de inserir uma página após a seção de boas-vindas, onde o participante necessitou responder em qual região residia, caso não fosse o Vale do Jiquiriçá, o participante não conseguiria continuar respondendo o teste, sendo redirecionado para o final.

Como o teste experimental foi dividido em duas etapas, conforme relatado, para a análise de dados foram verificadas inicialmente as respostas individuais dos participantes a fim de verificar a quantidade de erros nas sentenças verificadoras de atenção. Os participantes que erraram acima de sete sentenças verificadoras de atenção foram des-

cartados da análise de dados, e posteriormente, as respostas dos dois questionários foram agrupadas.

4 RESULTADOS E DISCUSSÕES

Os links para acesso aos questionários foram direcionados para os falantes de PB, com foco nos falantes residentes no Vale do Jiquiriçá. O Quadro 1 apresenta o número de participantes de cada questionário.

Quadro 1 - Total de participantes dos formulários 01 e 02.

Formulário	Participantes (total)	Participantes (vale do jiquiriçá)
01	51	33
02	51	31

Fonte: Elaborado pelos próprios autores.

Ao total, cento e duas pessoas participaram. Destes, trinta e oito foram eliminados por não serem do Vale do Jiquiriçá e treze foram eliminados por desatenção. Os casos de desatenção foram percebidos pelos participantes não acertarem mais que sete questões verificadoras de atenção. Mesmo sendo claro a quem o pronome se referia, os participantes assinalaram: I) o mesmo ponto da escala para ambos os protagonistas II) julgou de forma errada mais que sete questões verificadoras de atenção, conforme mostra o Quadro 2.

Quadro 2 - Total de participantes eliminados - outras regiões e casos de desatenção.

Formulário	Outra região	Eliminados por desatenção
01	18	07
02	20	05

Fonte: Elaborado pelos próprios autores.

Dessa forma, restaram cinquenta e um participantes para a análise de dados, sendo vinte e cinco para o formulário 01 e vinte e seis para o formulário 02, conforme demonstrado no Quadro 3.

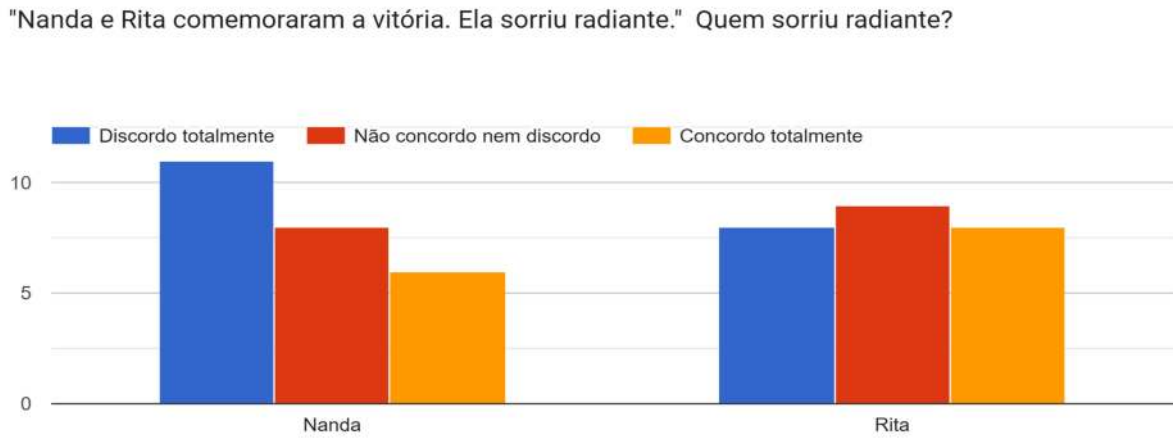
Quadro 3 - Total de respostas analisadas após as eliminações.

Formulário	Vale do jiquiriçá	Eliminados por desatenção
01	33	07
02	31	05
Total de respostas analisadas após eliminações		
Formulário 01		25 participantes
Formulário 02		25 participantes

Fonte: Elaborado pelos próprios autores.

Ao verificarmos as respostas dos participantes (formulários 01 e 02), foi possível observar que, em vinte e quatro das trinta sentenças experimentais os participantes tiveram uma predileção pelo primeiro protagonista mencionado. Ademais, no formulário 01, houve três casos de empate, os quais compreendemos que os participantes julgaram não haver predileção por nenhum dos protagonistas em suas interpretações. Houve também três ocorrências de preferência pela segunda menção. Entretanto, em dois desses três casos de preferência pela segunda menção, o protagonista escolhido tinha o nome menor, ainda que possuíssem a mesma quantidade de sílabas, conforme ocorrência apresentada na Figura 1.

Figura 1 - Ocorrência de preferência pela segunda menção.

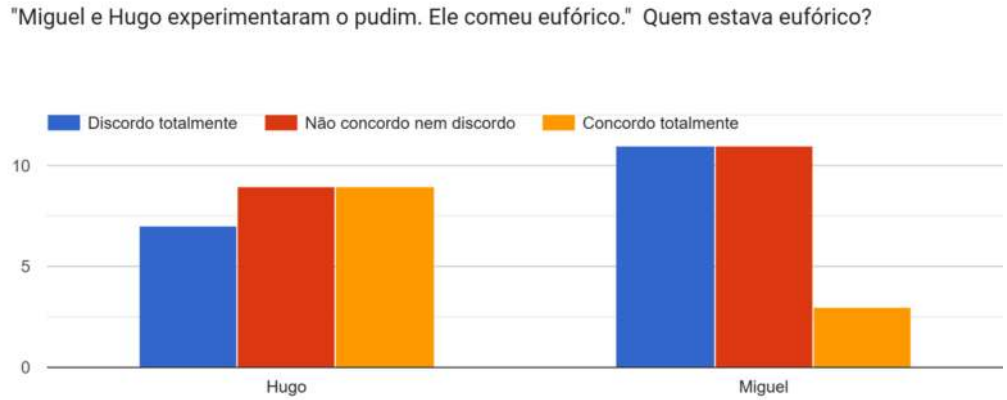


Fonte: Elaborado pelos próprios autores.

Ao analisarmos o gráfico é possível observar que na sentença “Nanda e Rita comemoraram a vitória. Ela sorriu radiante.”, a diferença entre a predileção pelo primeiro e segundo protagonista mencionado foi de apenas dois votos, onde oito dos participantes compreenderam que a protagonista Rita sorriu radiante, e seis participantes compreenderam que quem sorriu radiante foi Nanda. Além disso, oito participantes afirmaram não concordar nem discordar que Nanda, a primeira protagonista mencionada, sorriu radiante, enquanto nove participantes optaram por afirmar que não concordavam nem discordavam que Rita sorriu radiante.

De outro modo, na sentença "Miguel e Hugo experimentaram o pudim. Ele comeu eufórico.", apenas três participantes concordaram totalmente que Miguel, o primeiro protagonista mencionado, estava eufórico, enquanto nove concordaram totalmente que quem estava eufórico foi o segundo protagonista mencionado e, consequentemente, nesse caso, o protagonista com o menor nome no que diz respeito a quantidade de unidades gráficas. Ademais, nove participantes não concordaram nem discordaram que Hugo experimentou o pudim, enquanto onze participantes não concordaram nem discordaram que quem havia experimentado o pudim foi Miguel, conforme mostra o gráfico da Figura 2.

Figura 2 - Ocorrência de preferência pela segunda menção.

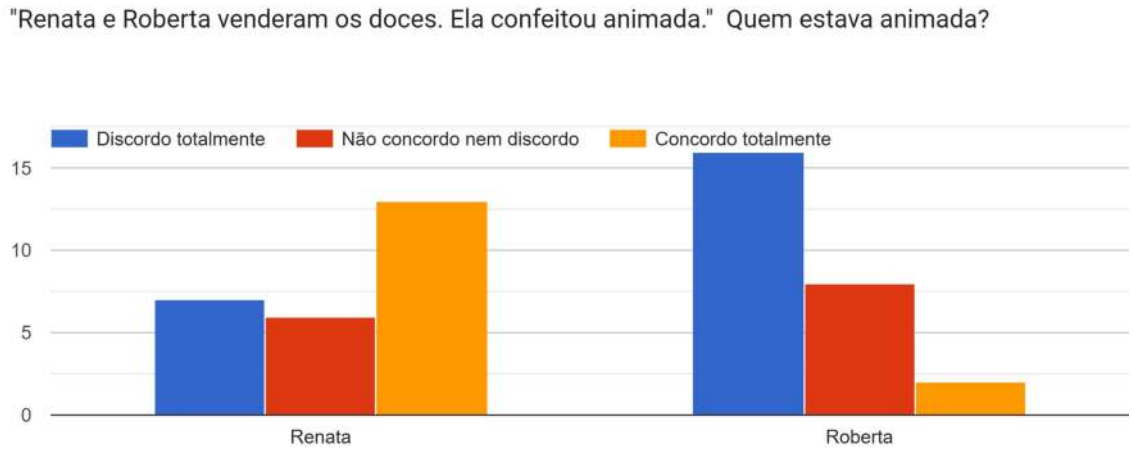


Fonte: Elaborado pelos próprios autores.

Talvez a escolha pelo segundo protagonista mencionado nas sentenças em questão tenha sido influenciada pela quantidade de letras que formavam o nome de cada protagonista.

No que concerne à preferência pela primeira menção, em alguns casos foi possível perceber que os participantes, em grande maioria, optaram pelo primeiro protagonista mencionado, como ilustrado pela ocorrência apresentada na Figura 3.

Figura 3 - Ocorrência de preferência pelo primeiro protagonista mencionado.

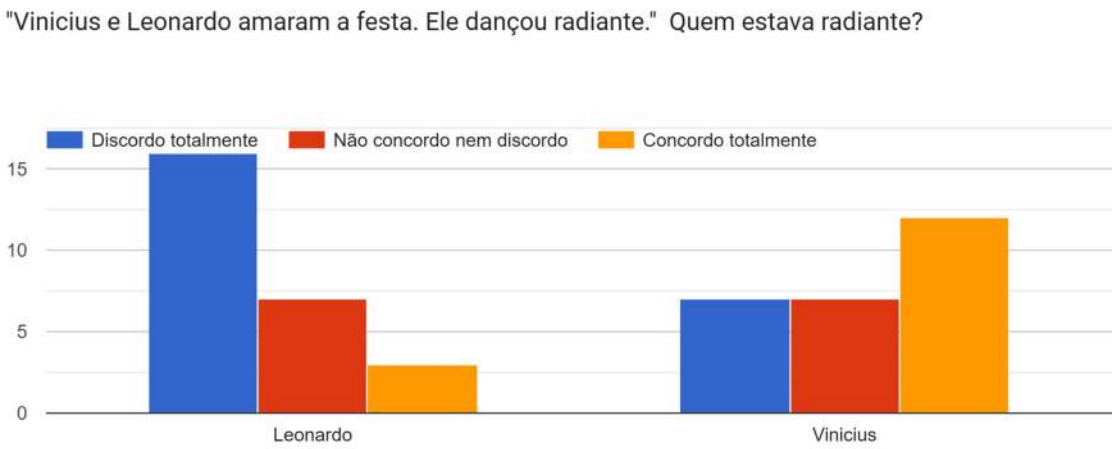


Fonte: Elaborado pelos próprios autores.

Trata-se de uma das sentenças experimentais utilizadas no formulário 02, onde nas quinze sentenças experimentais deste formulário, a preferência foi pelo primeiro protagonista mencionado. É possível perceber que houve uma diferença significativa no que diz respeito a predileção pelo primeiro protagonista, onde treze colaboradores concordaram totalmente que quem estava animada foi Renata, a primeira mencionada na oração, enquanto apenas três concordaram totalmente que Roberta quem estava animada, contra dezesseis participantes que discordaram totalmente que a segunda protagonista quem esteve animada.

Somado a essas questões, observamos também um dos casos em que inserimos o primeiro protagonista mencionado como segunda opção na escala a fim de verificar a atenção dos participantes e evitar a possibilidade de respostas repetitivas. A Figura 4 mostra uma das ocorrências.

Figura 4 - Exemplo de alternância na posição dos protagonistas nas respostas do questionário.



For
da
zendo referência ao primeiro protagonista mencionado (Vinicius), ao passo que, dos vinte e seis participantes, doze concordaram totalmente que Vinicius estava radiante, enquanto, dezesseis discordaram totalmente que Leonardo estava radiante e apenas três concordaram totalmente.

5 CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo analisar como o fator estrutural da primeira menção (*first-mention*) contribui no processamento de sentenças ambíguas por falantes nativos do PB residentes no Vale do Jiquiriçá, região do estado da Bahia. A pesquisa foi realizada utilizando a metodologia *off-line*.

Os resultados obtidos evidenciaram uma tendência expressiva dos participantes referente à preferência pelo primeiro protagonista mencionado, corroborando com as hipóteses sobre a primazia da primeira menção na interpretação de sentenças ambíguas, defendidos por autores como Cozijn et al (2011) e Kaiser (2011). Embora não tenhamos realizado testes estatísticos para verificar a significância dos resultados encontrados, reconhecemos a sua relevância e destacamos o desejo de utilizá-los em pesquisas futuras com um *corpus* maior.

Ademais, a análise das respostas demonstrou que o tamanho do nome dos protagonistas pode ter relevância na escolha dos participantes e, os nomes mais curtos podem ser os preferíveis em algumas situações, reforçando a importância de se levar em consideração fatores estruturais e semânticos em estudos de compreensão do processamento de sentenças ambíguas na mente humana. Entretanto, não foi foco dessa pesquisa verificar se o tamanho do nome referente ao protagonista interfere na interpretação de ambiguidade como a analisada no presente estudo, talvez uma sugestão para trabalhos futuros.

No que concerne à metodologia utilizada, incluindo a utilização do teste piloto e o uso da escala tipo Likert, estes se mostraram eficazes durante a coleta de dados, permitindo uma avaliação cuidadosa das escolhas feitas pelos participantes. Contudo, acreditamos que futuras investigações possam explorar outras variáveis que podem afetar as escolhas feitas pelos falantes, como o contexto ao qual estão inseridos, por exemplo.

O presente estudo não somente fortalece o campo dos estudos linguísticos, mas também oferece *insights* valiosos para a compreensão do processamento de sentenças ambíguas na mente humana, destacando que a estrutura das sentenças e as características dos protagonistas, no que concerne a nome, sílabas e letras que compõem esse nome, são elementos essenciais a serem levados em consideração e controlados em pesquisas dessa natureza.

REFERÊNCIAS

ALMEIDA, René Alain Santana de. **A PROSÓDIA E O PROCESSAMENTO ON-LINE DE SENTENÇAS NO PORTUGUÊS BRASILEIRO.** Orientador: Prof. Dr. Miguel Oliveira Jr. 2017. 149 p. Tese de Doutorado (Doutor) - Universidade Federal de Alagoas, Maceió-Alagoas, 2017.

ALMEIDA, René Alain Santana de; OLIVEIRA JUNIOR, Miguel; COZIJN, Reinier. **A INFLUÊNCIA DA PROSÓDIA DA FALA NA RESOLUÇÃO DE AMBIGUIDADE SINTÁTICA**: UM ESTUDO DE PROCESSAMENTO DE SENTENÇA. Cad. Est. Ling, [*s. l.*], v. 63, p. 1-23, 2021.

COZIJN, Reinier; COMMANDEUR, Edwin; VONK, Wietske; NOORDMAN, Leo G.M. **The time course of the use of implicit causality information in the processing of pronouns**: A visual world paradigm study**.** Journal of Memory and Language, Holanda, v. 64, p. 381-403, 4 maio 2011.

FINGER, Ingrid; ZIMMER, Marcia C. A preferência de interpretação de orações relativas curtas e longas em português brasileiro. In: MAIA, Marcus; FINGER, Ingrid. **Processamento da Linguagem.** Pelotas: EDUCAT, 2005, p. 111-130.

FORSTER, Renê. Aspectos da utilização do rastreamento ocular na pesquisa psicolinguística: Eye-tracking in psycholinguistic research. **D.E.L.T.A**, Rio de Janeiro, p. 610-644, 31 maio 2017. DOI https://doi.org/10.1590/0102-445095461720767529. Disponível em: https://www.scielo.br/j/delta/a/mSChQsvsKJkdvZPdJLnv8wh/?lang=pt. Acesso em: 18 nov. 2023.

FONSECA, Aline Alves. **Pistas prosódicas e o processamento de sentenças ambíguas do tipo "SN1-V SN2-Atributo" do português brasileiro**. 2008. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

FRAZIER, Lyn. **On comprehending sentences:** syntatic parsing strategies. PhD dissertation. Indiana University Linguistics Club. University of Connecticut, 1979.

GERNSBACHER, Morton A.; HARGREAVES, David J. Accessing sentence participants: The advantage of first mention. **Journal of Memory and Language,** 27, p. 699-717, 1988.

GODOY, Mahayana Cristina; CARVALHO, Renata Sabrinne Souza de. **EFEITOS SINTÁTICOS E TEMÁTICOS NA RESOLUÇÃO DE PRONOMES AMBÍGUOS EM PORTUGUÊS BRASILEIRO.** Revista do GELNE, Natal-RN, v. 22, n. 2, p. 131-142, 2020.

HARTSHORNE, Joshua K.; NAPPA, Rebecca; SNEDEKER, Jesse. **Development of the first-mention bias**. *Journal of Child Language*, Cambridge University Press, v. 41, n. 4, p. 1-24, 2014. Disponível em: https://doi.org/10.1017/S0305000914000075.

JARVIKIVI, Juhani; GOMPEL, Roger P.G. van; HYONA, Jukka; BERTRAM, Raymond. **Ambiguous Pronoun Resolution: Contrasting the First-Mention and Subject-Preference Accounts.** PSYCHOLOGICAL SCIENCE, [*s. l.*], v. 16, n. 4, p. 260-264, 2005.

KAISER, Elsi. **Focusing on pronouns: consequences of subjecthood, pronominalisation, and contrastive focus**. *Language and Cognitive Processes*, v. 26, n. 10, p. 1625-1666, 2011.

KAISER, Elsi. Experimental paradigms in psycholinguistics. In: PODESVA, Robert J.; SHARMA, Devyani. **Research Methods in Linguistics**. Cambridge: Cambridge University Press, 2013, p. 135-168.

LEITÃO, Márcio Martins. **O Processamento do objeto direto anafórico no Português Brasileiro.** Orientador: Marcus Antônio Rezende Maia. 2005. 160 p. Psicolingüística Experimental (Tese de doutorado-Faculdade de Letras) - UFRJ, Rio de Janeiro, 2005. Disponível em: https://www.leffa.pro.br/. Acesso em: 7 jun. 2024.

MAIA, Marcos; FINGER, Ingrid. **Processamento da Linguagem:** Série Investigações em Psicolingüística GT de Psicolingüística da ANPOLL. Pelotas-RS: EDUCAT, 2005. 536 p. ISBN 85- 7590-044-7.

MAIA, Marcos. Linguística experimental: aferindo o curso temporal e a profundidade do processamento: Experimental Linguistics: measuring the time course and the depth of processing. **Revista de estudos da linguagem**, Faculdade de Letras da UFMG, v. 21, ed. 1, p. 9-41, 2013. Disponível em: http://www.periodicos.letras.ufmg.br/. Acesso em: 6 jun. 2024.

MIYAMOTO, Edson T. Orações relativas ambíguas e a homogeneidade do processamento de sentenças. In: MAIA, Marcus; FINGER, Ingrid. **Processamento da Linguagem**. Pelotas: EDUCAT, 2005, p. 71-90.

PICKERING, Martin J.; TRAXLER, Mathew J. Plausibility and Recovery from Garden Paths: An Eye-Tracking Study. **Journal of Experimental Psychology:** Learning, Memory, and Cognition, vol. 24, n. 4, p. 940-961, 1998