

UNIVERSIDADE FEDERAL DO MARANHÃO
CENTRO DE CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS - PGLetras

RAYANE DE ANDRADE RODRIGUES

**SIGNIFICADOS SOCIAIS ASSOCIADOS A SEGUNDA PESSOA DO SINGULAR: A
percepção dos ouvintes caxienses**

São Luís-MA

2023

RAYANE DE ANDRADE RODRIGUES

SIGNIFICADOS SOCIAIS ASSOCIADOS A SEGUNDA PESSOA DO SINGULAR: A
percepção dos ouvintes caxienses

Área de Concentração: Descrição e Análise
do Português Brasileiro.

Orientadora: Profa. Dra. Cibelle Corrêa
Bélíche Alves

Coorientador: Prof. Dr. Wendel Silva dos
Santos

São Luís - MA

2023

Rodrigues, Rayane de Andrade.

SIGNIFICADOS SOCIAIS ASSOCIADOS A SEGUNDA PESSOA DO SINGULAR: : A percepção dos ouvintes caxienses / Rayane de Andrade Rodrigues. - 2023.

98 f.

Coorientador(a): Wendel Wendel Silva dos Santos.

Orientador(a): Cibelle Corrêa Béliche Alves.

Dissertação (Mestrado) - Programa de Pós-graduação em Letras/cch, Universidade Federal do Maranhão, São Luís, 2023.

1. Percepção sociolinguística. 2. Segunda pessoa do singular. 3. Variedade maranhense. I. Corrêa Béliche Alves, Cibelle. II. Wendel Silva dos Santos, Wendel. III. Título.

**SIGNIFICADOS SOCIAIS ASSOCIADOS A SEGUNDA PESSOA DO SINGULAR: A
percepção dos ouvintes caxienses**

RAYANE DE ANDRADE RODRIGUES

Banca Examinadora

Profa. Dra. Cibelle Corrêa Béliche Alves (UFMA)

Orientadora

Prof. Dr. Wendel Silva dos Santos (UFMA)

Coorientador

Prof. Dr. Marcus Garcia de Sene (UPE)

Examinador externo

Prof. Dr. Luís Henrique Serra (UFMA)

Examinador interno

Profa. Dra. Lívia Oushiro (UNICAMP)

Membro Suplente

São Luís - MA

2023

*As desilusões e decepções são necessárias, obrigada!
A Deus, porque ele ouviu a minha voz e a minha súplica.*

AGRADECIMENTOS

A Deus, por me fortalecer nos dias difíceis, por me ouvir e ajudar sempre e por todas as oportunidades que tem me concedido ao longo da minha vida.

Foi um caminho longo, que me proporcionou muitas experiências, conhecimentos e descobertas. Hoje posso dizer que finalizei uma das etapas mais difíceis que já enfrentei, sem dúvidas eu jamais conseguiria concluí sem o apoio e suporte de algumas pessoas, às quais dedico esta dissertação.

À minha família, que serviu de refúgio sempre que precisei. Em especial, meus pais, por serem meus maiores apoiadores e financiadores durante toda essa trajetória acadêmica.

À minha orientadora, professora Cibelle Béliche, obrigada! Por acreditar em mim e ter aceitado desenvolver esta pesquisa. Obrigada, por toda a paciência que teve comigo, por todas as conversas e reuniões. Agradeço por ter me ouvido nos meus momentos não tão bons. A senhora foi a calmaria em meio de muitas tempestades!

Agradeço ao meu querido coorientador, Wendel Santos, por quem nutro uma maior admiração e respeito. Por ter se disponibilizado em fazer parte desta pesquisa por livre espontânea vontade. Serei eternamente grata por esse gesto! Obrigada pelas orientações, pelas críticas, sugestões e por todas as vezes que o senhor fazia eu sair da minha zona de conforto em busca de mais conhecimentos. Saiba que suas contribuições foram significativas e marcaram na construção deste trabalho e na minha jornada acadêmica.

Agradeço à professora Conceição Ramos, por ter me acompanhado durante todo esse percurso do mestrado. A senhora foi uma das minhas maiores incentivadoras com suas palavras de consolo, jamais esquecerei de todas as vezes que a senhora dizia “que acreditava em mim e sempre enxergava uma evolução da Rayane como pesquisadora”. Metaforicamente, a senhora era meu posto de gasolina, pois eu sempre me abastecia quando lhe encontrava. Meu muito obrigada!

Agradeço à professora Livia Oushiro, por ter aceito participar da defesa do meu exame de qualificação. Esta pesquisa também é sua, pois suas sugestões foram bem-vindas e super válidas.

Aos professores membros da banca examinadora, a quem agradeço pela gentileza de aceitar participar da avaliação dessa dissertação.

Por fim, agradeço a todos que ajudaram mesmo que indiretamente nesta pesquisa!

RESUMO

Os trabalhos que se ocupam em acessar significados sociais potencialmente associados a variáveis morfossintáticas no estado do Maranhão são relativamente em menor número se comparadas a variáveis fonológicas (Santos, 2020). Diante disso, esta dissertação de mestrado visa sanar essa lacuna quando se propõe a investigar quais significados sociais são associados à segunda pessoa do singular na fala caxiense, a saber: *tu com concordância*, *tu sem concordância*, *você* e *cê*. Mais especificamente, interessa verificar quais significados sociais os ouvintes nascidos na cidade de Caxias, Maranhão associam a essas formas coocorrentes na sua variedade linguística. O trabalho fundamenta-se nos pressupostos teóricos da Sociolinguística Variacionista (Labov, 2006[1996]; 2008[1972], Eckert, 2008; 2012), e, metodologicamente, baseia-se na adaptação da técnica de *falsos pares* ou *estímulos pareados* (Lambert *et al*, 1960; Campbell-Kibler, 2006; Oushiro, 2015; Soriano, 2016; Mendes, 2017; Canever, 2017; Santos, 2020), que consiste na criação de estímulos auditivos, manipulados no sentido de que pareçam iguais em todos os aspectos, menos em relação à variável em foco. Assim, criou-se um experimento de percepção com disfarces gravados por quatro falantes caxienses (dois homens e duas mulheres), organizados em quatro grupos, cada um formado por quatro disfarces. Os estímulos, cuja temática versa sobre infância, são compostos por sentenças com duas ocorrências de cada uma das variantes de (segunda pessoa do singular). Uma vez elaborados, os conjuntos de estímulos foram apresentados a 60 ouvintes, cada ouvinte respondeu um questionário virtual, que correspondia a quatro estímulos, ou seja, totalizaram 240 respostas que passaram pelas análises de regressão linear, testes não paramétricos e teste qui-quadrado (R Core Team, 2023). As análises mostraram que os falantes foram percebidos, pelos ouvintes caxienses, como mais prestigiosos, quando ouvidos em seus disfarces com as formas linguísticas *tu com concordância*, *tu sem concordância* e *você*. Esses mesmos falantes foram percebidos, pelos ouvintes, como mais caxienses, quando avaliados em seus disfarces com as formas linguísticas *tu sem concordância* e *você*, ao passo que foram percebidos como menos caxiense em seus disfarces com as formas linguísticas *tu com concordância* e *cê*. Por fim, quando os falantes foram avaliados produzindo os estímulos à variante *cê*, os respondentes associaram o uso dessa forma à fala de pessoas com baixa escolarização, pertencentes a uma faixa etária mais jovem, e como sendo de classe social mais baixa. Os falantes foram percebidos como pessoas de escolarização média, classe média e como sendo da região sudeste, ao produzirem os estímulos com a variante *você*, mas sem diferença significativa para a percepção de faixa etária. Quando ouvidos realizando os estímulos com a variante *tu sem concordância*, os ouvintes associaram tais formas à fala de pessoas oriundas do norte e nordeste do país, enquanto os estímulos com *tu com concordância* foram associados à fala de pessoas com alto grau de escolarização, à fala de pessoas de segunda e terceira faixas etárias de origem da região Norte e Sul do Brasil.

Palavras-chave: Segunda pessoa do singular. Percepção sociolinguística. Variedade maranhense.

RESUMEN

Los trabajos que se centran en el acceso a significados sociales potencialmente asociados a variables morfosintácticas en el estado de Maranhão son relativamente menores en comparación con las variables fonológicas (Santos, 2020). Ante esto, esta tesis de maestría pretende subsanar esta laguna al proponer investigar qué significados sociales están asociados a la segunda persona del singular en el habla de Caxias, a saber: *tú con acuerdo*, *tu sin acuerdo*, *tú y cê*. Más específicamente, es interesante verificar qué significados sociales asocian los oyentes nacidos en la ciudad de Caxias, Maranhão, con estas formas concurrentes en su variedad lingüística. El trabajo se sustenta en los presupuestos teóricos de la Sociolingüística Variacionista (Labov, 2006[1996]; 2008[1972], Eckert, 2008; 2012), y, metodológicamente, se fundamenta en la adaptación de la técnica de los falsos pares o estímulos emparejados. (Lambert et al, 1960; Campbell-Kibler, 2006; Oushiro, 2015; Soriano, 2016; Mendes, 2017; Canever, 2017; Santos, 2020), que consiste en crear estímulos auditivos, manipulados para que aparezcan iguales en todos aspectos, menos en relación con la variable en foco. Así, se creó un experimento de percepción con disfraces grabados por cuatro hablantes de Caxias do Sul (dos hombres y dos mujeres), organizados en cuatro grupos, cada uno formado por cuatro disfraces. Los estímulos, cuya temática trata sobre la infancia, se componen de frases con dos apariciones de cada una de las variantes de (segunda persona del singular). Una vez creados, los conjuntos de estímulos fueron presentados a 60 oyentes, cada oyente respondió un cuestionario virtual, el cual correspondió a cuatro estímulos, es decir, hubo un total de 240 respuestas a las que se les realizaron análisis de regresión lineal, pruebas no paramétricas y chi-cuadrado. prueba (R Core Team, 2023). Los análisis mostraron que los hablantes eran percibidos, por los oyentes de Caxias, como más prestigiosos, cuando se les escuchaba en sus formas con las formas lingüísticas *tu con concordancia*, *tu sin concordancia* y *tú*. Estos mismos hablantes fueron percibidos por los oyentes como más caxienses, cuando fueron evaluados en sus disfraces con las formas lingüísticas *tu sem concordance* y *Você*, mientras que fueron percibidos como menos caxienses en sus disfraces con las formas lingüísticas *tu con concordancia* y *cê*. Finalmente, cuando se evaluó a los hablantes produciendo estímulos utilizando la variante *cê*, los encuestados asociaron el uso de esta forma con el habla de personas con bajo nivel educativo, pertenecientes a un grupo de edad más joven y como pertenecientes a una clase social más baja. Los hablantes fueron percibidos como personas con escolaridad media, de clase media y de la región Sudeste, al producir los estímulos con la variante *você*, pero sin diferencia significativa en la percepción del grupo de edad. Al escucharse realizar los estímulos con la variante *tu sin acuerdo*, los oyentes asociaron dichas formas con el habla de personas del norte y noreste del país, mientras que los estímulos con *tu con acuerdo* se asociaron con el habla de personas con un alto nivel de educación, el discurso de personas de segundo y tercer grupo de edad de las regiones Norte y Sur de Brasil.

Palabras- Clave: Segunda persona del singular. Percepción sociolingüística. Variedad Maranhão.

LISTA DE TABELAS

Tabela 1 - Ocorrências das formas tu e você/ocê/cê na amostra geral.	Erro! Indicador não definido.
Tabela 2 - Distribuição geral das formas tu e você, ocê e cê nas localidades estudadas.	44
Tabela 3 - Concordância verbal com o tu em São Luís, de acordo com a escolaridade.	46
Tabela 4 - Ocorrências das formas tu/você, conforme a Escolaridade, entre os falantes ludovicenses.	47
Tabela 5 - Ocorrências das formas tu/você, conforme a Faixa Etária, entre os falantes ludovicenses.	48
Tabela 6 - Ocorrências das formas tu/você, de acordo com a Classe Social, entre os falantes ludovicenses.	48
Tabela 7 - A concordância de segunda pessoa do singular na amostra Caxiense.	49
Tabela 8 - A concordância de segunda pessoa do singular na amostra Ludovicense.	49
Tabela 9 - Realização do pronome de segunda pessoa do singular, na amostra Caxiense.	49
Tabela 10 - A concordância de segunda pessoa do singular e a escolarização, na amostra Caxiense.	50
Tabela 11 – Números de ouvintes de Caxias e de outra localidade que participaram do experimento.	62
Tabela 12 – Componentes Principais gerados pelas correlações entre as respostas nas 5 escalas para ouvintes caxienses – com rotação Promax.	69
Tabela 13 – Resultado do Modelo de Regressão de Efeitos Mistos para o CP Prestígio.	73
Tabela 14 – Resultado do Modelo de Regressão de Efeitos Mistos para o CP Caxiense.	77

LISTA DE FIGURAS

Figura 1 - Frequência da ocorrência da variante rótica, de acordo com classe social econômica e estilo, no inglês de Nova Iorque (adaptado de Labov 2006 [1996], p. 118)	21
Figura 2 - Campo indicial de /t/.....	25
Figura 3 - Nuvem de palavras associadas ao uso de tu pelos participantes.	32
Figura 4 - Primeira parte do formulário de percepção.	35
Figura 5 - Segunda parte do formulário de percepção.	36
Figura 6 - Dispersões das respostas dos ouvintes ludovicenses e paulistanos no CP Competência.	37
Figura 7 – Mapa dos subsistemas dos pronomes de segunda pessoa do singular do português brasileiro falado no final do século XX e início do século XXI, o Brasil de 1709 e o Brasil do século XXI.	41
Figura 8 - Ocorrências das formas tu e você/ocê/cê de acordo com a faixa etária.....	45
Figura 9 - Organização dos estímulos	56
Figura 10 - Página de instrução do experimento	57
Figura 11 - Primeira Parte do Formulário de Percepção	58
Figura 12 - Segunda Parte do Formulário de Percepção	59
Figura 13 - Terceira Parte do Formulário de Percepção.....	60
Figura 15 – Dispersão Geral dos dados.....	65
Figura 16 – Screeplot da ACP das respostas dadas por caxienses às escalas do formulário de percepção.....	67
Figura 17 – Gráfico da direção da correlação	68
Figura 18 – Dispersões das respostas dos ouvintes caxienses no CP Prestígio, por Variante 70	
Figura 19 – Dispersão das respostas dos participantes por falantes de acordo com os estímulos ouvidos	71
Figura 20 – Dispersões das respostas dos ouvintes caxienses no CP Caxienseidade, por Variante	75
Figura 21 - Dispersão das respostas dos participantes por falantes de acordo com os estímulos ouvidos.....	76
Figura 22 - Distribuição das respostas para a variável Escolarização, por variante, pelos ouvintes caxienses	78

Figura 23 - Distribuição das respostas para variável Faixa Etária, por variante, pelos ouvintes caxienses.....	79
Figura 24 - Distribuição das respostas para variável Classe Social, por variante, pelos ouvintes caxienses	80
Figura 25 - Distribuição das respostas para variável Origem do Falante, por variante, pelos ouvintes caxienses	81

LISTA DE QUADROS

Quadro 1 - Estímulos para as gravações dos falantes.	55
--	----

SUMÁRIO

INTRODUÇÃO.....	15
1 ESTUDOS DA VARIAÇÃO E DO SIGNIFICADO SOCIAL	20
1.1 As três ondas da sociolinguística.....	20
1.2 Estudos de Percepção Sociolinguística	26
1.3 Estudos de Percepção Sociolinguísticas no Maranhão.....	32
2 TU X VOCÊ: A ESTRUTURA EM FOCO.....	39
2.1 Formas de Segunda Pessoa do Singular	39
2.2 Segunda Pessoa do Singular no Maranhão.....	42
3 MONTAGEM DO EXPERIMENTO.....	52
3.1 Definições acerca dos falantes	53
3.2 Preparação dos estímulos	54
3.3 O experimento de percepção	57
3.4 Coleta e perfil dos ouvintes	61
4 DESCRIÇÃO E ANÁLISE DOS DADOS.....	63
4.1 CP1- Prestígio.....	70
4.2 CP2 - Caxienseidade	74
4.3 Variáveis qualitativas	78
CONSIDERAÇÕES FINAIS	84
REFERÊNCIA.....	87
APÊNDICES	93

INTRODUÇÃO

Esta pesquisa é a primeira investigação que busca acessar a percepção sociolinguística no português maranhense sobre a variável da segunda pessoa do singular. Busca-se concentrar no nível de análise linguística morfossintática, ao propor a realização de um estudo de percepção acerca das formas de referência à segunda pessoa do singular no português brasileiro, que podem ser expressas com: i) *tu* + verbo com marca de concordância, como em ***tu estudastes para a prova amanhã***, ii) *tu* + verbo sem marca de concordância, a exemplo de ***tu fez a atividade***, iii) *você*, como em ***você falou com a professora*** e iv) *cê*, como em ***cê cancela o contrato***. A intenção mais específica é verificar quais significados sociais os ouvintes nascidos na cidade de Caxias, município do estado do Maranhão, percebem a essas formas coocorrentes na sua variedade linguística.

Há um discurso popular de que os maranhenses “falam a melhor variedade do português brasileiro”, justificado pelo fato de que utilizariam, em sua fala, a forma *tu com concordância*, seguindo, assim, o que prescrevem os compêndios gramaticais normativos. Entretanto, os estudos de que se têm notícia e que buscaram analisar a variável da segunda pessoa do singular do ponto de vista da produção linguística em diferentes comunidades de fala maranhense, como, por exemplo, Herênio (2006), Alves (2010, 2015), Carneiro (2011) e Miranda (2014), demonstraram que a variante *tu com concordância* é marca registrada pelos falantes, majoritariamente, da capital maranhense, ao passo que as variantes *tu sem concordância*, *você* e *cê* são formas linguísticas coocorrentes nas demais localidades do estado.

Os estudos das formas pronominais do Português Brasileiro (PB), a exemplo de Sherre *et al.* (2015); Carvalho (2019), mostram que os pronomes da segunda pessoa do singular admitem a possibilidade de cinco construções pronominais que coocorrem entre si, a saber: *tu com concordância*, *tu sem concordância*, *você*, *ocê* e *cê*. Essas formas, apesar de atuarem concomitantemente no sistema pronominal maranhense, apresentam probabilidades de uso que são distintas para cada forma, e, ainda que essas probabilidades mostrem diversidade no uso, essa variação é sistemática. Esses estudos, realizados em alguns municípios do Maranhão, na sua grande maioria, têm se concentrado, de um lado, na análise de produção dessas variáveis linguísticas, a partir de amostras de fala espontâneas (Carneiro, 2011; Alves, 2015) ou semiespontâneas (Herênio, 2006; Alves, 2010; Miranda, 2014) dos informantes. Por outro lado, são inexistentes as pesquisas que se concentram em acessar significados sociais associados a essas mesmas formas linguísticas, via estudos de percepção.

Em termos gerais, os resultados de pesquisa sobre a variação das formas pronominais da segunda pessoa do singular no estado do Maranhão demonstram que há uma forte correlação do uso dessas formas com fatores sociais e linguísticos, tais como *a localidade, a faixa etária, a escolaridade, a classe social, o sexo, a saliência fônica, os números de sílabas do verbo e o tempo verbal*. Contudo, é importante lembrar que a sociolinguística da produção e da percepção apresentam diferenças, isto é, nem sempre as formas variáveis mais frequentes são as mais percebidas pelos falantes, ou seja, precisamos correlacionar os fenômenos variáveis sob as duas óticas: produção e percepção (Freitag *et al.*, 2016).

A justificativa para a realização deste trabalho baseia-se no fato de que, se os usos e produções das formas linguísticas da segunda pessoa do singular no falar maranhense são bastantes variáveis, e se a variação linguística está diretamente associada à correlação dos fatores estruturais, próprios do sistema linguístico, e sociais, relacionados aos indivíduos que realizam essas formas variáveis, esses resultados podem endereçar significados sociais potencialmente associados a essas formas linguísticas, o que permite propor a realização de um estudo mais aprofundado, que pretende discutir o papel do ouvinte caxiense não apenas na construção do processo de variação, mas também na construção do significado social.

Acredita-se que este trabalho, sendo o primeiro trabalho de percepção no estado do Maranhão sobre uma variável linguística que parece identificar localmente os falantes, poderá nos dar uma visão mais aprofundada da variável que parece indiciar de significados sociais. Portanto, esta dissertação contribui para que se diminua um pouco mais essa lacuna, ao desenvolver um experimento numa comunidade onde os trabalhos de que se têm notícia analisam esse fenômeno apenas do ponto de vista da produção linguística.

Assim, o objetivo geral deste estudo é verificar quais significados sociais estão associados às formas pronominais da segunda pessoa do singular, a partir das percepções dos ouvintes caxienses.

Esse objetivo geral é orientado a partir de alguns questionamentos que direcionam o fazer desta pesquisa e são elencados a seguir:

- i. O uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu foste* e *tu gostastes*, é percebido pelos ouvintes caxienses como uma variante característica de sua localidade?
- ii. O uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu vais* e *tu és*, indicializa maior grau de escolarização entre os ouvintes caxienses?

- iii. O uso das formas pronominais, *tu* com concordância ou *você*, indiciam diferentes grau de formalidade para os ouvintes caxienses?

Essas perguntas direcionam as seguintes hipóteses de trabalho:

- i. O uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu foste* e *tu gostastes*, não será percebido pelos ouvintes caxienses como uma variante da sua localidade;
- ii. Os caxienses, ao ouvirem a realização do pronome *tu* seguido do verbo com marca de concordância, a exemplo de *tu vais* e *tu és*, indicializa maior escolarização;
- iii. Os usos das formas pronominais, *tu com concordância* ou *você*, indiciam graus de formalidade para os ouvintes caxienses.

Metodologicamente, a pesquisa foi desenvolvida com base na técnica de *falsos pares* ou dos *estímulos pareados*¹ (Lambert et al, 1960; Campbell-Kibler, 2006; Oushiro, 2015; Soriano, 2016; Mendes, 2016; Canever, 2017; Santos, 2020), que consiste na criação de estímulos auditivos, manipulados no sentido de que pareçam iguais em todos os aspectos, menos em relação à variável em foco.

A seleção do município de Caxias para a pesquisa se dá devido à diversidade linguística da variável de segunda pessoa do singular, a exemplo do que revela o estudo de Miranda (2014) sobre crenças, atitudes e usos variáveis da concordância verbal com o pronome *tu* nessa cidade maranhense, em que esse autor mostrou que as formas pronominais *tu* com concordância, *tu* sem concordância, *você* e *cê* coocorrem na comunidade de fala com frequência bastante variável.

Do ponto de vista histórico, no Maranhão, durante o período de colonização, houve duas entradas gerais, sendo uma pelo litoral e a outra pelo sertão, essas entradas geraram duas formações socioespaciais diferentes devido a influência de fatores naturais, culturais, econômicos e linguísticos. Pensando-se em termos linguísticos, a cidade de Caxias tem um espaço interessante no contexto maranhense, devido ao encontro das duas entradas ocorrerem exatamente na cidade, conseqüentemente, esse encontro ocasiona um efeito de variedades de povos e uma diversidade de dialetos.

¹ Tradução proposta por Oushiro (2015) para o termo *matched-guise*.

Por outro lado, do ponto de vista linguístico, a cidade de Caxias incluía-se entre as localidades que foram pensadas por Nascentes (1958) para que constituísse a rede de pontos do Atlas Linguístico do Brasil – ALiB². Destaca-se o fato de que a proposta de um Atlas Linguístico, e a seleção das redes de pontos, por Nascentes, foi feita com “base na importância de cada localidade para a região e estado, e dados de sua formação, além dos limites interestaduais e internacionais” (ALiB, 2023). No entanto, essas redes de pontos passaram por uma nova reformulação, o que resultou na retirada da cidade de Caxias da proposta do Atlas. Tal inserção foi feita anos depois, quando da proposta da elaboração do Atlas Linguístico do Maranhão – ALiMA³, evidenciando a importância geolinguística dessa cidade.

A cidade de Caxias tem uma área de 5.150,667 km² e uma população de 166.159 habitantes (IBGE, 2021). Localiza-se a latitude -04° 51’ 32’’ sul e longitude -43° 21’ 22’’ oeste, situada no Sertão Maranhense a cerca de 360 km de São Luís, capital do Estado e, 66 km de Teresina, capital do Estado do Piauí. Em relação aos limites interestaduais, é válido ressaltar que Caxias fica mais próxima da capital do Piauí, Teresina, comparativamente à capital do Maranhão, São Luís, fator que pode evidenciar uma aproximação sociolinguística ao estado vizinho do que necessariamente às influências linguísticas do seu próprio estado, na formação da identidade linguística do seu povo.

Esta dissertação é composta por quatro capítulos. No primeiro deles, denominado de “Estudos da Variação e do Significado Social”, apresenta-se uma revisão dos estudos sociolinguísticos, bem como discussões acerca do interesse nos estudos dos significados sociais de variantes fonológicas, morfológicas e sintáticas. O segundo capítulo, nomeado de “Tu X Você: Estrutura em foco”, discute a abordagem descritiva da variável referente à segunda pessoa do discurso, no singular, por meio da apresentação de estudos que analisaram os contextos de usos dessa variável, principalmente, no Maranhão. O capítulo terceiro, “Montagem do Experimento”, apresenta o desenho do experimento de percepção, a descrição da técnica — *matched-guise* — selecionada, que busca acessar quais significados sociais estão possivelmente associados às formas linguísticas da segunda pessoa do singular pelos ouvintes caxienses. No quarto capítulo, apresentam-se os resultados estatísticos obtidos a partir da análise dos dados obtidos no (R Core Team, 2023), de modo a se confirmar ou refutar as hipóteses inicialmente apresentadas. Por fim, nas considerações finais, apresentam-se as

² Um projeto que se desenvolve no campo da variação linguística, mais especificamente no campo da Dialetoлогия e com base na Geolinguística, caminho metodológico que se ocupa da cartografia dos fatos de língua, cuja produção de maior relevância se consubstancia nos atlas linguísticos.

³ O projeto descreve a realidade do português do Maranhão para identificar fenômenos fonéticos, morfossintáticos, lexicais, semânticos e prosódicos que caracterizam diferenciações ou define a unidade linguística no Estado.

conclusões que os falantes foram percebidos, pelos ouvintes caxienses, como mais prestigiosos, quando ouvidos em seus disfarces com as formas linguísticas *tu com concordância*, *tu sem concordância* e *você*. Esses mesmos falantes foram percebidos, pelos ouvintes, como mais caxienses, quando avaliados em seus disfarces com as formas linguísticas *tu sem concordância* e *você*, ao passo que foram percebidos como menos caxiense em seus disfarces com as formas linguísticas *tu com concordância* e *cê*.

CAPÍTULO 1

1 ESTUDOS DA VARIAÇÃO E DO SIGNIFICADO SOCIAL

A base da teoria sociolinguística se estabelece na indissociabilidade entre a língua e as relações sociais dos falantes. Como já mencionava Labov (2008), não se pode entender o desenvolvimento da mudança linguística fora da vida social de uma comunidade em que ela ocorre. A língua é inerentemente variável, e o modo como enxergamos a relação entre língua e sociedade é determinante para a realização de estudos que tomam a variação como objeto de estudo (Campbell-Kibler, 2009).

Neste capítulo, apresentam-se os trabalhos pioneiros nos estudos sociolinguísticos, depois de se discutir as chamadas “ondas” dos estudos variacionistas. Destacam-se as pesquisas sobre percepção sociolinguística, um campo de estudos que vem crescendo o número de pesquisas desenvolvidas no Brasil. As pesquisas que embasam o presente estudo são apresentadas neste tópico.

1.1 As três ondas da sociolinguística

A Sociolinguística é a subárea dos estudos linguísticos que se ocupa em analisar o uso linguístico pela comunidade de fala (Labov, 2008), e os diversos estudos empíricos vinculados a essa teoria contribuíram para discutir a ideia de homogeneidade linguística pensada, especialmente, pelos formalistas.

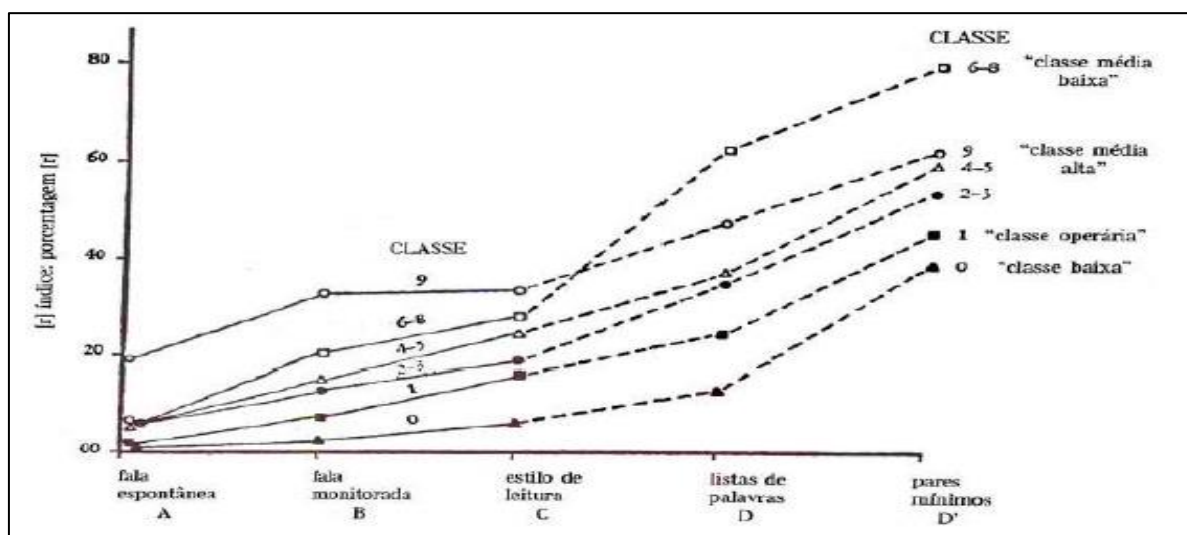
Uma das características mais importantes com as quais a Sociolinguística lida é o fato de estabelecer padrões de uso, já que compreende que os falantes compartilham regras linguísticas que não apenas revelam o conhecimento que o falante possui sobre o seu próprio sistema linguístico, como também trazem à tona as relações sociais estabelecidas entre os próprios falantes. A compreensão dessas duas máximas faz com que os falantes adaptem o seu modo de fala ao contexto em que estão inseridos.

A variação linguística é, assim, a materialização do fato de que “as línguas não existem sem as pessoas que as falam”, e de que “a história de uma língua é a própria história de seus falantes.” (Calvet, 2002, p.12). A variação é um aspecto tão importante para os estudos variacionistas que Eckert (2012, p. 286), ao propor que o estudo sociolinguístico se organiza em três ondas, entende que a “variação não se revela pela posição estrutural do falante em um sistema de produção, mas parte da produção ativa – estilística – da diferenciação social”, ou seja, o estilo se define pelo que o falante faz com a língua dentro do seu contexto social.

As três ondas dos estudos sociolinguísticos referem-se ao modo como diferentes modelos de análise variacionista lidaram com a questão da variação e da mudança linguística. Assim, a primeira onda estabelece amplas correlações entre as variáveis linguísticas e as categorias macrosociológicas de classe socioeconômica, sexo, etnicidade e idade. Um dos estudos característicos da primeira onda é o de Labov (2006 [1996]), sobre a estratificação do inglês nova-iorquino, que objetivou descrever os padrões regulares de variação e de mudança linguística quanto à realização ou não de /r/ em posição pós-vocálica, como em *car* ‘carro’ e *four* ‘quatro’. Nesse estudo, Labov partiu da discussão de que a variável linguística revela sensível estratificação social e estilística entre os nova-iorquinos.

Em linhas gerais, o autor chegou à conclusão de que, para falantes de classe baixa, a ocorrência da variante não rótica é mais frequente, ao passo que, para os falantes de classe mais alta, a ocorrência da variante rótica é mais frequente. Entretanto, há um resultado muito interessante quando se faz o cruzamento da classe econômica x estilo, como se pode observar na Figura 1.

Figura 1 - Frequência da ocorrência da variante rótica, de acordo com classe social econômica e estilo, no inglês de Nova Iorque (Labov 2006 [1996], p. 118)



Fonte: Labov (2006 [1996], p. 118)

De acordo com a Figura 1, o autor observou que, para todas as classes sociais, a depender do contexto de maior monitoramento, a pronúncia da variante rótica tende a aumentar, principalmente quando há mudança de estilo. Observou ainda que os falantes tendem a apresentar taxas maiores de realização de /r/ nos contextos de maior monitoramento, como em leitura de pares mínimos.

Eckert (2012) pontua que esses estudos estabeleceram um padrão regular de estratificação socioeconômica da forma linguística, e que as formas não padrão são mais difundidas na extremidade inferior da hierarquia socioeconômica, o que de fato pode-se verificar na figura acima.

Os estudos de segunda onda sociolinguística diferenciam-se dos de primeira onda, no sentido de que eles são voltados para métodos etnográficos em relação aos falantes, buscando evidenciar significados sociais mais locais, conforme afirma Eckert (2012). Os estudos de segunda onda se interessam, sobretudo, pela agência social do falante quanto ao uso do vernáculo como expressão da identidade local ou de classe. Enquanto na primeira onda a variação é definida por categorias macrossociais, na segunda, ela é definida por categorias localmente definidas, a partir de variantes vernaculares (Mendes, 2017).

A pesquisa de Eckert (2000) é um exemplo de trabalho que se caracteriza de segunda onda. Nesse estudo, a autora analisa dois grupos de alunos em Belten High, escola situada na periferia de Detroit, a saber, *jocks* e *burnouts*, ambos com características sociais distintas. Enquanto o perfil do primeiro grupo se caracterizava com alunos de classe média, atletas e com valores orientados ao colégio e dos professores, o segundo era caracterizado por alunos de classe mais baixa e com valores menos voltados para as orientações escolares.

Em linhas gerais, a autora identificou diferentes usos do alçamento e a posteriorização de vogais média e baixa, a exemplo de *bot* e *bought*, além da variável sintática de dupla negação, como em *I don't know nothing*, e verificou que essas formas linguísticas estavam diretamente associadas à construção dos estilos associados às duas comunidades de práticas distintas que mais se destacavam entre os alunos. Em outras palavras, os usos linguísticos desses grupos tendem manter uma diferenciação mediante a variação dos falantes, ou seja, normas linguísticas compartilhadas que atuam diretamente na construção/expressão de identidades.

No contexto da variedade do português falado em São Luís, um estudo que se enquadra nos estudos de segunda onda é o trabalho de Alves (2015), em que a autora analisa dados de fala de um grupo que tinha como centro dois colaboradores-alvo, um homem e uma mulher, gravados em situações de interação, a partir de observações participantes, ou seja, observando “a pessoa em seu contexto social natural – interagindo com a família ou com seus pares” (Labov, 2008 [1972], p. 63).

O objetivo de Alves (2015) foi descrever a alternância pronominal da segunda pessoa na comunidade de fala de São Luís, Maranhão. Em contrapartida, o *tu com concordância*, quando utilizado, é condicionado, dentre outros fatores, pelo contexto de interação e pelo papel do indivíduo na interação, no sentido de que o falante alterna o seu estilo, sendo essa variante

definida como um traço linguístico de diferenciação entre os ludovicenses escolarizados e não escolarizados, ou seja, os falantes ludovicenses escolarizados compreendem a relevância dessa variável na vida local, fazendo usos linguísticos da variante para manter uma diferenciação e também como forma de construção/expressão de identidade.

Em linhas gerais, os estudos apresentados permitem afirmar que, embora algumas variantes possam ser estigmatizadas em um nível mais amplo da estratificação social, elas também podem ser associadas a valores e práticas locais positivas, no sentido de poder expressar determinadas identidades. Assim, é possível compreender que “variáveis linguísticas não indicializam categorias, mas características, o que oferece uma base teórica e um enfoque metodológico inteiramente novos para o estudo da variação na terceira onda” (Eckert, 2012, p.278).

A terceira onda dos estudos sociolinguísticos, por sua vez, volta a sua atenção para a variação, não mais centrado na correlação a categorias macro ou micro sociais, mas a partir de práticas sociais por meio das quais o falante se posiciona, comunica e interage no mundo social, ou seja, a *prática linguística*. Um aspecto que caracteriza a terceira onda é a noção de estilo, entendido como o *locus* da emergência dos significados sociais (Eckert, 2012). O estilo é social, pois vai ser definido de acordo com o uso da língua pelo falante, e como tal, adquire significado social nas correlações as pessoas e os papéis que desempenham no contexto, em outras palavras, a noção de estilo assemelha à individualidade do falante. Eckert (2012) pontua que noção de estilo é “tomada de posição” dos falantes.

Destaca-se que os estudos de terceira onda se concentram no papel do indivíduo e na construção de *estilos* e *personae*, ou seja, um mesmo falante faz uso de diferentes formas linguísticas e não linguísticas que estão diretamente relacionadas a significados sociais diversos (Barcellos, 2020). Assim, a mudança linguística passa a ser tratada em segundo plano, e os significados sociais da variação passam então a ser o foco e as variáveis assumem significados específicos.

A pesquisa empreendida por Podesva (2007) é expressiva nesse sentido e pode ser considerada como típica de terceira onda. Nesse estudo, o autor analisou a fala de Heath, um médico gay norte-americano, a quem foi pedido que gravasse suas conversas em situações de diferentes interações: no trabalho (ambiente mais formal), no churrasco com os amigos (ambiente informal) e numa conversa com seu pai (ambiente informal). Em seguida, foi medida a entonação (o *pitch*) da fala de Heath nas gravações, com o intuito de observar com que frequência o falsete era usado por ele em cada uma das situações e em qual momento era mais favorecido o uso.

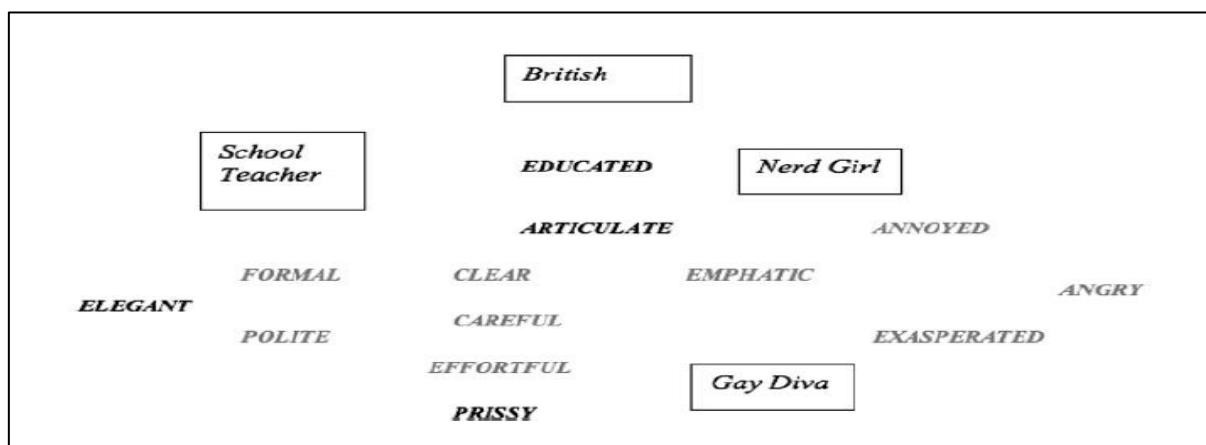
Com base nesse estudo, Podesva (2007) concluiu que a frequência do *pitch* aumentava e diminuía a depender das diferentes situações comunicativas em que Heath estivesse inserido, chegando em alguns momentos a ser mais longos e com frequência (F0) mais alta, especialmente nas interações com os amigos no churrasco, além de em momentos de interação no trabalho, sendo assim correlacionado a características “mais femininas”.

Podesva (2007), ao discutir quais significados sociais foram associadas às variantes, observou que o falante é agente, que atua sobre os significados quando está exposto em situação de interação diferente, ou seja, quando o médico Heath estava atuando em interação distintas, mudava o seu estilo a depender dos objetivos do seu engajamento na conversa, e por isso os significados sociais atribuídos as variantes eram distintas, devido às distintas situações interacionais do falante. O autor constatou que as variantes usadas pelo Heath foram associadas a mais de uns significados sociais. Por exemplo, na interação com os amigos no churrasco, o falsete foi percebido como parte de um estilo “diva”, grandioso e socialmente eloquente, ao passo que, no consultório, foi percebido como um estilo mais cuidadoso, de um médico prestativo com seus pacientes. Ou seja, uma mesma variante carrega significados sociais diferentes.

Eckert (2008; 2012) pontua que os significados sociais podem se relacionar entre si, formando o que a autora denomina de *campo indicial* (tradução do termo *indexical field*) da variação linguística, e se caracteriza por ser dinâmico e multável. Ou seja, certos elementos linguísticos podem estar direta ou indiretamente relacionados a certos elementos sociais, podendo ganhar outros valores, vários outros significados sociais.

Para melhor compreendemos a questão do campo indicial proposta por Eckert (2008), apresenta-se, na Figura 2, o campo indicial para a pronúncia hiperarticulada de /t/, em palavras como *but*.

Figura 2 - Campo indicial de /t/.



Fonte: Eckert (2008, p. 469)

Com base nos resultados de pesquisas associados à variante /t/, que ganha correlações com um conjunto de papéis sociais, como meninas nerds (Bucholtz, 2001), meninos judeus ortodoxos (Benor, 2001) e homens gays (Podesva *et al* 2002, apud Eckert, 2008), entre outros significados sociais, note-se que a variável não apresenta significado estático, mas apresenta significados específicos em diferentes contextos que se articulam ideologicamente, o que significa afirmar que há vários significados sociais que podem se correlacionar a essa variante, a qualquer momento da interação sociocomunicativa.

Nessa perspectiva, entende-se que, nos estudos de terceira onda, a identidade revela uma posição que é assumida pelo sujeito no interior das variadas instituições nas quais ele vive. Os significados sociais fazem parte da construção que coopera entre falantes e ouvintes (Ochs, 1993: 291 *apud* Soriano, 2016). São tantas posições quantas as comunidades interativas em que o sujeito atua, isto porque cada campo social ou comunidade de prática exige um posicionamento específico do sujeito, de acordo com os diferentes papéis sociais que exerce.

Entender o conceito de comunidade de prática é de suma importância para o entendimento do significado social da variação linguística. Enquanto Eckert (2000) define comunidade de prática como um agregado de pessoas que se juntam para engajar-se em algum empreendimento comum, Meyerhoff (2004, p. 530) afirma que comunidade de prática diz respeito a “um conjunto de indivíduos negociando e aprendendo práticas que contribuem para a satisfação de um objetivo em comum”. Para Meyerhoff, a comunidade de prática é caracterizada por três aspectos distintos: “engajamento mútuo”, “empreendimento negociada em conjunto” e “repertório compartilhado”, ou seja, as pessoas estão em contato direto ao se engajarem-se, trabalham em conjuntos, como em uma empresa, e negociam o empreendimento.

Em outras palavras, interagem entre si e compartilham as práticas sociais e usos linguísticos entre os falantes.

É quando discute as características dos estudos da terceira onda que Eckert (2012) lança mão do conceito de comunidades de práticas como *locus* primordial da construção estilística, facilitando a compreensão de que a construção de uma persona/identidade pode estar diretamente associada à agentividade linguística, entendida como a maneira como o indivíduo conduz os estilos.

Os estudos sociolinguísticos entendem a variação como “diferentes formas de dizer a mesma coisa” (Tarallo, 2007). Essa ideia foi bastante difundida e amplamente discutida, no sentido de que a variação era compreendida a partir da correção de fatores de ordens sociais, estruturais e linguísticos, levando em consideração o estilo.

Embora estilo tivesse uma visão diferenciada dentro dos estudos variacionistas de primeira e segunda onda, foi nos estudos da terceira onda que passou a ser retomado, de acordo com Eckert (2003, 2005, 2006, 2010, 2012). Desse modo, essa autora compreendia que faltava uma teoria para retomar os interesses do estudo do significado social das variáveis porque “a ênfase na prática estilística da terceira onda situa os falantes não como portadores passivos e estáveis de um dialeto, mas como agentes estilísticos, costurando estilos linguísticos em projetos, em desenvolvimento e de longa duração, de autoconstrução e diferenciação”.

1.2 Estudos de Percepção Sociolinguística

Os estudos sobre percepção linguística têm se consolidado cada vez mais dentro dos estudos variacionistas no Brasil, com o intuito de elaborar uma visão mais abrangente em relação à variação linguística, observável não apenas na produção da fala, mas também no processo de como as variantes são ouvidas e processadas por membros de uma comunidade (Oushiro, 2015). Esses estudos têm se ocupado em acessar significados sociais potencialmente associados a certas variantes linguísticas, e buscam a sistematicidade dos julgamentos feitos pelos ouvintes, pois, conforme Oushiro (2015), estamos a todo momento fazendo inferências a respeito das pessoas que escutamos.

Um dos primeiros trabalhos que se apoiaram em fazer a sistematicidade das percepções linguísticas foi o de Pear (1931, *apud* Oushiro, 2015). Oushiro explica que Pear fez um estudo em que buscava relacionar os perfis de vozes escutadas pela rádio BBC, ao pedir a ouvintes da rádio que fornecessem os perfis sobre a personalidade de diversas pessoas que eles ouviam; em seguida, os perfis obtidos pelos ouvintes foram comparados com os perfis fornecidos por indivíduos que conheciam os donos das vozes. No geral, Oushiro aponta que Pear observou que

as respostas eram distintas entre os grupos que julgaram as vozes, ainda que houvesse um amplo consenso nas impressões dos ouvintes em relação aos estereótipos que foram associados a cada uma das vozes.

Lambert *et al* (1960) propuseram fazer um experimento de percepção sociolinguística com 130 estudantes universitários – anglófonos e francófonos – em relação à língua inglesa e à francesa. Para a construção dos estímulos, utilizaram a leitura de um texto escrito em francês e inglês lido por quatro falantes bilíngues. Os ouvintes foram expostos às gravações dos dois textos e, em seguida, convidados a fazer julgamentos e expressar opiniões sobre as vozes, a partir das escalas altura, aparência física, aptidão à liderança, senso de humor, inteligência, religiosidade, autoconfiança, digno de confiança, jovialidade, gentileza, ambição, sociabilidade, caráter e capacidade geral de ser simpático. Durante essa tarefa, os ouvintes não foram informados que se tratava dos mesmos falantes.

Os resultados mostraram que os ouvintes não conseguiram de fato identificar que se tratava das mesmas vozes, e diante desse experimento os pesquisadores entenderam que estava sendo de fato avaliado não eram as gravações, mas a língua. Dos resultados obtidos, destaca-se que os falantes avaliaram a língua inglesa mais positivamente do que a língua francesa. Os estímulos em inglês se sobressaíram em sete características (altura, boa aparência, inteligência, confiabilidade, bondade, ambição e caráter), ao passo que os estímulos em francês foram avaliados positivamente no tocante à escala humor. As avaliações feitas pelos ouvintes francófonos coincidiram com as avaliações dos ouvintes anglófonos, julgando a língua inglesa positivamente em relação à sua própria língua.

Essa técnica, que ganhou uma grande relevância dentro dos estudos sociolinguísticos, caracteriza-se, de acordo com Oushiro (2015, p. 268) como um “método rigoroso e elegante que, ao controlar uma série de variáveis com a criação de estímulos comparáveis, permite investigar apropriadamente o papel da linguagem na formação de impressões e julgamentos sociais”. Para Labov (2006 [1966]), o estudo de Lambert *et al* (1960) permitiu acessar as reações subjetivas dos ouvintes, o que para ele é “a questão mais obscura e difícil” (p. 265) dentro dos estudos sociolinguísticos. A partir desse trabalho, Labov buscou aplicar essa mesma metodologia com base nos resultados de produção, com a finalidade de estudar as reações e atitudes de falantes nova-iorquino em relação as cinco variáveis fonológicas: (r), (oh), (æh), (th), (dh).

Labov (2006 [1966]) concluiu que os usos dessas variáveis estavam correlacionados a fatores extralinguísticos, como classe social e estilo. No que diz respeito à classe social, Labov verificou que a taxa de produção da variável /r/ da classe média alta excedia o da classe alta, o

que podia ser explicado pelo fato de que os falantes de uma classe social mais baixa desejam sempre soar como falantes de uma classe mais alta. Essa explicação se justifica, especialmente, a partir do processo de hipercorreção social. Diante desse resultado, Labov se interessou em acessar as atitudes/percepções desses falantes/ouvintes quanto ao uso dessas variáveis.

O autor selecionou algumas sentenças dos textos gravados e apresentou a 122 ouvintes (novaiorquinos de diferentes idades, sexos, classes socioeconômicas e etnias). Em seguida, pediu que imaginassem estar na condição de um gerente, e tivessem que selecionar as candidatas que estivessem prestes a ser contratadas. Todos os estímulos foram ouvidos em ordem aleatória, e, em seguida, os ouvintes deveriam preencher um questionário que indicava qual cargo era adequado para cada uma das vozes ouvidas: personalidade televisista, secretária executiva, recepcionista, operadora de telefone, vendedora, operária de fábrica ou nenhuma das anteriores.

Em linhas gerais, Labov (2006 [1966]) confirmou sua hipótese de que a pronúncia rótica tende a ser a variante mais socialmente prestigiada pelos novaiorquinos, chegando à conclusão de que “o princípio mais geral que aparece nesta pesquisa é que as reações subjetivas referentes às variáveis fonológicas formam uma estrutura profunda enraizada, reconhecida por toda a comunidade de fala⁴”. (Labov, 2006 [1966], p. 298).

Diante desse trabalho de Labov (2006 [1966]), além do aperfeiçoamento da técnica desenvolvida por Lambert *et al* (1960), muitos estudos se interessaram em acessar os significados sociais associados a variantes linguísticas (Podesva, 2007; Campbell-Kibler, 2006, 2007, 2009; Levon, 2014), ainda que esse interesse tenha sido recente. No contexto da sociolinguística brasileira, destacam-se os pioneiros trabalhos de Oushiro (2015), Mendes (2016, 2017, 2018), Soriano (2016) e Barcellos (2020), em São Paulo. No Rio de Janeiro temos os trabalhos, por exemplo, de Melo (2017) e Carvalho (2019). No contexto maranhense, destacam-se os estudos de Serra (2018) e Santos (2020).

Mendes (2016), analisou os significados sociais da concordância nominal (CN), na fala de quatro homens. O autor empregou a técnica *matched-guise* (Lambert *et al*, 1960; Campbell-Kibler, 2006) para a elaboração dos estímulos, feitos por dois trechos da fala de 4 falantes (pseudônimos Carlos, Robson, Jaime e Lucas), retirados da amostra do SP2010⁵. Com

⁴ Tradução do texto feita por Santos (2020): “*The most general principle which appears from this review is that subjective reactions to phonological variables form a deeply embedded structure which is recognized by the entire speech community*”.

⁵ O Projeto SP2010 reúne amostras da fala paulistana coletadas pelo Grupo de Estudos e Pesquisa em Sociolinguística da USP, na forma de gravações e transcrições, com o intuito de contribuir para a disseminação de trabalhos sociolinguísticos sobre essa variedade de fala.

auxílio do software PRAAT (Boersma & Weenink, 2014), foram elaborados pares, sendo um com o disfarce na versão padrão da CN, a exemplo de “os meninos”, e outro na versão não padrão da CN, como em “os menino-Ø”).

O experimento foi aplicado a 100 participantes, que responderam a um questionário de percepção, e eram eliciados a julgar cada uma das quatro vozes de acordo com as escalas de escolaridade, amigabilidade, feminilidade, formalidade, inteligência e classe. Como resultados gerais, Mendes (2016) observou que, a depender do estímulo em que foram ouvidos pelos ouvintes, os mesmos falantes foram avaliados diferentemente. Grosso modo, os falantes Jaime e Lucas soaram mais efeminados em seus dois disfarces (CN padrão ou não padrão) do que, Carlos e Robson, mas todos soaram mais efeminados com a CN padrão.

De modo mais específico, o autor mostrou que a variante padrão (CN) é associada a questões como efeminilidade de forma direta, o que, de acordo com Mendes, baseado em Eckert (2008), são significados sociais associados à concordância nominal que comporiam o campo indicial dessa variante na fala paulistana.

Como esta pesquisa se interessa em acessar significados sociais potencialmente associados ao uso dos pronomes *tu* e *você*, destaca-se, aqui, o trabalho de Carvalho (2019), que estudou a percepção sociolinguística das formas pronominais de segunda pessoa do singular (2SG) em posição de sujeito, no português do Rio de Janeiro, e conta com dois experimentos: o primeiro diz respeito a um julgamento de adequação sociolinguística, e, o segundo, um questionário de avaliação subjetiva (com base na técnica de *matched-guise*). Com a finalidade de testar a hipótese central de sua pesquisa, a de que a forma *você* é percebida pelos falantes cariocas como adequada a diferentes contextos sociointeracionais, como em “*poxa, mas você pode parcelar em até dez vezes sem juros*” e “*prazer, Carla. Você vai adorar a cidade*”, ao passo que a forma *tu* associa-se a contextos mais específicos, a depender da relação estabelecida entre os falantes – simétricas e assimétricas –, como em “*tu pode experimentar se quiser*” e “*sim, tu pode deixar as bolsas*”.

O primeiro experimento é baseado na técnica chamada de “Julgamento de aceitabilidade”. Nos testes desse tipo, os participantes são expostos a frases isoladas em telas brancas e devem julgar a aceitabilidade e a naturalidade de tais frases em sua língua. Fazendo as devidas adaptações, a pesquisadora selecionou cenas, disponibilizadas na internet por editoras de livros didáticos de língua estrangeira em situações diversificadas, de maneira que os pronomes estudados aparecessem como parte dos diálogos dessas cenas; a autora criou diálogos com frases contendo *tu* e *você*, para que os participantes fossem expostos aos pronomes em situações cotidianas e não de forma isolada.

Foi aplicado a 40 ouvintes, todos universitários do curso letras, nascidos e/ou moradores da região Metropolitana do Rio de Janeiro, sendo 30 mulheres e 10 homens, entre 17 e 40 anos. Eles foram instruídos oralmente e por escrito sobre como proceder em relação à tarefa a ser realizada. Para esse experimento a autora contou com três variáveis independentes: os pronomes de segunda pessoa do singular (*tu* e *você*), o tipo de relação (simétrica e assimétrica) e o tipo de frase (com verbo e sem verbo). Os ouvintes foram divididos em quatro grupos. Conforme a autora, a divisão foi necessária por dois motivos: a extensão e duração do teste, e a identificação do fenômeno observado. Assim, os grupos 1 e 2 assistiram a cenas de relações assimétricas com diferenças em termos dos padrões frasais postulados, enquanto os grupos 3 e 4 assistiram a cenas de relações simétricas. Em seguida eles deveriam atribuir uma nota (de 1 a 5) considerando a naturalidade e a adequação dos diálogos em relação à cena.

O segundo experimento elaborado foi uma adaptação da técnica *matched-guise*, no qual dois homens e duas mulheres, com ensino superior, com idades entre 21 e 27 anos, foram convidados a gravar um texto elaborado pela pesquisadora, um parágrafo com 29 palavras que apresentava os pronomes de segunda pessoa do singular duas vezes. Esse parágrafo se apresentava em quatro versões, o que, de acordo com a pesquisadora, “são as possibilidades de realizações dos pronomes e verbos correspondentes no Rio de Janeiro, além de uma versão controle, com a forma *tu* combinada a verbo com desinência de 2SG (*tu-2SG*)” (Carvalho, 2019, p. 62). Vejamos exemplos de cada versão criada:

Versão 1 (*tu*): Oi, Laura. Tu pode comprar o ingresso do filme pra mim? Eu acabei me atrasando e vou chegar em cima da hora. Tu sabe onde é o cinema, né?

Versão 2 (*você*): Oi, Laura. Você pode comprar o ingresso do filme pra mim? Eu acabei me atrasando e vou chegar em cima da hora. Você sabe onde é o cinema, né?

Versão 3 (*você/tu*): Oi, Laura. Você pode comprar o ingresso do filme pra mim? Eu acabei me atrasando e vou chegar em cima da hora. Tu sabe onde é o cinema, né?

Versão 4 (*tu-2SG*): Oi, Laura. Tu podes comprar o ingresso do filme pra mim? Eu acabei me atrasando e vou chegar em cima da hora. Tu sabes onde é o cinema, né?

Foram criadas 16 gravações, divididas em dois grupos. O experimento foi aplicado a 50 participantes, sendo 36 participantes do sexo feminino e 14 do sexo masculino, de faixa etária entre 17 e 56 anos.

Após ouvirem as gravações, os participantes eram convidados a preencher o questionário com 12 características que a pesquisadora julgou pertinentes como representativas

da fala carioca. Dentre elas, três se relacionam à identidade nacional e regional (*carioca*, *suburbano* e *favelado*), três a aspectos de solidariedade (*honestidade*, *intimidade* e *formalidade*), três se relacionam a status (*rico*, *trabalhador*, *letrado*) e três a atrativo pessoal (*elegante*, *cortês*, *velho*).

Os atributos eram apresentados com o seu correspondente oposto (*muito/pouco carioca*, *honesta/desonesta*) e os participantes deveriam avaliar o falante de acordo com uma escala contínua de 5 pontos. Além do questionário, a autora também solicitou as impressões dos participantes quanto ao uso da forma pronominal *tu*.

Os resultados alcançados pelo primeiro experimento aplicado na pesquisa mostraram que, nos disfarces que continham as formas pronominais *você*, os falantes foram mais bem percebidos, e foi a forma mais aceita em todos os contextos. Já em relação ao segundo experimento, ressaltamos o resultado que comprova a hipótese da autora, de que de fato o *tu* com concordância não seria associado à fala carioca. Sobre os resultados das características atribuídas aos disfarces temos: a) na dimensão solidariedade, *tu* com concordância se correlacionou ao polo mais formal, ao passo que as formas *tu* sem concordância e *você/tu* foram percebidos como mais informais e a forma *você* foi percebida como “neutra”; b) na dimensão do status, os falantes foram percebidos como mais ricos quando estavam usando a variante *tu* com concordância; e c) na dimensão atrativo pessoal, a variante *tu* com concordância foi associada a um falante mais cortês, elegante e mais velho (Carvalho, 2019).

Em síntese, a pesquisa de Carvalho (2019) permitiu concluir que a forma pronominal *você*, na fala carioca, apresenta um comportamento de *indicador*⁶, no sentido de Labov (2008, p. 360), enquanto a forma *tu* se comporta como um *marcador*⁷, isto é, está sujeito à variação estilística e tem a sua percepção/avaliação influenciada pelo tipo de interação entre os indivíduos, sendo avaliada como a forma mais adequada para situações de relação simétrica, além de estar associada à noção de informalidade pelos participantes. Os resultados das impressões dos participantes quanto ao uso da forma pronominal *tu* foi apresentado na forma de nuvem de palavras, conforme se pode observar na figura:

⁶ *Indicadores*: traços linguísticos encaixados numa matriz social, exibindo diferenciação segundo a idade e o grupo social, mas que não exibem nenhum padrão de alternância estilística e parecem ter pouca força avaliativa.

⁷ *Marcadores*: exibem estratificação estilística tanto quanto estratificação social. Embora possam estar abaixo do nível da consciência, produzirão respostas regulares em testes de reação subjetiva. De acordo com Labov as reações envolvidas na mudança linguística, segundo o tipo de avaliação social que recebem, podem ser classificadas como indicadores, marcadores e estereótipos, que são formas socialmente marcadas, rotuladas enfaticamente pela sociedade.

Figura 3- Nuvem de palavras associadas ao uso de tu pelos participantes.



Fonte: Carvalho (2019, p. 94)

Os participantes associaram o uso do *tu* a “informal”, e também associaram o uso do *tu* a outras palavras que fazem parte do mesmo campo semântico, como por exemplo, “informalidade”, “descontraída” e “situações informais”. Carvalho (2019, p. 95) conclui que “essas associações, sobretudo em relação à informalidade, [...] demonstraram que o uso de *tu* é avaliado mais positivamente em situações informais de comunicação.”

As pesquisas de percepção empreendida por Mendes (2018) e Carvalho (2019) são umas das poucas que, sobretudo no contexto da sociolinguística brasileira, não se ocupou em analisar uma variável fonético-fonológica, nível de análise em que se concentra a maioria dos estudos de percepção que vêm sendo desenvolvidos.

No tópico a seguir, elencam-se pesquisas de percepção sociolinguística realizadas no estado do Maranhão associadas à produção de variáveis morfossintáticas.

1.3 Estudos de Percepção Sociolinguísticas no Maranhão

Pesquisas que se ocupam em analisar os significados sociais potencialmente associados a variáveis morfossintáticas são menos frequentes se comparadas a variáveis fonológicas. De acordo com Santos (2020, p. 72), “as variáveis gramaticais têm recebido menos atenção dos estudos de percepção sociolinguística, relativamente a variáveis fonológicas”. De todo modo, alguns pesquisadores vêm contribuindo para diminuir essa lacuna. No Maranhão, por exemplo, encontramos as pesquisas de Serra (2018) e Santos (2020), que voltaram suas atenções para percepção em variáveis de nível morfossintático.

Serra (2018) se interessou em analisar a produção, as crenças e atitudes linguísticas da dupla negação no português maranhense, em duas localidades: a capital, São Luís, e a

comunidade quilombola Jamary dos Pretos. Essas localidades foram escolhidas porque a autora gostaria de observar se há diferenças entre a realidade urbana e rural quanto à relação da estrutura e das crenças. De acordo com a autora, o português do Brasil se destaca como a única língua românica a apresentar três estruturas negativas, a saber: a negação pré-verbal (NÃO + sintagma verbal) como, por exemplo em, “*não gosto muito disso*”⁸, a dupla negação (NÃO + sintagma verbal + NÃO) como, por exemplo, em “*já fui, hoje não sou mais não*”⁹ e a negação pós-verbal (sintagma verbal + NÃO) como, por exemplo, em “*ah, tenho essas coisas não*”¹⁰.

Para o seu trabalho de produção, Serra (2018) analisou uma amostra de fala composta por 24 pessoas nascidas nas localidades estudadas (16 de São Luís e 8 de Jamary dos Pretos), estratificadas de acordo por seu gênero/sexo (mulheres e homens), faixa etária (faixa I, de 18 a 30 anos, e faixa II, de 55 a 85 anos) e escolaridade (ensino fundamental e ensino superior, sendo este último nível de escolaridade analisado somente em São Luís).

Para a produção da dupla negação, Serra (2018) observou que, em sua amostra, 80% dos dados analisados estavam em estrutura pré-verbal, 17% em posição pré e pós-verbal (dupla negação) e 2% de negação em posição pós-verbal. Em Jamary dos Pretos, a negação pré-verbal totalizou 86% do total de ocorrências, enquanto a dupla negação foi produzida em 13% dos dados analisados, e a negação pós-verbal em 0,8%. Em São Luís, a negação em posição pré-verbal foi a forma mais recorrente na amostra (78% dos dados), seguida da dupla negação (19% dos dados analisados) e da negação pós-verbal (3% da amostra de fala de analisada).

Serra (2018) apresenta os resultados das variáveis separadamente, sob a justificativa de que se tenha uma melhor compreensão dos dados. Na variável localidade, os dados apontaram que não há grande diferença em relação à dupla negação nas localidades, mas sinalizam que a dupla negação é a segunda mais recorrente. Na variável sexo, a produção da dupla negação apresentou um favorecimento pelas mulheres. Na variável faixa etária, os falantes da primeira faixa etária tendem a favorecer o uso da dupla negação. Por fim, a variável escolaridade apresentou que os falantes do ensino superior tendem a favorecer o uso da dupla negação, resultado obtido com base apenas na amostra de São Luís.

A partir dos resultados do trabalho de produção, Serra (2018) elaborou um teste de percepção que objetivou acessar os significados sociais atribuídos às variantes para a realização de negação no português maranhense. A princípio, a autora fez uma entrevista aberta em relação às estruturas em estudo com os alunos do curso de Letras, a fim de obter metacomentários sobre

⁸ Extrato retirado do trabalho de Serra (2018).

⁹ Extrato retirado do trabalho de Serra (2018).

¹⁰ Extrato retirado do trabalho de Serra (2018).

as sentenças que continham as variáveis em estudo. Após essa etapa, a autora contou com a participação de dois falantes, um homem de 60 anos de idade, e uma mulher de 25 anos, ambos estudantes do curso de Teatro, que foram convidados a gravar sentenças que continham os contextos de negação pré-verbal, como em “*não sei, mas deve ter aqui perto*”, e dupla negação, a exemplo de “*não vi não, mas eu acho que tá em cima do armário*”.

A partir daí, foi elaborado um questionário com 12 perguntas, tais como *qual profissão essa pessoa exerce? essa pessoa fala corretamente? essa pessoa tem boa condição financeira?* divididas em cinco campos aspectuais: *inteligência/aparência, classe social, fala, caráter e estado emocional*. Esse teste foi aplicado a cinco ouvintes, mas a autora apresentou os resultados para apenas três desses participantes.

Quanto ao teste de percepção, os informantes relacionaram o uso da estrutura de dupla negação ao falar rural, caipira, interiorano, e a pessoas sem instrução, e, durante à realização do teste, os ouvintes podiam justificar suas escolhas. Desse modo, muitas justificativas dos ouvintes direcionavam à noção de erro. Assim, a autora concluiu que as respostas dadas pelos ouvintes estavam relacionadas as crenças e atitudes, em relação a estrutura da dupla negação. Serra (2018) observou que, ainda que os falantes maranhenses realizem com certa frequência a variante dupla negação, eles não a têm como uma variante de prestígio na comunidade.

Santos (2020), por sua vez, buscou investigar os efeitos da morfologia do subjuntivo e do indicativo na percepção de quão competentes, sérios, formais e antipáticos soam os ludovicenses e os paulistanos. O autor analisou as percepções associadas ao modo subjuntivo e indicativo, em contextos de subordinação substantiva (como em “*O chefe quer que a secretária permaneça/permanece*” e “*O candidato à vaga acredita que possam/podem chamá-lo*”) e contextos de subordinação adverbial (como em “*Embora permanecessem/permaneciam com dívidas, terminaram o trabalho*” e “*Talvez estivessem/estavam todos os interessados presentes para participar da reunião*”). O autor parte da discussão de que as características atribuídas ao subjuntivo e ao indicativo estariam estritamente atribuídas à morfologia a condição básica para a sua expressão.

Para eliciar os significados sociais associados ao uso do indicativo e do subjuntivo, foram gravados um homem e uma mulher paulistanos e um homem e uma mulher ludovicenses, todos com ensino superior completo e com faixa etária entre 20 e 30 anos, enunciando 4 pares de sentenças – uma no subjuntivo e uma no indicativo, sendo dois pares de orações adverbiais introduzidas por *embora* e *talvez*, e dois pares de orações subordinadas substantivas, introduzidas pelo verbo volitivo *querer* ou pelo verbo cognitivo *acreditar*. As sentenças foram manipuladas digitalmente com o auxílio do programa *Praat*, a fim de diferenciar os dois

falantes ludovicenses dos dois paulistanos. Assim, os estímulos sofreram alterações na forma verbal e na pronúncia do -/r/ em coda, tepe para os paulistanos e aspirado para os ludovicenses¹¹. Foram elaborados 32 os estímulos, 16 no subjuntivo e 16 no indicativo.

Antes de aplicar o experimento propriamente, os estímulos elaborados foram apresentados a um conjunto de ludovicenses, além de alunos da Universidade de São Paulo, a fim de que dessem suas impressões gerais sobre os áudios. Essas impressões foram, então, sintetizadas em nuvem de palavras. Em seguida, as palavras com maior destaque na nuvem foram utilizadas na elaboração do questionário de percepção.

Figura 4- Nuvem de palavras baseada nas principais impressões e avaliações dos ouvintes.

Fonte: Santos (2020, p. 138)

O questionário de percepção foi elaborado com base nas principais impressões e avaliações dos ouvintes:

Figura 4- Primeira parte do formulário de percepção.

Essa pessoa soa: *

Dê preferência ao uso de fones de ouvidos.

Nada amigável	☐	☐	☐	☐	☐	☐	Muito amigável
Nada antipático/a	☐	☐	☐	☐	☐	☐	Muito antipático/a
Nada educado/a	☐	☐	☐	☐	☐	☐	Muito educado/a
Nada escolarizado/a	☐	☐	☐	☐	☐	☐	Muito escolarizado/a
Nada formal	☐	☐	☐	☐	☐	☐	Muito formal
Nada inteligente	☐	☐	☐	☐	☐	☐	Muito inteligente
Nada paulistano/a	☐	☐	☐	☐	☐	☐	Muito paulistano/a
Nada sério/a	☐	☐	☐	☐	☐	☐	Muito sério/a

Fonte: Santos (2020, p. 138)

Os ouvintes também eram convidados a responder outra parte do questionário em relação ao como percebiam os falantes quanto a sua idade do falante, o seu grau de escolaridade e a classe social que poderia ser atribuída a eles.

¹¹ Foram controladas as variantes forma verbal e a pronúncia do - /r/ em coda. O autor ressalta que a alteração em relação ao - /r/ em coda se trata um fenômeno linguístico que mais marca diferenças dialetais no português falado no Brasil.

Figura 5 - Segunda parte do formulário de percepção.

Qual você acha que é o nível de escolaridade dele/dela? *

☐ Fundamental

☐ Médio

☐ Superior

Qual você acha que é a faixa etária dele/dela? *

☐ 20 a 30 anos

☐ 30 a 40 anos

☐ 40 anos ou +

A que classe social você acha que ele/ela pertence? *

☐ Baixa

☐ Média Baixa

☐ Média

☐ Média Alta

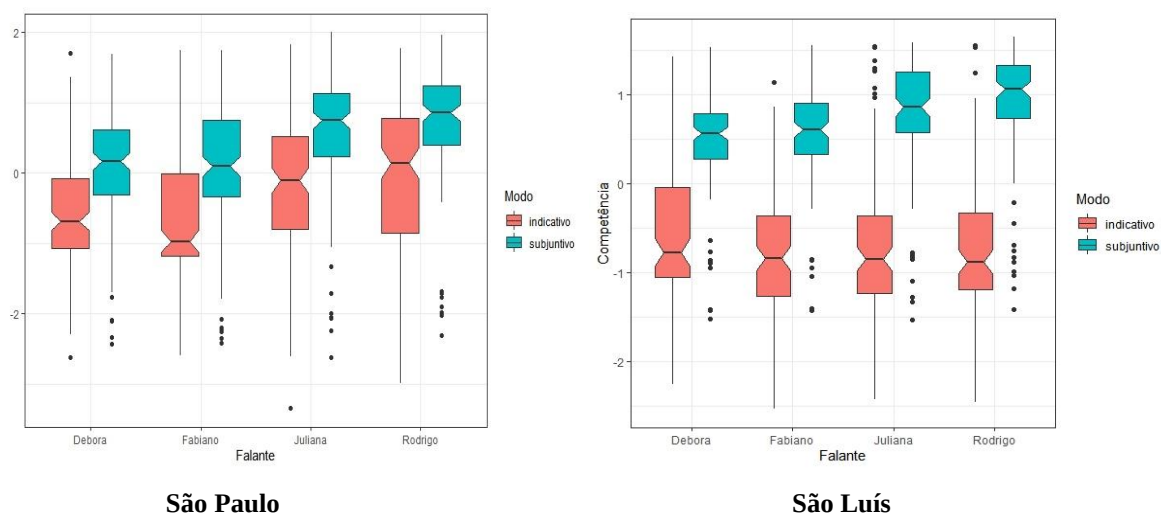
☐ Alta

Fonte: Santos (2020, p. 138)

Por fim, eles forneceram informações sociodemográficas sobre si. Participaram da pesquisa um total de 501 ouvintes: 217 de São Luís e 284 de São Paulo.

Santos (2020) observou que, de maneira geral, os falantes foram percebidos como mais competentes quando ouvidos nos seus disfarces com formas subjuntivas, tanto por ouvintes paulistanos quanto ouvintes ludovicenses. Os falantes também foram avaliados quanto ao seu grau de antipatia, competência, *paulistanidade*, seriedade e formalidade.

De maneira geral, os falantes foram percebidos como mais competentes, no sentido de escolarizados, quando ouvidos nos disfarces no subjuntivo, como evidenciam as respostas amostradas na Figura 6:

Figura 6 - Dispersões das respostas dos ouvintes ludovicenses e paulistanos no CP Competência.

Fonte: Santos (2020, p. 148)

Observa-se que, ainda que todos os falantes (paulistanos e ludovicenses) fossem avaliados como mais competentes quando ouvidos no modo subjuntivo, os falantes paulistanos foram avaliados como ainda mais competentes que os falantes ludovicenses, tanto por ouvintes paulistanos, quanto pelos próprios ouvintes ludovicenses.

Além disso, a percepção dos paulistanos como mais competentes poderia estar relacionada ao fato de que, como o autor observou no próprio discurso dos participantes, “paulistanos são reconhecidos como os que mais têm acesso a bens de cultura, comparativamente a ludovicenses”. Afirmarões desse tipo apareceram nas entrevistas abertas realizadas pelo pesquisador. Por fim, mas não menos importante, Santos (2020) discutiu o aspecto teórico de que variáveis morfossintáticas estão disponíveis para avaliação social, igualmente ao que ocorre com as variáveis fonológicas.¹²

Os trabalhos aqui apresentados possuem indiscutível importância para a pesquisa que aqui se desenvolve, no sentido de que subsidiam o interesse teórico-metodológico de acesso aos significados sociais associados ao uso dos pronomes *tu* e *você*, no português maranhense.

No próximo capítulo, são apresentados trabalhos de produção sociolinguística a respeito da variável de segunda pessoa do singular, no português brasileiro, e mais especificamente, no português falado no Maranhão. Esses trabalhos vão contribuir para o desenho do experimento que aqui se propõe a realizar, no sentido de que se compreende que os

¹² O autor especifica que o interesse na percepção de variáveis morfossintáticas parte das afirmações dos estudos de (Labov, 1993; Labov *et al*, 2011; Levon & Buchstaller, 2015; Mendes, 2018) de que “elas estariam menos propensas à avaliação social, comparativamente às variáveis fonológicas e lexicais”.

resultados obtidos por essas pesquisas podem encaminhar possíveis significados sociais associados às variáveis de interesse deste estudo.

CAPÍTULO 2

2 TU X VOCÊ: A ESTRUTURA EM FOCO

Neste capítulo, apresentaremos pesquisas referentes ao estudo da expressão da variável que mais admite possibilidades dentro do sistema pronominal português, uma vez que concorrem, entre si, diferentes variantes para a referência à segunda pessoa do singular (2SG), no português brasileiro (PB), a saber: *você*, *ocê*, *cê*, *tu sem concordância* e *tu com concordância* (Scherre *et al.*, 2021).

Descreve-se, de modo breve, a trajetória diacrônica dessa variável desde *vossa mercê*, *vossemecê*, *vomecê*, *vancê*, até chegar à forma atual *você/ocê/cê*, que variam com a forma *tu*, este com ou sem concordância, com base na literatura existente.

Na sequência, apresentaremos as mudanças sofridas pela forma *você*. Desde quando passou a ser inserida no quadro pronominal do PB, tal forma passa por um processo de gramaticalização e configuração, como pronome sujeito, além de pronome de tratamento. É importante lembrar que a forma *você* só entrou no sistema pronominal brasileiro durante o fim do século XIX e início do século XX (Lopes; Cavalcanti, 2011, p. 31), mas que ganhou espaço no português brasileiro (Carvalho, 2019).

Sintetizamos, em seguida, estudos com base na variedade maranhense do português que buscaram compreender os padrões de variação do fenômeno em questão, e que dialogam com a pesquisa que aqui se desenvolve.

2.1 Formas de Segunda Pessoa do Singular

O sistema pronominal tem sido o sinal linguístico que mais evidencia as diferenças entre o português europeu (PE) e o português brasileiro (PB), principalmente, em pesquisas referentes à segunda pessoa do singular, que vêm demonstrando que a forma *você* vem se gramaticalizando, passando da função de pronome de tratamento para pronome sujeito, substituindo, assim, o pronome *tu* (Carvalho, 2019).

Faraco (2017[1966]) explica que os primeiros registros de *Vossa Mercê*, forma originária do pronome *você*, se deu no ano de 1666, como forma de tratamento ao rei de Portugal, mas que, com o passar do tempo, perdeu essa exclusividade, começando a circular entre a nobreza e a burguesia, o que resultou na popularização do seu uso, gerando o “desgaste semântico e formal de tal forma nominal de tratamento, a ponto de originar um novo pronome (*você*) que, por sua vez, convive, na sincronia atual do PB, com o pronome *tu* nos mesmos domínios funcionais” (Rumeu, 2012, p. 37-38).

O desenvolvimento da forma linguística *você*, no Brasil, se deu de modo diferente em relação a Portugal. Segundo Nascentes (1956, p. 118), não temos uma data específica de quando houve o surgimento da forma *você* no Brasil e em Portugal, mas Lopes e Cavalcante (2011, p. 36) apontam para o fato de que, “em termos históricos, *Vosmecê*, *mecêa*, *vosse*, *você* e a própria forma *Vossa Mercê* aparentemente chegaram ao Brasil sem a força cortês dos primeiros tempos – séculos XIII-XIV”.

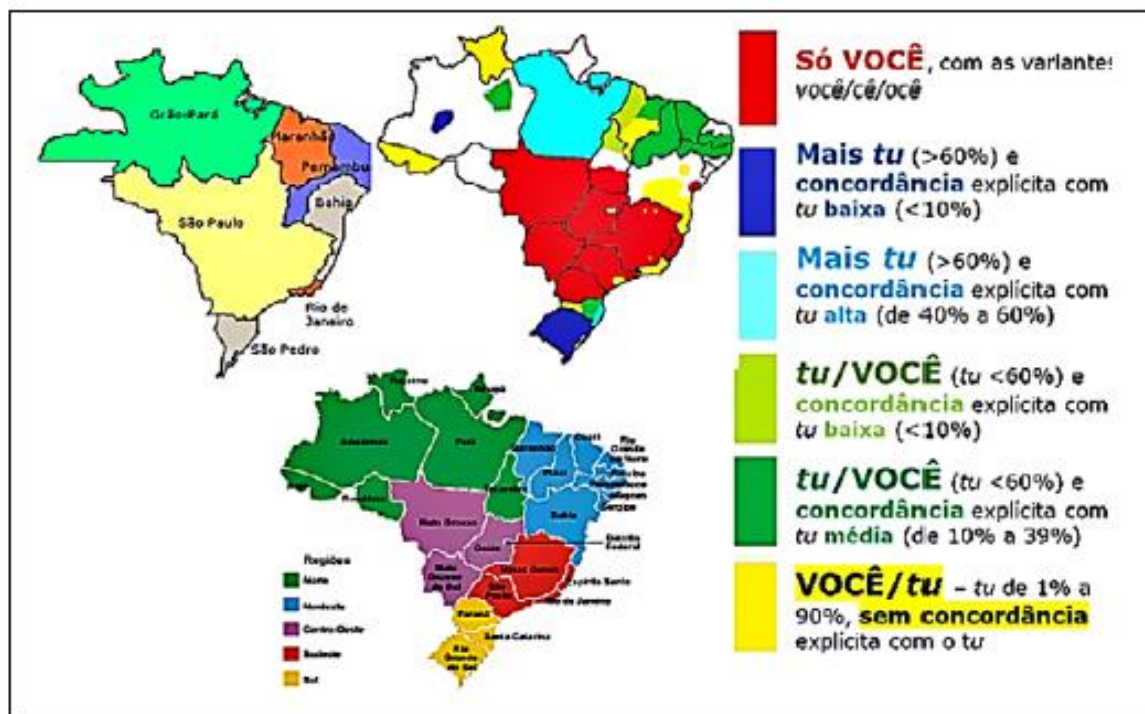
Rumeu (2004) investigou as formas de tratamentos em cartas de circulação pública (cartas oficiais) e de circulação privada (cartas não oficiais) escritas no Rio de Janeiro entre os séculos XVIII e XIX. Foram analisadas as formas nominais *vossa mercê*, *vossa excelência*, *vossa majestade*, *senhor* e *vossa senhoria*, além das formas pronominais de tratamento *tu*, *vós* e *você*. A partir do levantamento de 551 dados na posição de sujeito, a autora fez algumas constatações gerais: i) nas cartas oficiais do século XVIII há preferência (85%) para a forma pronominal *Vossa Excelência*; ii) nas cartas não oficiais há preferência (50%) para a forma *você*; iii) o uso da forma *tu* aparece nas cartas com baixa frequência tanto para as cartas oficiais (0,5%) e para as cartas não oficiais (0,7%). Ademais, a autora concluiu que, durante o século XVIII, existiam relações – assimétricas (falantes de nível superior para inferior) e simétricas (falantes que ocupam o mesmo nível) – sobre essas formas, de um modo mais claro, os resultados comprovam que o pronome *você* se faz presente nas relações assimétricas, enquanto o pronome *tu* se faz presente nas relações simétricas.

Lopes e Duarte (2003), por sua vez, analisaram peças teatrais brasileiras durante os séculos XVIII e XIX. Em geral, buscam identificar os fatores linguísticos e extralinguísticos que aceleraram o processo de gramaticalização de *vossa mercê* > *você* no português do Brasil, ocasionando sua inserção no nosso sistema pronominal. Verificaram que o uso da forma *você* apresentava baixa frequência (13%) de uso, ao passo que as formas *vossa mercê* (33%), *tu* (29%) e *vós* (25%) apresentavam frequência próximas durante o século XVII. Observaram também que, durante os séculos XVIII, e primeira metade do século XIX, a forma *tu* ganhou um aumento significativo de (63%), chegando a 90% na primeira metade do século XIX, mas houve uma queda na segunda metade do século XIX. Com isso, as formas *Vossa Mercê* e *vós* passam a deixar de ser usadas durante a segunda metade do século XIX.

Na variedade do português brasileiro, foram diversas as mudanças observadas na estrutura do sistema pronominal. As mudanças diacrônicas ocorridas na língua evidenciam as diferenças entre o PE e o PB, o que fez com que muitos pesquisadores buscassem estabelecer os padrões que diferenciam uma variedade linguística da outra. Volta-se o olhar, especialmente, para as pesquisas desenvolvidas por Scherre *et al.* (2015, p. 142), que propuseram a estruturação

do sistema pronominal brasileiro em seis subsistemas, levando em consideração a presença ou ausência das formas de segunda pessoa e a concordância verbal do pronome tu:

Figura 7 – Mapa dos subsistemas dos pronomes de segunda pessoa do singular do português brasileiro falado no final do século XX e início do século XXI, o Brasil de 1709 e o Brasil do século XXI.



Fonte: Scherre *et al.* (2015, p. 142)

Scherre *et al.* (2015) comparam o mapa proposto para ilustrar os subsistemas pronominais usados nas variedades linguísticas do país ao próprio mapa geopolítico do Brasil, datado de 1709. Tal comparação possibilitou a produção de uma série de discussões a respeito da distribuição dessas formas linguísticas. As autoras explicam que

na grande área em vermelho, simbolizando o subsistema só VOCÊ no Brasil Central, corresponde ao estado de São Paulo do mapa de 1709, que abarca os atuais estados de Rondônia, Mato Grosso, Mato Grosso do Sul, Tocantins, Minas Gerais, Goiás, São Paulo e parte do Paraná. [...] A região Nordeste se mostra bastante diversificada, com, pelo menos, quatro subsistemas: para o Maranhão, subsistema 4 (*tu/VOCÊ*, *tu* com concordância baixa), subsistema 5 (*tu/VOCÊ*, *tu* com concordância média) e subsistema 6 (*VOCÊ/tu*, *tu* sem concordância); para a Bahia, subsistema 1 (só VOCÊ) e subsistema 6 (*VOCÊ/tu*, *tu* sem concordância); para o Piauí, Ceará, Paraíba e Pernambuco, projeção do subsistema 5 (*tu/VOCÊ*, *tu* com concordância média), sem dados à época para o desenho do Rio Grande do Norte, de Alagoas e de Sergipe. (Scherre *et al.*, 2015)

Na região Nordeste, em especial o estado do Maranhão, *locus* da nossa pesquisa, são admitidas, para a expressão de formas de segunda pessoa do singular, cinco possibilidades, que funcionam sintaticamente em posição de sujeito. Ao observar o mapa, concluímos que no

Maranhão há um sistema bastante diversificado, representado pelas formas *você* (com as variantes *você/cê/ocê*) *tu* com concordância e *tu* sem concordância. De acordo com Scherre *et al.* (2015), essas “formas pronominais competem entre si nos mesmos espaços discursivos”, o que vem a ser comprovado com as pesquisas voltadas para a produção linguística da forma de segunda pessoa (Herênio, 2006; Alves, 2010; 2015; Carneiro, 2011; Miranda, 2014). Tais pesquisas nos permitem refletir e voltar o nosso olhar mais minuciosamente para a variedade maranhense, a fim de compreender a singularidade presente na fala maranhense.

Na sequência, descrevem-se os trabalhos de produção de segunda pessoa do singular realizados no Maranhão, voltando nossa atenção para os resultados mais expressivos de cada pesquisa. Esses trabalhos podem evidenciar o porquê de ainda existir no imaginário dos falantes o mito ‘de que o maranhense fala o melhor português’.

2.2 Segunda Pessoa do Singular no Maranhão

A revisão de trabalhos realizados no estado do Maranhão sobre a variação dos pronomes de segunda pessoa do singular importa para a pesquisa que aqui se desenvolve, uma vez que estudos dessa natureza descrevem e analisam quantitativamente o fenômeno da segunda pessoa do singular como marcadores de categorias sociais, ao utilizarem metodologias de coletas de falas, a partir de uma amostra representativa da comunidade.

Em outras palavras, essas pesquisas sociolinguísticas categorizam as formas linguísticas de segunda pessoa como traço de identificação social, apresentando resultados que mantêm padrões semelhantes, consequentemente, se compreende que o falante, ao favorecer uso de uma forma linguística, projeta sua filiação à comunidade linguística em que está inserido.

Os resultados dessas pesquisas contribuem para o nosso estudo sobre as formas pronominais de segunda pessoa, no sentido de nos auxiliarem a identificar quais categorias sociais que atuam no padrão linguísticos dessas formas, e de modo mais preciso, a compreender quais as práticas construídas na dinâmica das interações sociais.

No contexto maranhense, Herênio (2006) teve como objetivo principal investigar a variação entre as formas linguísticas dos pronomes pessoais do caso reto *tu* e *você* no português brasileiro falado contemporaneamente em duas regiões diferentes do país: Uberlândia - MG e Imperatriz - MA. A hipótese inicial da autora era a de que o pronome *tu* ocorresse com muita frequência em Imperatriz, comparativamente a Uberlândia, localidade em que não eram esperadas ocorrências desse pronome. A autora construiu um *corpus* com grupos de 43 informantes da cidade de Uberlândia e 43 informantes da cidade de Imperatriz, totalizando 86

entrevistas de fala semiespontânea, em que os participantes foram estratificados de acordo com sua faixa etária e classe social.

Segundo Herênio (2006), a classe social não foi um fator que se correlacionasse às variantes analisadas, pois não há diferenças significativas quanto ao uso das formas entre as classes, que empregaram mais o pronome *você*, com percentuais de 80,4%, 94,5% e 84,2%, correspondentes às classes alta, média e baixa, respectivamente. Ressalta-se que a autora “utilizou o Critério Brasil do IBGE para o levantamento socioeconômico dos informantes. Esse critério utiliza a tabela de bens de consumos e graus de escolaridade do chefe de família” (Herênio, 2006, p. 62).

Em relação à faixa etária, os falantes da faixa etária de 20 a 30 anos apresentaram uma taxa de 18% no emprego da forma *tu* e 82% na forma *você*. Os representantes da faixa etária de 31 a 45 anos apresentaram 7% de emprego da forma *tu* e 93% de emprego da forma *você*. Aqueles indivíduos com idade com 46 anos ou mais apresentaram uma taxa de 13% emprego no emprego da forma *tu*, no total da amostra, e 87% de emprego da forma *você*.

Alves (2010), por sua vez, analisou o uso das formas *tu* e *você*, em referência à segunda pessoa do discurso, no Maranhão. O objetivo mais geral foi o de fazer uma ‘fotografia sociolinguística’ do português falado no Maranhão referente ao uso desses pronomes. Para tanto, utilizou entrevistas realizadas pelo Atlas Linguístico do Maranhão (ALiMA) nas mesorregiões Norte (São Luís e Pinheiro), Centro (Bacabal e Tumtum) e Sul (Alto Parnaíba e Balsas) do estado do Maranhão. Os falantes, no total de 28, foram estratificados por sua escolaridade (ensino fundamental e ensino superior, em São Luís, e ensino fundamental nas demais localidades do estado), faixa etária (faixa etária I, de 18 a 30 anos, e faixa etária II, de 50 a 65 anos) e sexo (masculino e feminino).

Os resultados apontaram que a forma pronominal *você* é a mais frequente no falar maranhense, ao apresentar um percentual de 61,6% das ocorrências, enquanto a forma *tu* apresentou 38,4% dos dados analisados.

Quando são verificadas as distribuições dos dados entre as localidades estudadas, observamos que as cidades de Bacabal e Balsas foram as que mais realizaram a forma *tu* (56,5% e 56,7%, respectivamente).

Tabela 1 - Distribuição geral das formas *tu* e *você*, *ocê* e *cê* nas localidades estudadas.

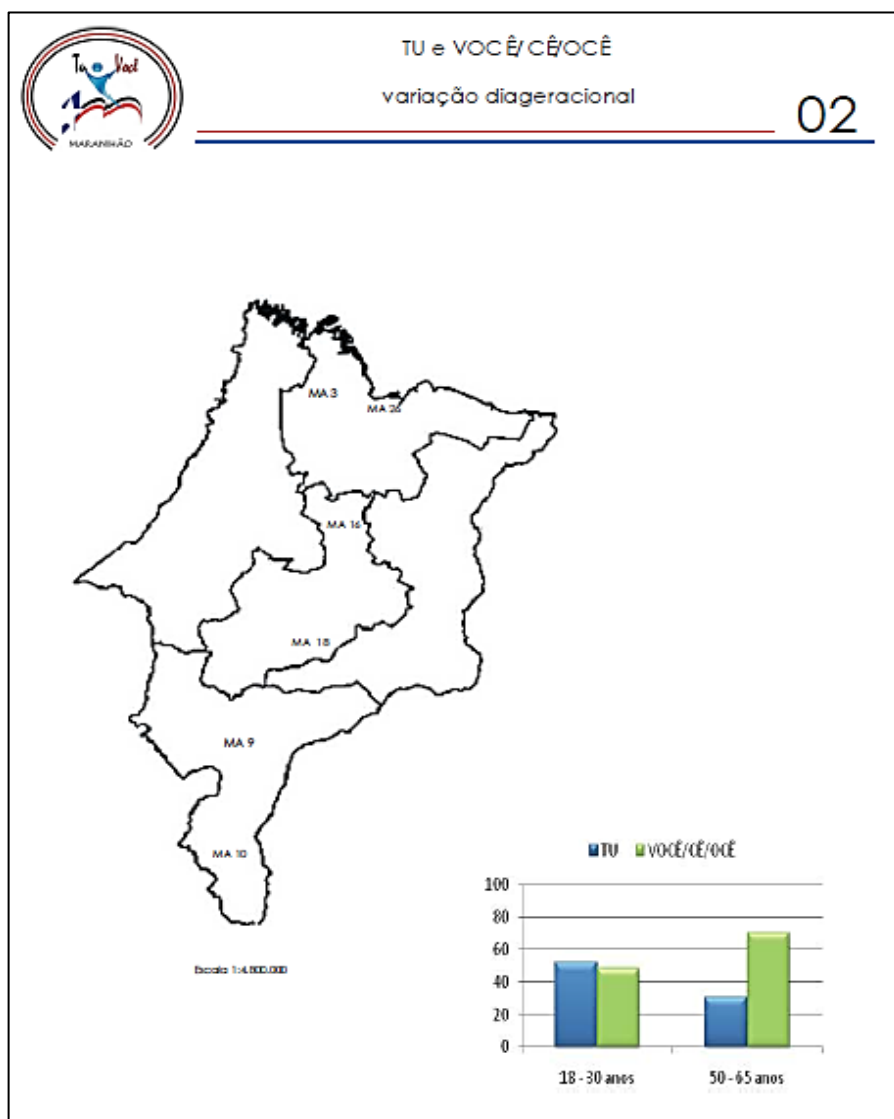
LOCALIDADE	TU	VOCÊ	CÊ	OCÊ	TOTAL
Bacabal	13 56,5%	9 39,1%	1 4,3%	0 0%	23
Balsas	17 56,7%	13 43,3%	0 0%	0 0%	30
Alto Parnaíba	5 15,2%	20 60,6%	4 12,1%	4 12,1%	33
Tuntum	15 35,7%	18 42,9%	6 14,3%	3 7,1%	42
Pinheiro	31 36,9%	46 54,1%	7 8,2%	0 0%	84
São Luís	45 38,8%	61 52,2%	10 8,6%	0 0%	116
Total	126 38,8%	168 51,1%	27 8,5%	7 2,1%	328 100%

Fonte: Elaboração própria, com base em Alves (2010, p. 65)

Alves (2010) evidenciou ainda que as categorias macrosociais dos falantes, a exemplo de localidade, faixa etária, sexo, tipo de relato e tipo de referência, são fatores que condicionam a variação do uso dessas variantes, chegando à conclusão de que a alternância do *tu* e *você* é condicionada pela faixa etária dos falantes e pela localidade à qual pertence.

No que diz respeito à faixa etária, Alves (2010, p. 71) revela que os mais jovens realizaram mais as formas de *tu*, indicando um processo de ressignificação do uso dessa forma, em contraposição à faixa etária de falantes com mais idade, que tendem mais ao uso do pronome *você*. Uma possível explicação está no fato de que os mais jovens são os que utilizam as formas mais inovadoras, com verbo sem marca da concordância:

Figura 8 - Ocorrências das formas *tu* e *você/ocê/cê* de acordo com a faixa etária.



Fonte: Alves (2010, p. 70)

Ainda que não tenha sido selecionada como uma variável estatisticamente significativa, os resultados para a escolaridade serão reportados aqui porque pensa-se que contribuirão para os interesses do experimento de percepção, cujo desenho vai-se detalhar no capítulo 3.

Alves (2010) verificou que, na distribuição dos dados, os falantes mais escolarizados são os que mais utilizaram a forma *você*, enquanto os falantes com menos escolarização não se diferenciam muito quanto ao uso do pronome *tu* ou *você*, com as taxas muito próximas entre si. Embora não tenham sido consideradas estatisticamente significativa, essas distribuições interessam, sobretudo porque subsidiam o interesse em verificar se os falantes podem ser considerados muito ou nada escolarizados a depender se usam *tu* com concordância, *tu* sem

concordância, a forma *você* e *cê*, o que vai nos permitir uma compreensão mais profunda do processo e/ou variação de mudança linguística.

Alves (2010) fez uma análise apenas com os dados coletados em São Luís, isso porque apenas na amostra dessa localidade havia falantes com ensino universitário. Ressalte-se que a proposta dessa análise importa, sobretudo por causa do imaginário popular de que é em São Luís que se fala o melhor português do Brasil, afirmação baseada, sobretudo, pelo uso das formas pronominais de segunda pessoa do singular que vêm à fala das pessoas quando elas afirmam coisas dessa natureza (Miranda, 2014; Santos, 2015; 2020). Nos dados analisados houve frequência da forma *tu com concordância*. Foram analisadas todas as variáveis, mas que, para o interesse desta pesquisa, e para as discussões que serão feitas, importam os resultados para a escolaridade.

Tabela 2 - Concordância verbal com o *tu* em São Luís, de acordo com a escolaridade.

Escolaridade	Aplicação/ Total		Peso relativo
	N	%	
Fundamental	1/19	5,3	0.17
Superior	12/26	46,2	0.76
Total	13/45		

Fonte: Elaboração própria, com base em Alves (2010, p.104)

Alves (2010) observou que a forma *tu*, com marca de concordância verbal, é favorecida entre falantes ludovicenses com maior escolaridade (0.76), revelando que há um prestígio em relação ao grupo de maior *status* social quanto ao uso dessa variante, ou seja, os ludovicenses mais escolarizados compreendem a relevância dessa variável na vida local, fazendo uso linguístico dessa variante como construção/expressão de identidade (Alves, 2010).

Carneiro (2011) objetivou investigar a alternância das formas de tratamento *tu/você* no português ludovicense. Para tanto, analisou uma amostra, organizada de acordo com os critérios de escolaridade (fundamental, médio e superior), sexo/gênero (masculino e feminino), faixa etária (faixa etária I, de 15 a 25 anos, faixa etária II, de 26 a 55 anos, e faixa etária III, acima de 55 anos) e classe social (média alta, média, média baixa e baixa), totalizando 90 falantes, com dados alcançados em situações espontâneas da fala em São Luís, Maranhão.

De um modo geral, os resultados obtidos constataram que essas formas coocorrem na fala dos ludovicenses, sendo de maior predominância o pronome *tu*, apresentando uma frequência de 69,3% do total das ocorrências, seguido do pronome *você*, com uma frequência de 30,6% do total dos dados. Os dados referentes à concordância verbal com a forma *tu*, o verbo

com a marca de concordância apresenta percentual de 5,73%, contra o percentual de 91,67% para o verbo sem essa marca.

Carneiro (2011) observou que os ludovicenses com maior grau de escolaridade tendem a favorecer o uso do pronome *você*, enquanto os falantes de escolaridade menor tendem a favorecer o uso do pronome *tu*. Os resultados da autora vão de encontro àqueles alcançados por Alves (2010).

Tabela 3 - Ocorrências das formas *tu/você*, conforme a Escolaridade, entre os falantes ludovicenses

Nível de escolaridade	Em relação ao conjunto da amostra			
	Tu		Você	
	N	%	N	%
Fundamental	72/192	37,5%	25/85	30%
Médio	71/192	36,9%	13/85	15,3%
Superior	49/192	25,5%	46/85	54,1%
Total	192/192	100%	85/85	100%

Fonte: Elaboração própria, com base em Carneiro (2011, p. 95)

Carneiro (2011, p. 95), ao se referir aos falantes com mais escolaridade, reitera que “esse resultado pode ser explicado pelo fato de que estas pessoas tenham entrado na escola utilizando *tu*, mas que, à proporção que foram galgando níveis mais altos de escolaridade, estejam passando a usar mais a forma *você*”. Os falantes com escolaridade mais baixa estariam, por sua vez, sensíveis às pressões escolares impostas ao uso da gramática normativa, ao passo que os falantes de escolaridade mais alta estão passando por um processo de variação, que pode ser explicado devido às necessidades de estarem inseridos em ambientes diversos, em que a forma *você* tende a ser mais aceita pelo grupo. A autora verificou, entre outros aspectos, que, embora tenham sido menos numerosas, as ocorrências de *tu com concordância* foram produzidas somente por pessoas com nível de ensino superior (63,64%) ou médio (36,36%), a maioria com idade entre os 26 e os 55 anos, quase sempre em contextos de total informalidade, com amigos e parentes.

Quanto à faixa etária, Carneiro (2011) observou que a tendência geral do uso das formas linguísticas se mantém para as faixas etárias de 15 a 25 anos (com 32,29% e 16,47% para a realização das formas *tu* e *você*, respectivamente) e de mais de 55 anos (com 19,27% para o uso de *tu*, contra 16,47%, para o uso de *você*). Na faixa intermediária, dos 26 aos 55 anos, a tendência se modifica, o que permitiu a Carneiro (2011) observar que, em São Luís, as pessoas nesta faixa de idade, proporcionalmente, privilegiam mais a utilização do pronome *você* (67,06% para a realização de *você*, contra 48,44%, para *tu*).

Tabela 4 - Ocorrências das formas *tu/você*, conforme a Faixa Etária, entre os falantes ludovicenses.

Faixa Etária	Em relação ao conjunto da amostra			
	Tu		Você	
	N	%	N	%
15 a 25 anos	62/192	32,2%	14/85	16,4%
26 a 55 anos	93/192	48,4%	57/85	67%
Acima de 55 anos	37/192	19,2%	14/85	16,4%
Total	192/192	100%	85/85	100%

Fonte: Elaboração própria, com base em Carneiro (2011, p. 96)

Quanto à classe social, a pesquisa de Carneiro (2011) mostrou que o uso do *tu* é muito mais frequente em relação ao uso do *você* quando se comparam falantes da denominada classe baixa e da classe média baixa. Entre os falantes de classe média e de classe média alta, os números revelam que os ludovicenses tendem a fazer mais uso do pronome *você*, comparativamente ao pronome *tu* (tabela 6):

Tabela 5 - Ocorrências das formas *tu/você*, de acordo com a Classe Social, entre os falantes ludovicenses.

Classe social	Em relação ao conjunto da amostra			
	Tu		Você	
	N	%	N	%
Média alta	17/192	8,5%	15/85	17,6%
Média	45/192	23,4%	37/85	43,5%
Média baixa	59/192	30,7%	19/85	22,3%
Baixa	71/292	36,9%	14/85	16,4%
Total	192/192	100%	85/85	100%

Fonte: Elaboração própria, com base em Carneiro (2011, p. 99)

Segundo Carneiro (2011, pag. 99), esses resultados podem estar ancorados na premissa de que as pessoas pertencentes a classes alta, média alta e média têm maior acesso aos bens culturais e chegam mais facilmente ao ensino superior, tendo uma maior preferência de uso pela forma *você*, comparativamente à forma *tu*. Uma outra explicação para esse resultado é a de que possivelmente “esteja havendo uma pressão no sentido de atribuir ao pronome *você* o estatuto de variante prestigiada e que vem de fora, que vem se impondo, através da mídia, através da vinda de pessoas com maior poder aquisitivo de outros centros do país”.

Miranda (2014) realizou um estudo voltado para o exame da concordância verbal na segunda pessoa do singular na fala maranhense, nas cidades de Caxias e em São Luís. A amostra foi composta por: (1) corpus do projeto Atitudes Linguísticas dos Falantes do Maranhão (ALFMA), que integra a amostra Caxiense, e (2) *corpus* da amostra ludovicense, disponibilizada por Carneiro (2011), que contém os dados de São Luís.

A pesquisa empreendida por Miranda (2014) caracteriza-se como um estudo de caso, que analisa o imaginário de um grupo de pessoas da cidade de Caxias em relação à afirmação popular de que no Maranhão se fala o melhor português.

Para uma melhor compreensão dos dados, o autor analisa os resultados de cada amostra separadamente, constatando que os falantes realizaram o verbo com a marca de concordância apenas em 4,5% e 5,1% nas amostras das cidades, respectivamente.

Tabela 6- A concordância de segunda pessoa do singular na amostra Caxiense.

Concordância verbal	Total de ocorrências	%
Concorda	6	4,5%
Não concorda	127	95,5%
Total	133	100%

Fonte: Elaboração própria, com base em Miranda (2014, p. 101)

O total de ocorrências na amostra caxiense foi menor, se comparada à amostra ludovicense, mas com base nos resultados das duas amostras foi possível observar que há fortes indícios de uma mudança linguística no paradigma do sistema pronominal em análise.

Tabela 7 - A concordância de segunda pessoa do singular na amostra Ludovicense.

Concordância verbal	Total de ocorrências	%
Concorda	10	5,1%
Não concorda	187	94,9%
Total	197	100%

Fonte: Elaboração própria, com base em Miranda (2014, p. 101)

Miranda (2014) apresenta dados da amostra da variável de segunda pessoa do singular quando realizado na função de pronome com função sintática de sujeito. Os resultados para a análise da amostra caxiense nos interessam na medida em que nos darão auxílio para as nossas definições metodológicas que serão apresentadas no próximo capítulo. Os resultados abaixo nos mostram que as variantes *você*, *tu* e *cê*, coocorrem na fala dos caxienses, vejamos:

Tabela 8 - Realização do pronome de segunda pessoa do singular, na amostra Caxiense

Tipo de pronome	Aplicação	Total da ocorrência	%
Pronome <i>você</i>	771	771	67,6%
Pronome <i>tu</i>	6	133	11,6%
Pronome <i>zero</i>	9	225	19,7%
Pronome <i>cê</i>	12	12	1,1%
Total	798	1141	100%

Fonte: Elaboração própria, com base em Miranda (2014, p. 99)

Com a realização de sua pesquisa, o autor concluiu que, em Caxias, o pronome *você* (67,6%) é o mais frequente na amostra analisada, seguido do *pronome zero*¹³ (19,7%), da forma *tu* (11,6%), e, por último, do pronome *cê* (1,1%).

Ao analisar apenas a realização de 133 dados com o pronome *tu*, com ou sem concordância, Miranda (2014) observou que, desse total, 6 dados, 4,5% do total de casos, foram expressos com concordância verbal, enquanto a grande maioria dos dados, 127, foram realizados com a ausência da marca de concordância, configurando, assim, 95,5% do total. Os grupos de fatores selecionados pelo programa *GoldVarbX* (Tagliamonte *et al.*, 2012) como variáveis predictoras que condicionam a concordância verbal da segunda pessoa do singular na fala caxiense foram três, todas de natureza estrutural, a saber o *tipo de sentença*, o *número de sílabas do verbo* e a *frequência do verbo*, e apenas uma variável de natureza externa ao sistema: *escolarização*.

No sentido de analisar a seleção da escolarização dos falantes para explicar o fenômeno em questão, a pesquisa revelou que os falantes do ensino médio favorecem a regra de concordância, ao passo que os falantes do ensino superior desfavorecem a regra. Com esses dados, Miranda (2014, p. 114), não confirmou a hipótese que esperava de que “os falantes do ensino superior com mais anos de escolaridade, maior experiência com a língua realizassem mais a aplicação da regra de concordância verbal”.

Tabela 9 - A concordância de segunda pessoa do singular e a escolarização, na amostra Caxiense.

Escolarização	Aplicação	Total da ocorrência	P. R
Ensino Médio	3	27	.59
Ensino Superior	3	32	.41
Total	6	59	<i>Input .011</i>

Fonte: Elaboração própria, com base em Miranda (2014, p.114)

Além do estudo de produção, Miranda (2014) também buscou acessar as avaliações dos caxienses quanto a crença de que são os maranhenses aqueles que melhor falam o português brasileiro. Para isso, Miranda observou que, de certa maneira, tal crença existe, uma vez que os informantes tiveram uma tendência de avaliar positivamente tal mito linguístico. Os resultados para o estudo de avaliação das crenças e das atitudes linguísticas da amostra caxiense foram obtidos a partir de questões direcionadas aos informantes: “*Você acredita que o maranhense*

¹³ Denominado na pesquisa como sujeito não preenchido, a exemplo de “Ø fez a atividade”.

fala o melhor português?”, “*O português que você fala é melhor do que o falado pelos seus pais?*”.

Em linhas gerais, os resultados obtidos na pesquisa revelaram que ainda há no imaginário dos falantes o mito ‘de que o maranhense fala o melhor português’, e tal imaginário está relacionado às crenças e às atitudes dos informantes acerca do “sotaque”, em expressões como ‘falar sem gíria’, ‘falar bem’, ‘o modo de falar’, ‘a expressão’, ‘o falar mais bonito’, ‘falar muito bem’, ‘o português mais explicado’, ‘linguagem muito bonita’.

Em suma, apresentamos os estudos empíricos de análises sociolinguísticas sobre a variável da segunda pessoa do singular no Maranhão. Tais pesquisas embasam o interesse da proposta que aqui se lança, que é a de se realizar um estudo de percepção sociolinguística, a fim de que se acessem quais potenciais significados sociais emergem quando diferentes ouvintes são estimulados, inconscientemente, a julgar falantes por meio do uso de um sinal linguístico, a saber, aqui, o uso dos pronomes *tu* e *você* na referência à segunda pessoa do discurso, no singular.

CAPÍTULO 3

3 MONTAGEM DO EXPERIMENTO

Um dos principais interesses dos estudos de percepção sociolinguística reside no fato de que se busca compreender quais elementos estão envolvidos numa comunidade quando um ouvinte atribui um significado social a uma forma linguística. É fato indiscutível que a percepção é parte do processo do uso da língua, e que faz parte das nossas relações sociais atribuir valores às formas linguísticas movimentadas em nossas interações sociolinguísticas. Cabe ao sociolinguista, portanto, valer-se de métodos bem definidos para acessar e estabelecer padrões aos significados sociais potencialmente associados a certas formas linguísticas, que podem mudar conforme as práticas sociais da comunidade à qual esse falante/ouvinte se vincula.

Neste capítulo, descreve-se como se desenvolveu a montagem do experimento de percepção que pretende acessar os significados sociais possivelmente associados ao uso dos pronomes *tu com concordância*, *tu sem concordância*, *você* e *cê*, no português falado no Maranhão, especialmente na cidade de Caxias.

De um modo geral, busca-se entender se: i) O uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu foste* e *tu gostaste*, é percebido pelos ouvintes caxienses como uma variante característica de sua localidade; ii) O uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu vais* e *tu és*, é percebido pelos ouvintes caxienses como índice de maior grau de escolarização; iii) O uso das formas pronominais, *tu* com concordância ou *você*, indiciam diferentes graus de formalidade para os ouvintes caxienses.

Evidencia-se a informação de que essas perguntas foram elaboradas a partir dos resultados alcançados por estudos de produção, cujos resultados foram apresentados no capítulo anterior e mostraram que a variável de segunda pessoa do singular se correlaciona significativamente aos graus de formalidade entre os interlocutores, além de informações sociais dos informantes como escolarização, localidade e outras dimensões que se referem ao *status* dos falantes.

Tais resultados nos possibilitam propor a realização do estudo de percepção sociolinguística que aqui é proposto, a fim de se discutir o papel do ouvinte caxiense não apenas no processo de produção, como também na atribuição de significados sociais às formas linguísticas objeto deste estudo, uma vez que se entende, como Eckert (2005), que toda variação tem potencial para receber significado social.

O aparato teórico-metodológico desta pesquisa é o da Sociolinguística Variacionista, especialmente dos estudos Labov (2006[1966], 2008[1972]). O interesse é o de verificar se as dimensões sociais, tais como *grau de formalidade*, *escolarização*, *localidade* e atributos como *gentileza* e *inteligência* são alterados para os ouvintes a depender do uso das formas pronominais pelos falantes participantes do estudo. A proposta de elaboração dos estímulos que foram usados no experimento é apresentada a seguir, em paralelo à apresentação das definições para a realização do experimento, bem como o seu desenho, a formulação e a aplicação do questionário de percepção, pensado com base na técnica *matched-guise* (Lambert et al, 1960; Campbell-Kibler, 2006; Mendes; Oushiro, 2015; e Mendes, 2016).

3.1 Definições acerca dos falantes

Com a intenção de verificar sistematicamente quais possíveis significados sociais estão associados ao uso da variável da segunda pessoa *tu com concordância*, a exemplo de “*tu és mesmo filho de uma égua mesmo*”, *tu sem concordância*, como em “*tu desenha outro nariz*”, *você*, a exemplo de “*você sabe falar bom dia*” e *cê*, como em “*cê ficou triste por isso*”, elaborou-se um experimento de percepção sociolinguística, cujo desenho é aqui descrito.

Trata-se de uma adaptação da técnica dos estímulos pareados (Lambert et al, 1960; Campbell-Kibler, 2006; Oushiro, 2015; Soriano, 2016; Canever, 2017; Mendes, 2017; Santos, 2020), que consiste na elaboração de estímulos, geralmente auditivos, extraídos da fala de um certo número de falantes, manipulados digitalmente, no sentido de que sejam iguais em todos os aspectos, diferenciando-se apenas com respeito à variável linguística que esteja sendo analisada. Esses diferentes falsos pares são apresentados a diferentes grupos de ouvintes, que além de darem suas impressões gerais sobre os estímulos que ouvem, avaliam-no por meio do questionário de percepção apropriado para a pesquisa em foco. A técnica consiste em uma avaliação indireta em que o falante não tem consciência do foco da pesquisa, isto é, do objeto que está sendo investigado.

A técnica dos estímulos pareados levantou uma série de investigações sistemáticas das atitudes linguísticas e passou a ser aplicada em diversos outros estudos interdisciplinares à Psicologia Social, nos quais incluem-se os trabalhos sobre percepção (Oushiro, 2015).

Para a nossa pesquisa, foram elaborados “falsos pares” que se diferenciaram apenas quanto à morfologia pronominal. Em relação ao conteúdo das sentenças, trabalhamos com uma temática que fosse relacionado a todos.

Para tanto, foram gravados quatro falantes, dois homens e duas mulheres, nascidos e residentes da cidade de Caxias, todos com ensino superior completo e faixa etária entre 25 a 30

anos. O perfil social aparentemente homogêneo dos falantes, exceto por seu sexo, é assim definido com o objetivo de minimizar o efeito de outras variáveis que possivelmente possam agir sobre a percepção dos ouvintes, tal como se estabeleceu para o conteúdo dos trechos.

A elaboração das sentenças justifica-se porque o presente estudo busca acessar significados sociais de variáveis morfossintáticas de segunda pessoa do singular, e precisaríamos de contextos específicos, com realizações específicas para que garantíssemos a análise dos contextos estruturais, como por exemplo, o tempo verbal, a quantidade de sílabas do verbo, a saliência fônica, conjugação verbal, tipo de sentença e contexto fônico precedente, como se pode verificar a seguir, no item 2.2. Desse modo, não conseguimos extrair sentenças de falas, como esses conjuntos de fatores específicos, das entrevistas sociolinguísticas já realizadas. O que nos possibilitou em elaborar sentenças-alvos e pedir a falantes anteriormente contatados que realizassem as sentenças o mais natural possível.

3. 2 Preparação dos estímulos

As sentenças foram gravadas a partir de um aparelho celular Iphone 13s pro max. Os quatro falantes convidados que cederam suas vozes foram levados a um ambiente silencioso, no intuito de que não houvesse interferência de ruídos externos durante a gravação dos estímulos.

Abaixo, são reproduzidas as quatro versões das sentenças que foram gravadas, considerando-se as variantes que interessam a este estudo: *tu com concordância*, *tu sem concordância*, *você* e *cê*. Essas variantes foram inseridas em trechos cujos conteúdos eram distintos para cada uma delas, também com o intuito de que os ouvintes não desviassem sua atenção para outros aspectos dos estímulos que não fossem os sinais linguísticos que aqui importam. Assim, solicitou-se a cada um dos falantes que gravassem as diferentes sentenças, até que se obtivesse um trecho lido de modo mais espontâneo possível. Cada falante foi gravado três vezes falando as quatro versões. Selecionou-se uma melhor versão dentre as quatro gravadas, servindo de sentença-base para manipulação do áudio e posterior elaboração dos falsos pares.

Quadro 1 - Estímulos para as gravações dos falantes.

Homem 1 (Breno)	Alguém pode passar e carregar a bola, então tu corres pra pegar logo. Teus pais vão brigar de novo, pois tu perdes brinquedo demais Alguém pode passar e carregar a bola, então tu corre pra pegar logo. Teus pais vão brigar de novo, pois tu perde brinquedo demais. Alguém pode passar e carregar a bola, então você corre pra pegar logo. Teus pais vão brigar de novo, pois você perde brinquedo demais Alguém pode passar e carregar a bola, então cê corre pra pegar logo. Teus pais vão brigar de novo, pois cê perde brinquedo demais
Mulher 1 (Luciana)	Eu brinquei de tudo que tu podes pensar. Tu sabes brincar de amarelinha? Era a minha brincadeira preferida. Eu brinquei de tudo que tu pode pensar. Tu sabe brincar de amarelinha? Era a minha brincadeira preferida. Eu brinquei de tudo que você pode pensar. Você sabe brincar de amarelinha? Era a minha brincadeira preferida. Eu brinquei de tudo que cê pode pensar. Cê sabe brincar de amarelinha? Era a minha brincadeira preferida.
Homem 2 (Raimundo)	Tu compraste pipa onde? Lá no seu Zé? Tu botaste pra subi? Ouvi dizer que ele não sabe fazer pipa. Tu comprou pipa onde? Lá no seu Zé? Tu botou pra subi? Ouvi dizer que ele não sabe fazer pipa. Você comprou pipa onde? Lá no seu Zé? Você botou pra subi? Ouvi dizer que ele não sabe fazer pipa. Cê comprou pipa onde? Lá no seu Zé? Cê botou pra subi? Ouvi dizer que ele não sabe fazer pipa.
Mulher 2 (Paula)	Tu ficaste brincando sempre no meio da rua, é perigoso. Por isso que quando tu chamaste pra jogar bola, eu não fui. Tu ficou brincando sempre no meio da rua, é perigoso. Por isso que quando tu chamou pra jogar bola, eu não fui. Você ficou brincando sempre no meio da rua, é perigoso. Por isso que quando você chamou pra jogar bola, eu não fui. Cê ficou brincando sempre no meio da rua, é perigoso. Por isso que quando cê chamou pra jogar bola, eu não fui.

Os estímulos apresentados no quadro 1, acima, foram manipulados, utilizando-se o programa Audacity (2023), um software livre que disponibiliza de uma função que nos ajuda a editar as gravações originais, ou seja, substitui dos trechos de áudios originais as produções controladas da variável de segunda pessoa, através do comando “copiar/colar” do programa para a criação dos falsos pares. Dessa maneira, criaram-se, assim, 16 estímulos controlados, organizados em 4 conjuntos de dados, contendo 1 estímulo por cada falante.

Figura 9 - Organização dos estímulos

Grupo A1	H1	M1	H2	M2
Tu com concordância				
Tu sem concordância				
Você				
Cê				

Grupo A3	H1	M1	H2	M2
Tu com concordância				
Tu sem concordância				
Você				
Cê				

Grupo A2	H1	M1	H2	M2
Tu com concordância				
Tu sem concordância				
Você				
Cê				

Grupo A4	H1	M1	H2	M2
Tu com concordância				
Tu sem concordância				
Você				
Cê				

Fonte: Elaboração própria.

Na figura 10 apresenta-se a organização dos estímulos que foram assim organizados após serem feitas todas as combinações possíveis, a fim de que diferentes ouvintes ouçam todos os quatro falantes para este estudo realizando pelo menos uma das quatro variantes. Em verde-escuro estão os estímulos que formam cada grupo, e em verde-claro estão os “falsos pares” desses estímulos. Assim, cada “voz” foi ouvida em todas as formas pronominais, o que é possível porque foram elaborados 4 conjuntos de dados, em que cada ouvinte ouviu cada um dos quatro falantes apenas uma vez. Por exemplo, no grupo A1, o homem 1 foi ouvido apenas na versão *tu com concordância*; no grupo A2, ele foi ouvido realizando a versão *tu sem concordância*; no grupo A3, ele foi ouvido realizando a versão *cê*; e no grupo A4, ele foi ouvido realizando a versão *você*. O mesmo aconteceu com os demais falantes, seguindo-se a mesma lógica de apresentação dos estímulos. Desse modo, todos os falantes foram avaliados pelo ouvinte realizando algum dos disfarces elaborados. Em seguida à preparação e organização dos estímulos, elaborou-se o formulário de percepção, apresentado no tópico a seguir.

3.3 A formulação do teste de percepção

Com este experimento, busca-se acessar quais significados sociais estão associados às formas pronominais de segunda pessoa do singular; além disso, interessa saber se o uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu foste* e *tu gostaste*, é percebido pelos ouvintes caxienses como uma variante representativa de sua variedade linguística; se o uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu vais* e *tu és*, indicializa maior grau de escolarização; e se o uso das formas pronominais, *tu* com concordância ou *você*, indiciam diferentes graus de formalidade para os ouvintes caxienses.

Uma página de instrução foi criada para os participantes com as devidas informações necessárias, a fim de que os ouvintes respondessem ao formulário (Figura 10).

Figura 10 - Página de instrução do experimento

Pesquisa de Linguística

Olá, meu nome é Rayane de Andrade Rodrigues. Sou aluna do mestrado em Letras da Universidade Federal do Maranhão - UFMA.

O meu objetivo é verificar como as pessoas percebem a fala uma das outras.

Antes de passar algumas instruções sobre como responder o questionário, gostaria de lhe agradecer a por participar, com sua contribuição. Muito obrigada!

Dê preferências pelo uso de um fone quando for ouvir os áudios. Ouça esses áudios quantas vezes quiser, antes de responder às perguntas sobre as vozes dos áudios. Depois de ouvir quantas vezes quiser cada um dos áudios, deverá responder algumas perguntas a respeito desses áudios.

Em seguida, dará sua opinião sobre informações sociais desses ouvintes.

Por fim, responderá a algumas perguntas sobre você, mas sem precisar se identificar.

Lembramos que não tem respostas certas ou erradas. O que levaremos em consideração é a sua impressão sobre as vozes que ouvir.

Fique à vontade para entrar em contato comigo caso tenha qualquer dúvidas em relação ao experimento.

Email: rayane.andrade@discente.ufma.br

Fonte: Elaboração própria.

O questionário de percepção foi pensado com base nos modelos de testes de percepção dos estudos de Campbell-Kibler, 2006; Oushiro, 2015; Soriano, 2016; Mendes, 2016; Canever, 2017; Santos, 2020. Foram incluídas cinco escalas de variáveis quantitativas com graus de

diferenciais semânticos de seis pontos (escolarização¹⁴, inteligência, gentileza, formalidade e caxiense¹⁵), conforme se podem verificar:

Figura 11 - Primeira Parte do Formulário de Percepção

Essa pessoa soa como: *							
	1	2	3	4	5	6	
Nada escolarizado	☐	☐	☐	☐	☐	☐	Muito escolarizado

Essa pessoa soa como: *							
	1	2	3	4	5	6	
Nada inteligente	☐	☐	☐	☐	☐	☐	Muito inteligente

Essa pessoa soa como: *							
	1	2	3	4	5	6	
Nada gentil	☐	☐	☐	☐	☐	☐	Muito gentil

Essa pessoa soa como: *							
	1	2	3	4	5	6	
Nada formal	☐	☐	☐	☐	☐	☐	Muito formal

Essa pessoa soa como: *							
	1	2	3	4	5	6	
Nada caxiense	☐	☐	☐	☐	☐	☐	Muito caxiense

Fonte: Elaboração própria.

¹⁴ O termo *escolarização* foi usado para se referir a variável quantitativa e o termo *escolaridade* para a variável qualitativa. Entende-se que há diferenças entre os termos. De acordo com o dicionário *Houaiss* da língua portuguesa, a *escolarização* é fazer passar por aprendizado em escola. Ou seja, são os conhecimentos que estão intimamente ligados ao ensino escolar, mas também, por outros indicadores, como por exemplo, os bens de cultura das pessoas. No que diz respeito a *escolaridade*, segundo o dicionário *Houaiss* da língua portuguesa, é o tempo de frequência ou permanência do indivíduo no ambiente escolar.

¹⁵ Inclui-se a escala *caxiense* a fim de que se observe os ouvintes caxienses se reconheçam como usuários da forma linguística *tu com concordância*. No Maranhão o uso dessa variante está presente na comunidade de fala dos falantes da capital, como apontam os estudos, por exemplo, de Alves (2010; 2015) e Miranda (2014).

A segunda parte do questionário que cada um dos ouvintes respondeu dizia respeito a informações qualitativas sobre a escolaridade que eles acham que os falantes têm (se fundamental, médio ou superior), sobre a faixa etária a que acham que os falantes pertencem (de 20 a 30 anos; de 30 a 40 anos, ou de 40 anos ou mais), sobre a classe econômica à qual eles pensam que o falante pertence (se baixa, média ou se alta). Além dessas informações, pediu-se aos participantes que respondessem se o áudio que estivessem ouvindo era parecido com a sua própria fala, pois, com base nessa questão, buscamos comparar os resultados aqui alcançados àqueles alcançados por pesquisas de produção. A figura 12 representa a segunda parte do questionário está apresentada:

Figura 12 - Segunda Parte do Formulário de Percepção

Você acha que essa pessoa pertence a qual escolaridade? * ☐ Ensino Fundamental ☐ Ensino Médio ☐ Ensino Superior
Você acha que essa pessoa pertence a qual faixa etária? * ☐ 20 a 30 anos ☐ 30 a 40 anos ☐ 40 anos ou mais
Você acha que essa pessoa pertence a qual classe social? * ☐ Classe alta ☐ Classe média ☐ Classe baixa
Você acha que essa pessoa é de qual lugar do Brasil? * ☐ Centro-Oeste (Goiás, Mato Grosso e Mato Grosso do Sul) ☐ Norte (Acre, Amapá, Amazonas, Pará, Rondônia, Roraima e Tocantins) ☐ Nordeste (Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe) ☐ Sudeste (Espírito Santo, Minas Gerais, Rio de Janeiro e São Paulo) ☐ Sul (Paraná, Rio Grande do Sul e Santa Catarina)

Você falaria igual essa pessoa? *

☐ Sim

☐ Não

Fonte: Elaboração própria.

Por fim, a aplicação do questionário se encerrava com os ouvintes cedendo informações sociodemográficas acerca de si próprios, tais como sua idade, seu sexo, sua escolaridade, a cidade em que mora e local de nascimento, nos casos em que more em uma cidade diferente da que nasceu.

Figura 13 - Terceira Parte do Formulário de Percepção

Quantos anos você tem? *

Texto de resposta curta
.....

Qual o seu sexo ? *

☐ Masculino

☐ Feminino

Qual a sua escolaridade? *

☐ Ensino Fundamental

☐ Ensino Médio

☐ Ensino Superior

Você é de Caxias/MA? *

☐ Não

☐ Sim

Fonte: Elaboração própria.

3.4 Coleta e perfil dos ouvintes

No total, 60 ouvintes residentes em Caxias responderam aos formulários *online* da plataforma de pesquisa *Google Forms*. As redes sociais utilizadas para a divulgação do formulário e recrutamento dos ouvintes foram o *WhatsApp*, *Instagram* e *Facebook*, com a ajuda de voluntários, colegas da autora da pesquisa, que cordialmente compartilharam o *link* com pessoas próximas.

O formulário do *Google Forms* dividia-se em 3 seções: na primeira seção, havia um texto com algumas orientações necessárias para que os ouvintes respondessem ao formulário; na segunda, disponibilizaram-se os clipes com as vozes dos falantes produzindo as variantes cujo significado social se pretende acessar, por meio dos diferenciais semânticos. Ressalte-se que os falantes não podiam pular as etapas da pesquisa, uma vez que era obrigatório que os falantes respondessem às questões de cada uma das etapas.

A respeito da plataforma *Google Forms*, informa-se que ela oferece duas possibilidades para a adição dos áudios no formulário. A primeira possibilidade é compartilhando o *link* do áudio no formulário, para que o participante clicasse e fosse direcionado a uma nova página, quando, somente aí, os áudios poderiam ser reproduzidos. A segunda possibilidade é adicionar os áudios no formulário em formato de vídeo exportado do *youtube*, o que permitia que o ouvinte ouvisse os áudios quantas vezes quisesse, sem sair da mesma página em que estivesse respondendo aos questionários de percepção.

Optou-se pela segunda possibilidade de elaboração de questionário, pois entendeu-se que sair da página do próprio questionário para se ouvir os áudios podia não apenas dispersar os participantes na tarefa da resposta, como deixar o exercício de responder ao que se pedia mais cansativo.

De um modo geral, alguns participantes relataram ser bem extenso o questionário, mas isso não os impossibilitou de responderem ao que o questionário lhes pedia. Por outro lado, outros participantes se engajavam na tarefa, chegando a afirmar que nunca tinham parado para observar a fala do outro, bem como a sua fala. Outros, ainda, apontaram que não tiveram dificuldades em julgar determinadas formas linguísticas e possíveis significados sociais potencialmente associados a essas formas linguísticas em tela, como, por exemplo, a noção de caxienseidade, já que, de acordo com alguns participantes, era fácil de identificar a variável presente na fala dos caxienses.

A tabela 11 mostra a distribuição dos 60 participantes de acordo com suas características sociais:

Tabela 10 – Números de ouvintes de Caxias e de outra localidade que participaram do experimento.

	CAXIAS	NÃO CAXIENSES
FEMININO	29	5
MASCULINO	23	3
ENSINO FUNDAMENTAL	7	1
ENSINO MÉDIO	19	2
ENSINO SUPERIOR	26	5
TOTAL	52	8

Fonte: Elaboração própria.

O formulário de percepção foi preenchido por 60 ouvintes, sendo 52 caxienses, os respondentes, e 8 não caxienses. Como cada um dos 60 ouvintes respondeu a 4 questionários, obteve-se um total de 240 respostas dadas pelos ouvintes, foram realizadas as análises estatísticas na plataforma R (R Core Team, 2023), a fim de verificar quais possíveis significados sociais carregam as formas linguísticas de segunda pessoa do singular, bem como também verificar se as hipóteses levantadas neste capítulo são consolidadas. No próximo capítulo serão apresentados os resultados das análises de regressão.

CAPÍTULO 4

4 DESCRIÇÃO E ANÁLISE DOS DADOS

Este capítulo vai apresentar os resultados de percepção dos ouvintes caxienses em relação às formas pronominais de segunda pessoa do singular, com base na adaptação da técnica dos estímulos pareados (Lambert et al, 1960; Campbell-Kibler, 2006; Oushiro, 2015; Soriano, 2016; Canever, 2017; Mendes, 2017; Santos, 2020). Antes de expor os resultados, apresentam-se as hipóteses deste estudo: (i) o uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu foste* e *tu gostaste*, não será percebido pelos ouvintes caxienses como uma variante da sua localidade; em (ii) o uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu vais* e *tu és*, indicia maior escolarização; e (iii) os usos das formas pronominais de segunda pessoa do singular, *tu com concordância* e *você*, indiciam maior grau de formalidade para os ouvintes caxienses, ao passo que a forma *tu sem concordância* indicia menor grau de formalidade, revelando, assim, maior proximidade entre os interlocutores.

Como se viu no capítulo anterior, o questionário foi respondido por 60 ouvintes, que preencheram um formulário para cada um dos áudios produzidos por cada um dos quatro falantes, obtendo-se, assim, um total de 240 respostas. A análise estatística dos dados obtidos a partir dessas respostas foi feita com o auxílio do programa R (R Core Team, 2023).

Inicialmente foram feitas análises da dispersão geral das respostas por escala para facilitar a compreensão dos resultados da análise de componentes principais e entender a diferença nos julgamentos.

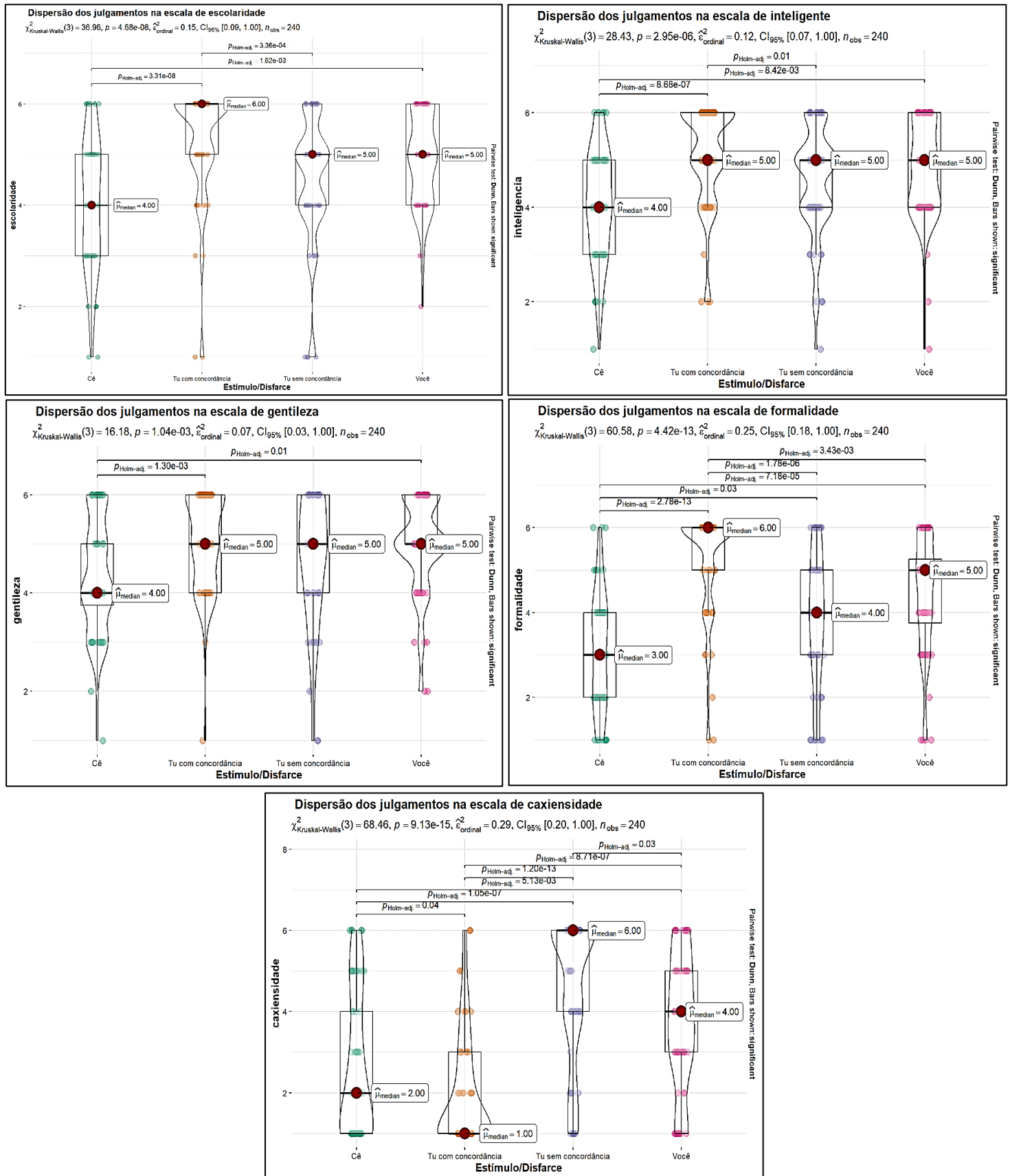
Para tanto, trabalhou-se com o teste estatístico não paramétrico *Kruskal Wallis*. Esse teste foi proposto por Kruskal e Wallis (1952), que é a versão não paramétrica do teste *Wilcoxon*, em razão de termos acima de três ou mais grupos de amostras independentes, que foram comparadas com base na média dos pontos.

O teste baseia-se no método não paramétrico de comparação de médias e variância de duas amostras pareadas, a fim de contrapor as notas das amostras pareadas. Além disso, esse teste mostra a significância das diferenças por base no valor de p . O valor de p serve para considerar as diferenças significativas entre as variáveis (escolaridade, inteligência, gentileza, formalidade e caxienseidade), quando o valor de p é igual ou menor que 0,05, e não significativas quando o valor de p é maior do que 0,05. Observa-se nos gráficos (figura 14) que os resultados das escalas indicam um $p < 0.05$. Assim, embora seja possível atestar que há diferença

significativa na comparação entre as amostras, é válido lembrar que significância não é sinônimo de importância e correlação não é sinônimo de motivação (Oushiro, 2015).

Outra decisão tomada em relação às análises diz respeito ao fato de que os resultados serão apresentados em gráficos de violino, em razão de trazerem informações mais detalhadas acerca dos resultados estatísticos. Além disso, conforme Sene (2021; 2022), esse tipo de gráfico reporta o tamanho do efeito, que é o resultado do E^2 ordinal, ou seja, o tamanho do efeito mede a magnitude e a força da variável independente sobre a dependente, como, por exemplo, o tamanho do efeito de soar escolarizado para explicar o julgamento social atribuído as quatro variantes. Segundo Sene (2021; 2022), o tamanho do efeito pode ser seguido com base no estudo de Cohen (1988) cuja referência é classificada como (0.01 a 0.06: pequeno; 0.06 a 0.14: moderado; ≥ 0.14 : grande).

Figura 14 – Dispersão Geral dos dados



Fonte: Elaboração própria.

Na figura 14, é possível observar que nas escalas de escolaridade, inteligência e gentileza, diante da dispersão das respostas dos atribuídas pelos respondentes no questionário, nota-se que não há uma diferença estatisticamente significativa, portanto, os resultados parecem revelar que, para os ouvintes caxienses, os falantes foram percebidos como menos escolarizados, menos inteligentes e menos gentis quando foram ouvidos, produzindo os estímulos com a variante *cê*. A dispersão das respostas dos ouvintes para os estímulos em que os falantes produziam a variante *tu com concordância* parece revelar maior escolaridade, maior inteligência, maior gentileza, maior formalidade e menor caxienseidade.

Do mesmo modo, os falantes foram percebidos como mais formais quando ouvidos produzindo os estímulos com a variante *tu com concordância*. Apenas quando avaliados realizando os trechos em que aparecem a variante *cê* é que os quatro foram percebidos como menos formais.

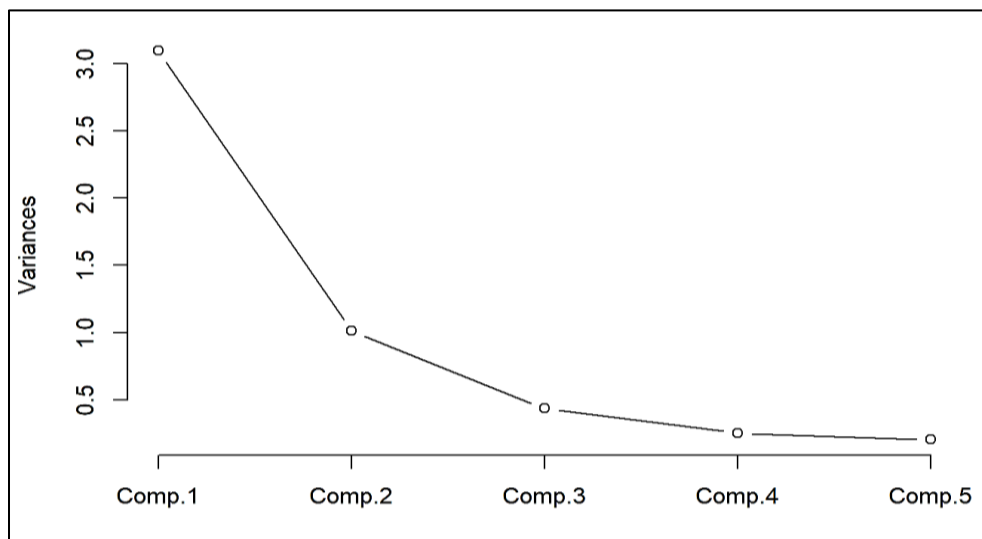
Considerando o que foi discutido, e levando em consideração os resultados dos estudos de produção realizados no Maranhão, pode-se inferir que há um gradiente de formalidade em relação às formas linguísticas de segunda pessoa do singular na variedade maranhense, em que formas *tu com concordância* e *você* estão em extremidades opostas desse gradiente de formalidade, indicializando maior formalidade nas relações comunicativas dos falantes, ao passo que a forma *tu sem concordância* estaria em seu extremo oposto, indicializando maior proximidade entre os interlocutores, confirmando, assim, a proposta de Alves (2015). Contudo, o nosso estudo de percepção vem nos apresentando que esse gradiente de formalidade não estaria atuando dessa mesma forma para os falantes caxienses, visto que os resultados alcançados nesta pesquisa nos revelam que as formas *tu com concordância*, *tu sem com concordância* e *você* estão nas mesmas extremidades do gradiente de formalidade, ou seja, para os ouvintes caxiense o uso dessas formas linguísticas dentro das relações comunicativas indicia maior formalidade.

Por fim, a escala caxienseidade parece revelar que os falantes foram percebidos como mais caxienses quando estavam usando as formas *tu sem concordância* e *você* comparativamente à quando foram avaliados usando as formas *tu com concordância* e *cê*, casos em que foram percebidos como menos caxienses.

Em seguida, realizaram-se as Análises de Componentes Principais (ACP), cuja função é para ver se há interação entre as escalas (Levshna, 2015). A análise de ACP é recomendada para os casos em que se objetiva determinar o número mínimo de fatores que respondem pela máxima variância nos dados, sendo os fatores chamados componentes principais (Malhotra, 2012).

A ACP determina quais componentes explicam a variância das respostas dadas pelos ouvintes, o eixo x corresponde ao número de componentes extraídos e o eixo y que corresponde aos *eigenvalues*, aplicando o teste de Cattell (Cattell, 1966 *apud* Santos, 2020). O intuito é encontrar o valor que corresponde ao ponto em que a variância apresenta uma inclinação linear (denominado de *ponto do cotovelo*):

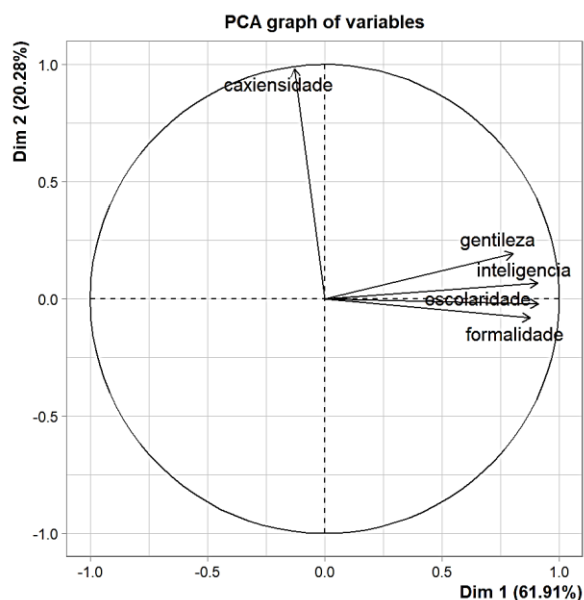
Figura 15 – *Screeplot* da ACP das respostas dadas por caxienses às escalas do formulário de percepção



Fonte: Elaboração própria.

No *screeplot* (figura 15), pode-se identificar o ponto do cotovelo das dimensões observadas nas respostas dadas pelos ouvintes caxienses. A partir do ponto 2, fica visível o ponto mais baixo do *eigenvalues*. Os CP1 e CP2 com os valores (3,09 e 1,01), respectivamente, foram os componentes considerados para a amostra de dados aqui analisada.

Na sequência, apresenta-se o gráfico em relação aos cinco CPs analisados, bem como as direções da correlação de cada uma.

Figura 16 – Gráfico da direção da correlação

Fonte: Elaboração própria.

O Coeficiente de correlação de Pearson (r) é uma medida adimensional que pode assumir valores no intervalo entre -1 e +1. O coeficiente mede a intensidade e a direção de relação lineares. A direção diz respeito ao tipo de correlação (positiva ou direta e negativa e inversa) entre as escalas. Já a intensidade diz respeito ao grau de relacionamento entre duas variáveis. Na figura 16, observa-se o direcionamento das escalas e suas correlações, as variáveis de escolaridade, inteligência, gentileza e formalidade apresentam um padrão similar na sua direcionalidade. Em outras palavras, à medida em que se observa, que certos estímulos indicializam maior escolarização, também parecem indicializar maior grau de inteligência, gentileza, formalidade. Essas escalas parecem interagir com a percepção de menor relação com a escala de caxiensiidade, ou seja, os falantes são percebidos como menos caxienses quando avaliados como mais formais, inteligentes, mais escolarizados e mais gentis.

Tabela 11 – Componentes Principais gerados pelas correlações entre as respostas nas 5 escalas para ouvintes caxienses – com rotação *Promax*¹⁶

	Componentes	
	(Prestígio)	(Caxiensiidade)
	CP1	CP2
Escolaridade	0,91	- 0,02
Inteligência	0,91	0,06
Gentileza	0,80	0,19
Formalidade	0,87	- 0,08
Caxiensiidade	- 0,12	0,98
<i>Eigenvalues</i>	3, 09	1,01
% Variância	61,8	20,3
% Acumulativa	61,8	82,2

Fonte: Elaboração própria.

Na tabela 12, encontram-se destacadas as escalas com os scores que apresentam valores mais elevados e que interagem diretamente com o componente principal. Seguem-se as decisões de Soriano (2016), Canever (2017), Mendes (2017), Sene (2019), Santos (2020) e Barcelos (2020), que determinaram os scores acima de 0,70 como média para composição dos CPs de suas respectivas pesquisas.

O primeiro dos componentes principais foi denominado prestígio, por revelar que as escalas para *escolaridade* (índice = 0,91), *inteligência* (índice = 0,91), *gentileza* (índice = 0,80), *formalidade* (índice = 0,87) interagem entre si, e que, juntas, explicam 61,8% da variância para as respostas.

Mantém-se a denominação *caxiensiidade* para se referir ao CP2, já que não houve interação dessa escala com nenhuma outra. A variância nas respostas dadas pelos ouvintes para esse CP é de 20,3, com (índice 0,98).

Como os resultados das Análises dos Componentes Principais por si só não permitem identificar um padrão de correlação entre as respostas e as escalas que apresentaram os scores mais altos, é necessário que sejam feitas análises de regressão linear para os dois CPs e para as variáveis qualitativas.

¹⁶ Promax é um método de rotação oblíquo do qual o pressuposto de independência entre os fatores é retirado, permitindo que os mesmos rodem livremente de maneira a simplificar-se a sua interpretação (Hair et al., 2006).

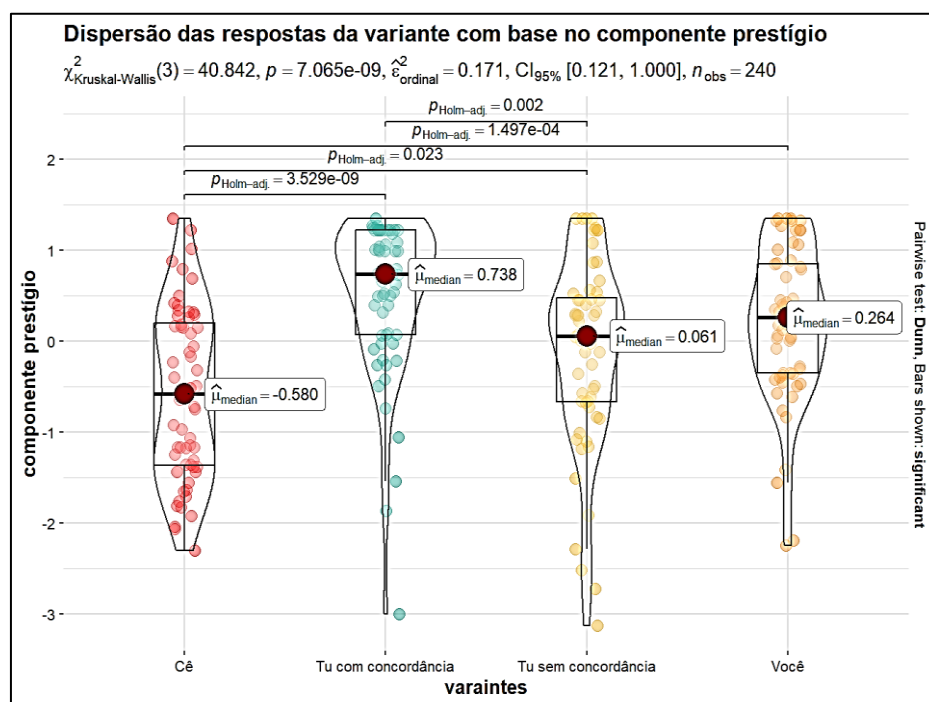
4.1 CP1- Prestígio

O primeiro componente principal é composto pelas escalas de escolaridade, inteligência, gentileza e formalidade. No início da subseção anterior, de maneira global, as dispersões das respostas dos ouvintes para todas as escalas revelam que há diferenças nas respostas dadas pelos ouvintes.

Reporta-se o entendimento de que, a partir da análise de componentes principais, que busca correlações entre as escalas inicialmente consideradas para a análise, são os *scores* gerados para cada um dos CPs (aqui, prestígio e caxienseidade) que passa a ser levado em consideração para a análise estatística, uma vez que eles são as variáveis em foco, a partir daqui. Por essa razão é que se vai considerar a variação dos *scores* de -3 a +2, e não as escalas de 1 a 6 dos questionários respondidos pelos falantes.

Note-se que os ouvintes caxienses perceberam os falantes como mais prestigiosos quando estes foram ouvidos realizando a variante *tu com concordância*. Como as diferenças nas dispersões das respostas são pequenas em relação as variáveis, parece-nos afirmar que para os ouvintes caxienses a forma *tu com concordância* indicializa maior prestígio, ainda que se possa considerar que, para os ouvintes caxienses, a variante *cê* parece ser a forma que indicializa menor prestígio social entre os falantes:

Figura 17 – Dispersões das respostas dos ouvintes caxienses no CP Prestígio, por Variante

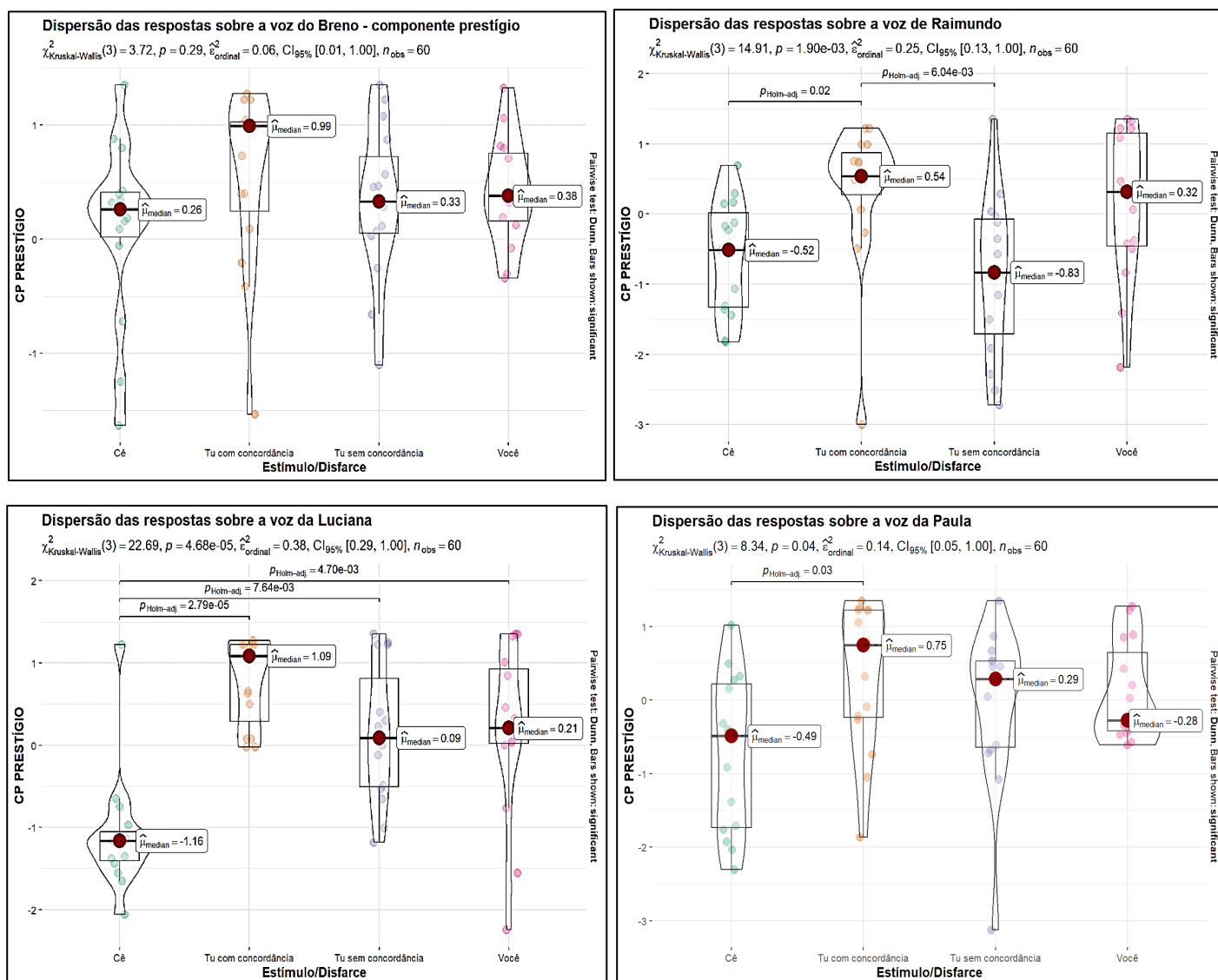


Fonte: Elaboração própria

A partir daí, interessa verificar se a diferença de percepção das respostas dos ouvintes é estatisticamente significativa. Antes, porém, observa-se, com base na figura 18, que a distribuição entre as respostas dadas pelos ouvintes parece revelar pouca diferença entre os falantes. De um modo geral, os gráficos mostram que os falantes Raimundo, Paula e Luciana, em geral foram percebidos como menos prestigiosos nos disfarces em que realizam variante *cê*, comparativamente ao falante Breno, que parece não revelar diferenças nas avaliações dos ouvintes para nenhuma das variantes.

Os dados de distribuição mostram também que todos os falantes foram percebidos como mais prestigiosos quando ouvidos com os disfarces com *tu*, *seguido da concordância*.

Figura 18 – Dispersão das respostas dos participantes por falantes de acordo com os estímulos ouvidos



Fonte: Elaboração própria

Observando o gráfico, pode-se inferir que as respostas dos ouvintes tendem a revelar e confirmar duas das hipóteses levantadas neste estudo: referente a nossa segunda hipótese, de que o uso do pronome *tu*, seguido do verbo com marca de concordância, a exemplo de *tu vais* e *tu és*, indicializaria maior escolarização, maior inteligência, maior grau de gentileza e maior formalidade. A terceira hipótese, a de que os usos das formas pronominais *tu* com concordância e *você* de segunda pessoa do singular indiciariam graus de formalidade para os ouvintes caxienses, é confirmada. Em outras palavras, pode afirmar que os ouvintes perceberam os falantes como mais formais nos estímulos realizados com a variante *tu* com concordância, *você* e *tu* sem concordância. Isso nos permite levar à observação de que soar informal para os caxienses é realizar a forma *cê*.

Esse resultado se torna interessante para o nosso estudo na medida em que se alinha à pesquisa de produção acerca da variável de segunda pessoa do singular no Maranhão (Alves, 2015), que tem apontado que as formas tendem a revelar maior ou menor grau de formalidade entre os interlocutores, a depender de qual variante mobilizem no ato da conversação. Dizendo de outro modo, para Alves (2015, p. 139), “[...] é possível afirmar que há um valor social dado à forma *tu* tendo em vista sua regularidade em ambientes informais (78%) e formais (58%) [...] O *tu* com concordância é favorecido em contextos mais formais (21,4%) [...] O *você* encaminharia nessa mesma direção (20,1%), ao apontar uso mais restritos a contextos de maior formalidade [...]”.

Análises estatísticas de regressão linear não dão conta de confirmar tal proposição elaborada por Alves (2015), dado que o desenho do experimento que aqui se projetou não permite avaliar tal afirmativa. No entanto, é possível aferir o efeito do sinal linguístico aqui analisado sobre a percepção de quão prestigiosos soam os falantes, ao serem ouvidos com cada umas das quatro maneiras de se referir à segunda pessoa do discurso, no singular.

As análises de regressão linear tomam como *intercept* da análise o falante Breno, a forma linguística *cê* e a interação entre o falante Breno, realizando a forma *cê*.

Tabela 12 – Resultado do Modelo de Regressão de Efeitos Mistos para o CP Prestígio

Efeitos Fixos	Estimativa	Erro padrão	Valor – <i>t</i>	<i>P</i>
(<i>Intercept</i>)	0,132	0,225	0,584	0,559
Luciana	-1,213	0,321	-3,778	< 0, 002***
Paula	-0,906	0,311	-2,914	< 0, 003**
Raimundo	-0,733	0,321	-2,280	< 0, 023*
Tu com concordância	0,371	0,311	1,195	0,233
Tu sem concordância	0,190	0,321	0,592	0,554
Você	0,196	0,320	0,614	0,540
Luciana: Tu com concordância	1,480	0,448	3,300	< 0, 001**
Paula: Tu com concordância	0,763	0,485	1,575	0,117
Raimundo: Tu com concordância	0,574	0,448	1,281	0,201
Luciana: Tu sem concordância	0,987	0,460	2,143	< 0, 033*
Paula: Tu sem concordância	0,544	0,448	1,213	0,226
Raimundo: Tu sem concordância	-0,436	0,512	-0,851	0,396
Luciana: Você	1,086	0,509	2,133	< 0,0351*
Paula: Você	0,674	0,448	1,504	0, 134
Raimundo: Você	0,473	0,460	1, 030	0, 304

Fórmula: CP (Prestígio) ~ falante * variante + (1|ouvinte), Data: PCA.S

Intercept: Breno; Cê

Fonte: Elaboração própria

Os valores da estimativa na tabela 13 devem ser lidos com relação ao *intercept*, ou nível de referência. Assim, na primeira linha da estimativa, 0,132 é o valor estimado pelo modelo de análise adotado para a análise do CP Prestígio, para Breno, em seu disfarce com *cê*. Em seguida, na mesma coluna, o valor -1,213 deve ser somado ao valor 0,132 do *intercept*, a fim de que se obtenha a estimativa para Luciana. O valor *p*, 0,559, indica as chances de se obter resultados caso a hipótese nula fosse verdadeira, ou seja, no caso em que não haja diferenças estatisticamente significativas entre Luciana e Breno.

Assim, observa-se que, comparativamente as falantes Luciana e Paula apresentam diferenças estatísticas relevantes quando comparadas ao falante Breno, soando menos prestigiosas, de acordo com os ouvintes. Do mesmo modo, há diferenças significativas entre os falantes homens, o falante Raimundo foi percebido como menos prestígio quando comparado a Breno.

Relembre-se o fato de que esse CP é composto pelas escalas iniciais de escolaridade, inteligência, gentileza e formalidade, o que importa informar que os três falantes em questão soaram menos escolarizados, menos inteligentes, menos gentis e menos formais comparativamente ao falante Breno.

Pode-se interpretar, de acordo com o modelo de regressão (tabela 13), que, para os ouvintes caxienses, Luciana e Paula foram percebidas como tendo menos prestígio, comparativamente a Breno ($p < 0,002***$ e $p < 0,003**$, respectivamente). O falante Raimundo também apresentou diferenças significativas, quando comparado ao falante Breno, sendo percebido como menos prestígio.

Já em relação à forma linguística, parece não haver diferenças significativas entre as formas linguísticas, quando analisadas fora da interação com o falante. Nesses casos, em que se analisa a interação falante e forma linguística, observa-se que Luciana, em sua versão com *tu com concordância* ($p < 0,001**$), foi percebida como soando mais prestigiosa, comparativamente a Breno, em seu disfarce com *cê*, resultado que se assemelha aquele alcançado para Luciana, em seu disfarce *tu sem concordância* ($p < 0,033*$). Tal diferença significativa se impõe quando essa mesma falante é analisada em seu disfarce com a forma linguística *ocê*, em que ela parece indiciar maior prestígio quando comparada ao falante Breno, em seu disfarce com a forma *cê* ($p < 0,0351*$).

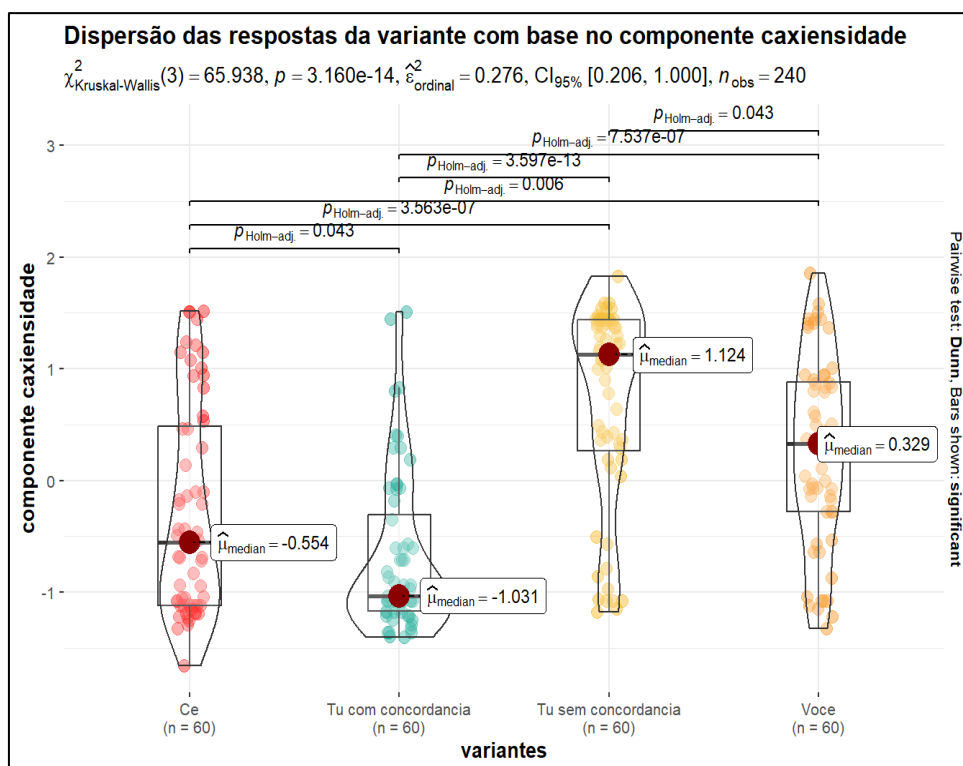
Não foram estabelecidas diferenças significativas entre os falantes Raimundo, comparativamente a Breno, e nem entre a falante Paula, em sua interação com o pronome *ocê*, e Breno, em seu estímulo com a forma *cê*.

De modo geral, a interação entre os falantes e as formas linguísticas revelam diferenças significativas, no sentido de que as mulheres foram percebidas como menos prestigiosas, o que, de certa maneira, nos distancia dos resultados obtidos pelos estudos de produção com maranhenses, em que as mulheres tendem a usar formas mais prestigiosas nas comunidades de falas (Alves, 2015).

4.2 CP2 - Caxienseidade

O segundo componente principal, caxienseidade, diferentemente do primeiro, é composto unicamente pela escala de mesmo nome. As dispersões nas respostas dadas pelos ouvintes ao CP2 são apresentadas no gráfico a seguir.

Figura 19 – Dispersões das respostas dos ouvintes caxienses no CP Caxienseidade, por Variante



Fonte: Elaboração própria

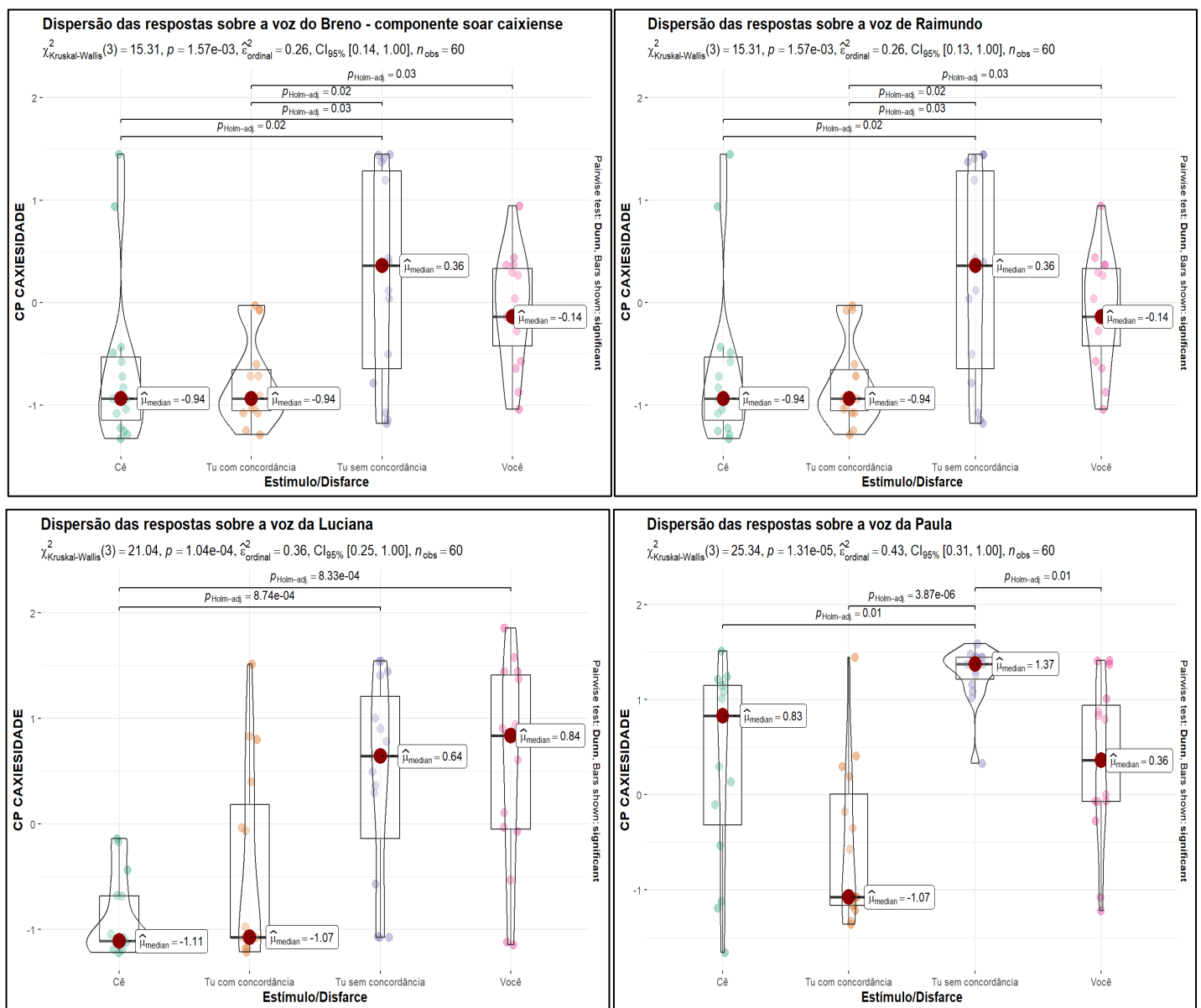
Note-se que os ouvintes caxienses perceberam os falantes (analisados conjuntamente) como mais caxienses quando ouviram os estímulos produzidos nos disfarces com *tu sem concordância* e *você*, ao passo que foram percebidos como menos caxienses quando os estímulos estavam sendo realizados nos disfarces com *tu com concordância* e *cê*. Esses resultados respondem a hipótese que para os ouvintes caxienses o uso do pronome *tu* seguido do verbo com marca de concordância, a exemplo de *tu foste* e *tu gostaste*, não seria percebido pelos ouvintes como uma variante da sua localidade. Esperava-se que os estímulos *tu sem concordância* e *você* fossem percebidos como formas linguísticas mais características dos falantes caxienses.

Miranda (2014), em estudo de produção sobre as formas pronominais *tu* e *você*, revelou que a variante *tu com concordância* é realizada em 4,5% de toda a amostra analisada. Nessa direção, observa-se que os resultados obtidos nesta pesquisa se coadunam à pesquisa de produção, realizada na mesma comunidade, de modo que, além de os falantes caxienses não produzirem a forma *tu com concordância*, esses falantes não se reconhecem/percebem como usuários dessa variante linguística.

Para a escala caxienseidade interessa saber se a diferença nas respostas dos falantes é estatisticamente significativa.

Observa-se que todos os falantes foram percebidos soando mais caxienses, quando foram ouvidos realizando os disfarces com os estímulos com *tu sem concordância*, com especial destaque para a falante Paula, que foi percebida como mais caxiense quando comparada aos demais falantes. Em contrapartida, os ouvintes parecem não relacionar a variante *cê* à sua variedade linguística, já que todos os falantes foram percebidos como menos caxienses quando avaliados produzindo os estímulos com esse sinal linguístico. No caso dos falantes do sexo masculino, Breno é quem parece soar menos caxiense quando ouvido na versão *com você*.

Figura 20 - Dispersão das respostas dos participantes por falantes de acordo com os estímulos ouvidos



Fonte: Elaboração própria

As diferenças entre os falantes Breno, Luciana, Paula e Raimundo podem ser observadas para além dos gráficos, mais especificamente por meio da análise de regressão linear multivariada realizada.

Tabela 13 – Resultado do Modelo de Regressão de Efeitos Mistos para o CP Caxienseidade

Efeitos Fixos	Estimativa	Erro padrão	Valor – <i>t</i>	<i>P</i>
(<i>Intercept</i>)	- 0,647	0,205	- 3,159	<0, 001**
Falante Luciana	-0,251	0,290	-0,867	0,386
Falante Paula	0,977	0,288	3,392	<0,001***
Falante Raimundo	0,711	0,290	2,452	0,015
Tu com concordância	-0,138	0,288	-0,481	0,631
Tu sem concordância	0,883	0,290	3,043	<0,002**
Você	0,568	0,290	1,959	0,051
Luciana: Tu com concordância	0,606	0,409	1,483	0,139
Paula: Tu com concordância	-0,735	0,421	-1,744	0,082
Raimundo: Tu com concordância	-0,903	0,409	-2,209	<0, 028*
Luciana: Tu sem concordância	0,459	0,411	1,105	0,270
Paula: Tu sem concordância	0, 055	0,409	0,136	0,892
Raimundo: Tu sem concordância	-0, 075	0,427	-0,177	0,860
Luciana: Você	0,879	0,426	2,062	<0,040*
Paula: Você	-0,545	0,409	-1,334	0,183
Raimundo: Você	-0,405	0,411	-0,985	0,325

Fórmula: CP (Caxienseidade) ~ falante * variante + (1| ouvinte), Data: PCA.S

Intercept: Breno; Cê

Fonte: Elaboração própria

Esse modelo de regressão apresenta os resultados das análises estatísticas para o componente principal caxienseidade, em função das variáveis previsoras, bem como da interação entre elas.

Para esse CP, observa-se que, comparativamente ao *intercept* Breno, há diferenças significativas apenas com a falante Paula, que tende a ser percebida como menos caxiense ($p < 0,001***$) do que o falante homem. Esse resultado importa porque, quando se observam os resultados para o CP prestígio, todos os falantes (Luciana, Paula e Raimundo) se diferenciaram significativamente de Breno, resultado que não se repete aqui.

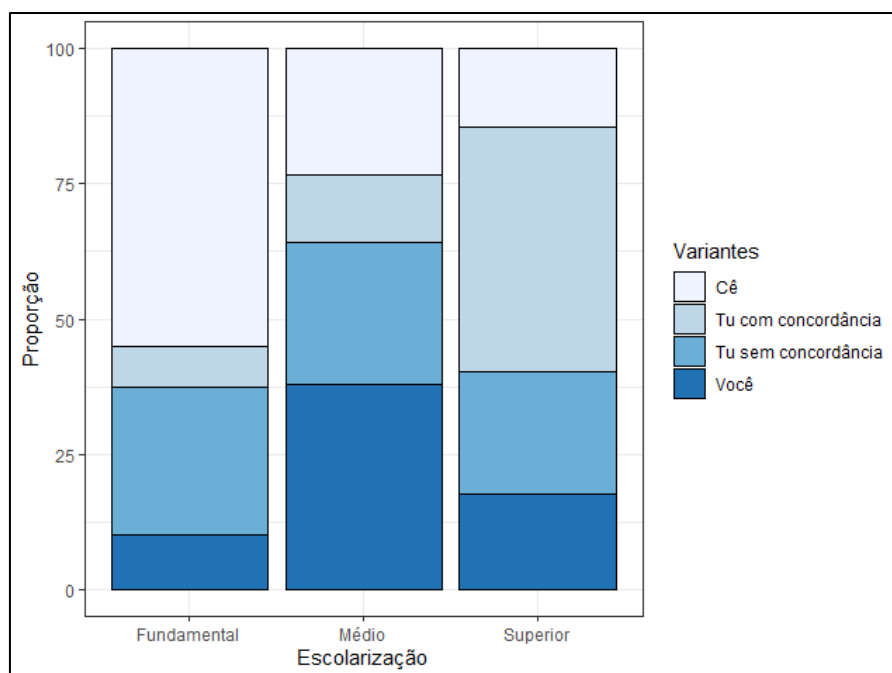
Quando analisadas individualmente, apenas a forma *tu sem concordância* mostra diferença significativa em relação à forma *cê*, *intercept* do modelo de análise. Nesse caso, é a forma *tu sem concordância* quem melhor representa o falar caxiense ($p < 0,002**$) em relação a variante *cê*.

Vão ser observadas diferenças significativas mais evidentes entre os falantes quando se analisam o falante em interação com as formas linguísticas: aqui, observa-se que há diferença significativa entre os falantes homens, com Raimundo, em sua versão *tu com concordância*, soando como alguém mais caxiense do que Breno, em sua versão com o mesmo estímulo. Também se observa diferença entre a falante Luciana, em sua versão com o estímulo *você*, comparativamente a Breno, que soa menos caxiense que Paula ($p < 0,040^*$). Nas análises com interação não se observam diferenças entre Paula e Breno.

4.3 Variáveis qualitativas

Os resultados apresentados nesta subseção levaram em conta as variáveis qualitativas escolarização, faixa etária, classe social e origem dos falantes. Essas respostas foram colhidas com base nas perguntas que fizeram parte da segunda seção aplicadas no questionário, logo após os ouvintes responderem às perguntas em relação as escalas. Para análise dessas variáveis, utilizou-se, no R, o teste de qui-quadrado, a fim de testar a significância observada entre duas variáveis (nominais/qualitativas), a variável dependente (escolarização, faixa etária, classe social e origem dos falantes) e a variável independente (a variável linguística). Os resultados para a variável escolarização pode ser visualizados na Figura 21.

Figura 21 - Distribuição das respostas para a variável Escolarização, por variante, pelos ouvintes caxienses



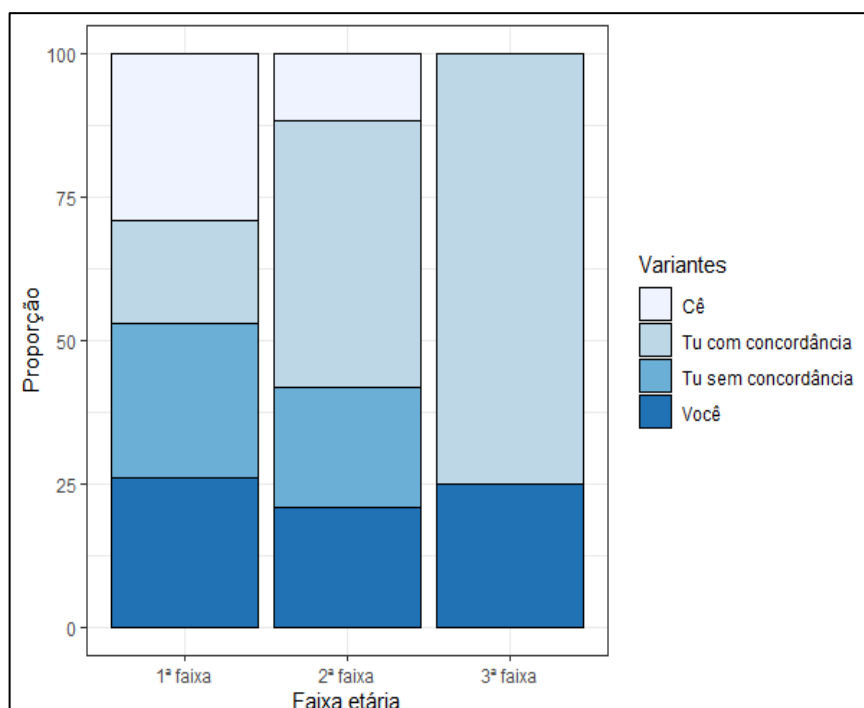
X-squared = 59.104, p-value = < 0,001

Fonte: Elaboração própria

Observa-se que, quando os estímulos foram realizados com a variante *cê* e com a variante *tu sem concordância*, os ouvintes caxienses associaram essas formas à fala de pessoas que aparentam ter uma baixa escolarização, ao passo que os mesmos ouvintes associaram os estímulos realizados com a variante *tu com concordância* à fala de pessoas que aparentam ter um nível de escolarização mais alto. Além disso, os ouvintes caxienses associaram os estímulos com a variante *você* à fala de pessoas que aparentam ter escolarização média.

Os resultados para o estudo de percepção aqui obtidos não se alinham, aos resultados obtidos por Miranda (2014, p. 114), que analisou dados de produção na fala de caxienses. Em seu estudo, o autor mostrou que falantes de nível médio favorecem a regra da concordância de segunda pessoa do singular em Caxias, que é *tu com concordância*, ao passo que os falantes de nível superior não favorecem a aplicação da regra, ainda que o autor esperasse “que falantes de nível superior com mais anos de escolarização realizassem a aplicação”, observa-se que os ouvintes caxienses tendem a associar o uso dessa variante à fala de pessoas com nível de escolarização alta, mas ressaltasse que “ter mais anos de escolarização nem sempre é garantia de realizar concordância verbal”.

Figura 22 - Distribuição das respostas para variável Faixa Etária, por variante, pelos ouvintes caxienses



X-squared = 28.954, p-value = < 0,001**

Fonte: Elaboração própria

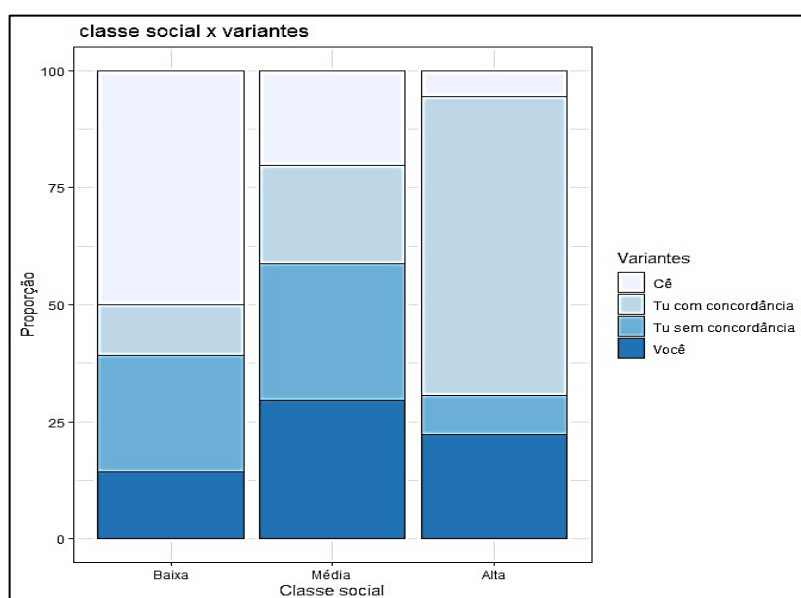
Para a variável Faixa Etária, observa-se que as diferenças significativas, estabelecidas pelo teste de qui-quadrado, mostram que, enquanto a forma *cê* é associada à fala de indivíduos

mais jovens, com especial destaque para os de primeira faixa etária. A forma *cê* é associada à fala de pessoas da segunda faixa etária. A forma *tu sem concordância* também não é associada à fala de pessoas de terceira faixa etária, estando essa associação à fala de pessoas de segunda e de primeira faixa etária.

Por outro lado, os ouvintes tendem a associar o uso da forma *tu com concordância* à fala de pessoas com mais idade, ao passo que vai diminuindo para os ouvintes que associam tal uso a falantes da segunda faixa etária e a falantes de primeira faixa etária. Esses resultados parecem coadunar-se aos resultados dos estudos de produção sobre a variação *tu* e *você*, na expressão de pronome de segunda pessoa, especialmente, aqui, aqueles desenvolvidos por Alves (2010) e Miranda (2014), que mostraram que falantes de terceira faixa etária tendem a favorecer o uso da variante *tu*, seguida do verbo com marca de concordância, à medida em que a forma *você* tende a ser favorecida por falantes mais jovens, em contextos que exigem maior formalidade, já que esses falantes parecem não se diferenciar quanto ao uso do *tu* sem marca de concordância em contextos de maior informalidade e aproximação entre os falantes (Alves, 2015).

Os estímulos realizados com a forma *você*, por sua vez, parece ser igualmente associado a falantes das três faixas etárias, com uma pequena diferença para os falantes de segunda faixa etária. Essa forma foi associada a falantes de primeira faixa etária, a falantes de segunda faixa etária e a falantes de terceira faixa etária.

Figura 23 - Distribuição das respostas para variável Classe Social, por variante, pelos ouvintes caxienses

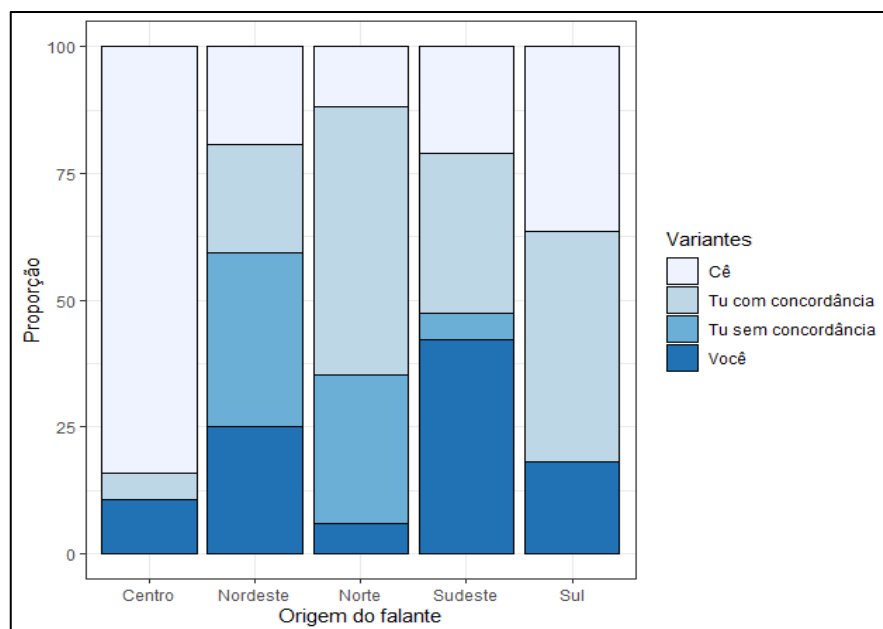


X-squared = 57.071, p-value = < 0,001

Fonte: Elaboração própria

Para a variável Classe Social, verifica-se que a diferença foi estatisticamente significativa ($X^2 = 57.071$, $p = 0.001$) quando os ouvintes avaliaram os falantes como sendo de classe social mais alta à medida em que foram ouvidos realizando os estímulos *tu com concordância*. Por outro lado, os ouvintes perceberam os falantes como sendo de classe social mais baixa quando estes foram ouvidos realizando os estímulos com a variante *cê*.

Figura 24 - Distribuição das respostas para variável Origem do Falante, por variante, pelos ouvintes caxienses



$X^2 = 69.763$, $p\text{-value} = < 0,001$

Fonte: Elaboração própria

Quando perguntados sobre de onde achavam que fossem os falantes, a depender do disfarce ouvido, os ouvintes caxienses associaram à região Centro Oeste do Brasil, ouviram as vozes produzindo os estímulos com a variante *cê* e *você*. Tal resultado coaduna-se à proposta de Scherre *et al.* (2015), que, ao organizar o sistema pronominal do português brasileiro em seis subsistemas, insere o uso de *cê* como característico do subsistema da região Central do Brasil, juntamente com a forma *você*. Esse subsistema inclui, ainda, os estados de Rondônia, Mato Grosso, Mato Grosso do Sul, Tocantins, Minas Gerais, Goiás, São Paulo e parte do Paraná.

Os ouvintes caxienses também avaliaram os falantes como sendo da região nordeste, mas não foram capazes de identificar uma forma linguística que fosse característica dessa região, os ouvintes associaram a produção do uso de *você* à região nordeste, o uso da forma *tu com concordância* e o uso da forma *cê*. A diferença relativamente maior foi para os estímulos

com *tu sem concordância* produzidos pelos falantes, percebidos como sendo dessa região pelos ouvintes.

De acordo com Scherre *et al.* (2015, p. 166), a região Nordeste

se mostra bastante diversificada, com, pelo menos, quatro subsistemas: para o Maranhão, subsistema 4 (tu/VOCÊ, tu com concordância baixa), subsistema 5 (tu/VOCÊ, tu com concordância média) e subsistema 6 (VOCÊ/tu, tu sem concordância); para a Bahia, subsistema 1 (só VOCÊ) e subsistema 6 (VOCÊ/tu, tu sem concordância); para o Piauí, Ceará, Paraíba e Pernambuco, projeção do subsistema 5 (tu/VOCÊ, tu com concordância média), sem dados à época para o desenho do Rio Grande do Norte, de Alagoas e de Sergipe.

Os ouvintes caxienses perceberam os falantes como sendo de origem da região Norte do Brasil, quando estes foram ouvidos realizando os estímulos com a variante *tu com concordância* e *tu sem concordância*. Desse modo, a percepção dos ouvintes caxienses se assemelha à classificação do subsistema proposto por Scherre *et al.* (2015). Na região Norte, os trabalhos voltados para a produção da variável de segunda pessoa do singular, como por exemplo, as pesquisas de Soares e Leal (1993) e Martins (2010), que mostram que o subsistema dessa região é diversificado, formado, basicamente, pelas variantes *tu com concordância* e *tu sem concordância*.

Os ouvintes caxienses avaliaram os falantes como sendo de origem da região Sudeste do Brasil, quando estes produziram os estímulos com a variante *você*, seguida dos estímulos com a variante *tu com concordância* e a variante *cê*. Pesquisas de produção de pronomes de segunda pessoa, no singular, realizadas na região sudeste parecem coadunar-se com os resultados aqui obtidos. As pesquisas de Modesto (2007), que mostrou que há alternância do uso da variante *tu* (32% de ocorrências) e *você* (67% de ocorrências) em Santos, cidade do litoral do Estado de São Paulo, Mota (2008), que estudou a variação dos pronomes *tu* e *você*, na cidade de São João da Ponte, em Minas Gerais, e mostrou que a variante *você* (89% de ocorrências) e a variante *tu* (10% de ocorrências), além de Santos (2012), que estudou a frequência dos pronomes sujeitos de segunda pessoa na fala carioca, constatando que a variante *tu* aparece com percentual de 12%, enquanto o pronome *você* é produzido por 49% dos participantes da pesquisa, são alguns exemplares de estudos que parecem confirmar os resultados deste estudo.

Esses estudos contribuem para a proposta por Scherre *et al.* (2015), discutindo o fato de que essa autora classifica a região Sudeste do Brasil como pertencente ao subsistema 1, formado pelas formas *você/cê/ocê* e subsistema 6, que inclui, além da forma *você*, a forma *tu*.

Os ouvintes caxienses perceberam os falantes como sendo de origem da região Sul do Brasil, quando ouvidos realizando os estímulos com a variante *tu sem concordância*, *você* e *cê*.

A pesquisa de produção de Loregian-Penkall (2004) mostrou que, ao analisar a alternância do tu/você e a concordância verbal com o pronome tu, nas localidades de Florianópolis, Santa Catarina, Porto Alegre, no Rio Grande do Sul e Ribeirão da Ilha, em Santa Catarina, o uso do tu é favorecido em relação a você, nas cidades estudadas. Scherre *et al.* (2015), classifica o subsistema da região Sul do Brasil como parte do subsistema 2, quando os falantes tendem a realizar *tu com concordância baixa*, com uso médio acima de 60%, e *tu com concordância* com índices abaixo de 10%.

No geral, os resultados das variáveis qualitativas permitem observar que a percepção dos ouvintes se aproxima dos resultados de produção da variável da segunda pessoa no singular. Em suma, quando os falantes foram ouvidos produzindo os estímulos com a variante *cê*, os ouvintes associaram esse uso à fala de pessoas com baixa escolarização, pertencentes à primeira faixa etária, e como sendo de classe social mais baixa, com origem na região Centro-Oeste do país.

Quando os falantes foram ouvidos produzindo os estímulos com a variante *você*, os ouvintes os avaliaram como pessoas de escolarização média, igualmente distribuídas entre as pessoas das três faixas etárias, além de serem mais fortemente associadas à fala de pessoas de classe média cujo nascimento teria se dado na região Sudeste do Brasil.

Quando ouvidos realizando os estímulos com a variante *tu sem concordância*, os ouvintes associaram tais formas à fala de pessoas oriundas do norte e nordeste do país, enquanto os estímulos com *tu com concordância* foram associados à fala de pessoas alto grau de escolarização, à fala de pessoas de segunda e terceira faixas etárias de origem da região Norte e Sul do Brasil.

CONSIDERAÇÕES FINAIS

Esta pesquisa de mestrado teve como objetivo verificar quais significados sociais estariam associados às formas pronominais da segunda pessoa do singular, a partir das percepções dos ouvintes caxienses. Os maranhenses são reconhecidos nacionalmente por, com base no discurso popular, “falarem o português melhor ou mais correto”. Esse discurso se prolifera devido os falantes utilizam a forma *tu com concordância*, seguindo, assim, o que prescrevem os compêndios gramaticais normativos. Porém, os estudos de variação e mudança sobre as formas de segunda pessoa do singular, realizados até o momento (Alves, 2010; 2015; Miranda, 2014), em relação à variedade de fala maranhense, tiveram as análises voltadas mais precisamente para a descrição da produção. Com isso, este trabalho mostra que é possível estudar a variação por outro viés, da percepção.

Os resultados obtidos por estudos de (Alves, 2010; 2015; Miranda, 2014) possibilitaram a desenvolver esta dissertação de mestrado. Utilizou-se o aporte teórico-metodológico de (Labov, 2006 [1966]; 2008 [1972]; Eckert, 2012), especificamente, os estudos denominados de terceira onda, ou seja, estudar o processo de variação e mudança sobre as formas de segunda pessoa do singular, mas observando-a categorias sociais que atuam no padrão linguístico e na construção do significado social. Metodologicamente, utilizou-se a técnica dos estímulos pareados (Lambert et al, 1960; Campbell-Kibler, 2006; Mendes; Oushiro, 2015; Mendes, 2016).

A nossa pesquisa veio nortear os estudos de percepção de segunda pessoa do singular na região. Apresentam-se quais significados sociais foram correlacionadas às formas linguísticas *tu com concordância*, *tu sem concordância*, *você* e *cê* com base nas percepções dos ouvintes caxienses. Esta pesquisa vem somar com o trabalho de produção desenvolvido por Miranda (2014), pois entende-se que “a associação entre variantes linguísticas e os significados sociais não decorrem diretamente das correlações verificadas nas análises de produção linguística” (Oushiro, 2015, p.325). Entretanto, essa pesquisa mostrou que muitos significados sociais foram relacionados a produção, acredita-se que essa relação deva a ser pelo status da variável na localidade.

Desse modo, retornaremos as hipóteses iniciais levantadas nesta pesquisa. Desde já, adiantaremos que todas as nossas hipóteses foram respondidas e confirmada. No tocante à primeira hipótese, achava-se que *o tu com concordância não seria associado pelos ouvintes caxienses como uma variante de sua localidade*.

Dentre os estímulos (analisados conjuntamente) aplicados no experimento, as formas linguísticas *tu* com concordância e *cê* foram percebidas como nada caxiense, ao passo que as formas linguísticas *tu sem concordância* e *você* foram percebidas como muito caxiense. Levando em consideração as análises feitas pelos estímulos (analisados individualmente), os resultados mostraram que todos os falantes foram percebidos soando mais caxienses quando foram ouvidos realizando os disfarces com os estímulos com *tu sem concordância*. Nessa direção, observa-se que os resultados obtidos nesta pesquisa se coadunam à pesquisa de produção, realizada na mesma comunidade, de modo que, além de os falantes caxienses não produzirem a forma *tu com concordância*, esses falantes não se reconhecem/percebem como usuários dessa variante linguística. É bem verdade, que fazer correlações entre usos linguísticos e categorias sociais não implicam relações diretas entre variáveis, mas o que se pode notar aqui é que a relação entre produção e percepção é uniforme (Mendes, 2018).

A segunda hipótese, acreditava-se que *os ouvintes caxienses ao ouvirem os estímulos sendo realizados na forma linguística tu, seguido do verbo com marca de concordância, indicializariam maior escolarização*. Dentre os estímulos (analisados conjuntamente), os resultados parecem revelar que, para os ouvintes caxienses, os falantes foram percebidos com maior escolaridade, maior inteligência, maior gentileza, maior formalidade e menor caxiensidade quando realizavam os estímulos com a forma *tu com concordância*. Em contrapartida, esses mesmos falantes foram percebidos como menos escolarizados, menos inteligentes e menos gentis quando foram ouvidos produzindo os estímulos com a forma *cê*.

Por fim, a terceira hipótese, *os usos das formas pronominais tu com concordância e você de segunda pessoa do singular indicariam graus de formalidade para os ouvintes caxienses*, é confirmada. Infere-se que os ouvintes perceberem os falantes como mais formais nos estímulos realizados com a variante *tu com concordância*, *você* e *tu sem concordância*, levando-se à observação de que soar informal para os caxienses é realizar a forma *cê*.

Além disso, os resultados das variáveis qualitativas permitiram observar que a percepção dos ouvintes se aproxima dos resultados de produção da variável da segunda pessoa no singular. No geral, quando os falantes foram ouvidos produzindo os estímulos com a variante *cê*, os ouvintes associaram esse uso à fala de pessoas com baixa escolarização, pertencentes à primeira faixa etária, e como sendo de classe social mais baixa, com origem na região Centro-Oeste do país.

Quando os falantes foram ouvidos produzindo os estímulos com a variante *você*, os ouvintes os avaliaram como pessoas de escolarização média, igualmente distribuídas entre as

peessoas das três faixas etárias, além de serem mais fortemente associadas à fala de pessoas de classe média cujo nascimento teria se dado na região Sudeste do Brasil.

Quando ouvidos realizando os estímulos com a variante *tu sem concordância*, os ouvintes associaram tais formas à fala de pessoas oriundas do norte e nordeste do país, enquanto os estímulos com *tu com concordância* foram associados à fala de pessoas alto grau de escolarização, à fala de pessoas de segunda e terceira faixas etárias de origem da região Norte e Sul do Brasil.

Em suma, esta dissertação veio contribuir para os estudos de percepção, abrindo caminhos no desenvolvimento de mais pesquisas de percepção em variáveis morfossintáticas, apresentando quais significados sociais correlacionados a uma variável linguística nunca antes estudada, o que a torna diferente dos estudos já desenvolvidos na variedade de fala maranhense.

REFERÊNCIA

- ALMEIDA, N. M. *Gramática metódica da Língua Portuguesa*. São Paulo: Saraiva, 1985.
- ALVES, Cibelle Corrêa Béliche. *O uso do tu e do você no português falado no Maranhão*. 2010. Dissertação (Mestrado em Linguística), Programa de Pós-Graduação em Linguística,
- ALVES, Cibelle Corrêa Béliche. *Pronomes de Segunda Pessoa no Espaço Maranhense*. 2015. Tese (Doutorado em Linguística), Universidade de Brasília, UnB, Brasília, 2015.
- ANDRADE, Carolina Queiroz. *O Tu e mais quantos? – A segunda pessoa na fala brasiliense*. 2010. Dissertação (Mestrado em Linguística). Instituto de Letras / Universidade de Brasília, 2010.
- BARCELLOS, Maria Eugênia Martins. *O falar paulist [É Ì:] no e os significados sociais de (AN): correlações entre origem do ouvinte e percepção*, 2020. Tese de Doutorado.
- BUCHOLTZ, Mary. The whiteness of nerds: Superstandard English and racial markedness. *Journal of linguistic anthropology*, v. 11, n. 1, p. 84-100, 2001.
- BUNIN BENOR, Sarah. The learned/t: phonological variation in Orthodox Jewish English. *University of Pennsylvania Working Papers in Linguistics*, v. 7, n. 3, p. 2, 2001.
- CAMPBELL-KIBLER, Kathryn. Accent, (ING), and the social logic of listener perceptions. *American speech*, v. 82, n. 1, p. 32-64, 2007.
- CAMPBELL-KIBLER, Kathryn. I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society*, v. 37, n. 5, p. 637-659, 2008.
- CAMPBELL-KIBLER, Kathryn. The nature of sociolinguistic perception. *Language Variation and Change*, v. 21, n. 1, p. 135-156, 2009.
- CAMPBELL-KIBLER, Kathryn. *Listener perceptions of sociolinguistic variables: The case of (ING)*. Tese de Doutorado. Stanford University, 2006. 282f.
- CANEVER, Fernanda. *Infinitivo flexionado em português brasileiro: Frequência e percepções sociolinguísticas*. 2017. Tese de Doutorado. Universidade de São Paulo.
- CARNEIRO, Honorina Maria Simões. *As formas de tratamento tu/você no português falado ludovicense*. 147 f. Tese (Doutorado em linguística e Língua Portuguesa), Faculdade de Ciências e Letras, UNESP, Araraquara, 2011.
- CARVALHO, Bruna Brasil Albuquerque. *O que você acha do uso de tu?": a percepção da variação dos pronomes de 2SG no dialeto carioca*. 2019. Dissertação de Mestrado. Faculdade de Letras, Universidade Federal do Rio de Janeiro.
- COUTINHO, Milson. *Caxias das aldeyas altas: subsídios para a sua história*. São Luís, Edição Prefeitura de Caxias, 1980.

DIAS, Edilene Patrícia. *O uso do tu no português brasileiro falado*. 2007. Dissertação (Mestrado em Linguística). Instituto de Letras / Universidade de Brasília.

DIAS, Edilene Patrícia. *O uso do tu no português brasileiro falado*. 2007. Dissertação (Mestrado em Linguística). Instituto de Letras / Universidade de Brasília.

DUARTE, Aspectos do sistema pronominal do português falado nas regiões Sudeste e Centro-Oeste. In: SIMPÓSIO NACIONAL DO GT de SOCIOLINGUISTICA DA ANPOLL, 1996, João Pessoa. *Anais do XI Encontro Nacional da ANPOLL*. Campinas: ANPOLL, 1997, p. 541-509.

ECKERT, P. *Linguistic variation as social practice*. Oxford: Blackwell, 2000.

ECKERT, P. Variation, convention, and social meaning. *Paper presented at the Annual Meeting of the Linguistic Society of America*. Oakland CA. Jan 7, 2005.

ECKERT, Penelope. “As três ondas do estudo da variação: a emergência do significado no estudo da variação linguística.” *Organon*, vol 37(73), 268 – 291, 2022 [2012]. Tradução de Samuel Gomes de Oliveira, Livia Majolo Rockenbach e Athany Gutierrez.

ECKERT, Penelope. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual review of Anthropology*, v. 41, n. 1, p. 87-100, 2012.

ECKERT, Penelope. Variation and the indexical field 1. *Journal of sociolinguistics*, v. 12, n. 4, p. 453-476, 2008.

FARACO, C. E. O tratamento você em português: uma abordagem histórica. *Revista Fragmenta*, Curitiba, nº 13, 1996. p. 51-82.

FREITAG, Raquel Meister Ko et al. Como os brasileiros acham que falam? Percepções sociolinguísticas de universitários do Sul e do Nordeste. *Todas as Letras-Revista de Língua e Literatura*, v. 18, n. 2, 2016.

FREITAG, Raquel Meister Ko. et al. *Sociolinguística e política linguística: olhares contemporâneos*. São Paulo: Blucher, p. 141-160, 2016.

HAIR.Jr., J.; F., et al. *Multivariate Data Analysis*. New Jersey: Pearson International Edition. 2006.

HERÊNIO, Kelly Karine Pereira. *Tu e você em uma perspectiva intra-linguística*. 2006. Dissertação (Mestrado em Linguística). Instituto de letras e linguística da Universidade Federal de Uberlândia – Minas Gerais.

HOUAISS, A.; VILLAR, M.S. *Minidicionário Houaiss da língua portuguesa*. Elaborado no Instituto Antônio Houaiss de Lexicografia e Banco de Dados da Língua Portuguesa S/C Ltda. 3ª Ed. ver. e aum. Rio de Janeiro: Objetiva, 2009.
http://www.linguisticalfal.org/25_linguistica_030_065.pdf. Acesso em: 20 de junho de 2022.

IBGE – Instituto Brasileiro de Geografia e Estatística. 2021. Disponível em: <http://www.ibge.gov.br/>. Acesso em: 29 de abril de 2022.

KRUSKAL, W.H.; WALLIS, W.A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, v.47, n.260, p.583-621, 1952.

LABOV, William. Sociolinguistic patterns. Philadelphia: University of Pennsylvania Press, 1972b. *Padrões sociolinguísticos*. (Tradução Marcos Bagno, Marta Maria Pereira Scherre e Caroline Cardoso) São Paulo: Parábola Editorial, 2008 [1972b].

LABOV, William. *The social stratification of English in New York City*. São Paulo: Cambridge University, 2006 [1966].

LAMBERT, W. E. et al. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, v. 60, n. 1, p. 44-51, 1960.

LEVSHINA, N. *How to do Linguistic with R*. Amsterdam/Philadelphia: John Benjamins, 2015.

LEVON, Erez. Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society*, v. 43, n. 5, p. 539-566, 2014.

LOPES DE MELO, M. A. S. *Direcionalidade da mudança sonora: o papel do item lexical e da avaliação social*. 2017. Tese de Doutorado. Tese de Doutorado, Universidade Federal do Rio de Janeiro.

LOREGIAN-PENKAL, Loremi. *(Re)análise da referência de segunda pessoa na fala da Região Sul*. 2004. 260 f. Tese (Doutorado em Letras. Área de concentração: Estudos Linguísticos) Universidade Federal do Paraná, Curitiba, 2004.

LOREGIAN-PENKAL, Loremi. Alternância tu/você em Santa Catarina: uma abordagem variacionista. In: Estudos Linguísticos. 1, 2005, Campinas. *Anais do XXXIV Estudos Linguísticos – GEL*. Campina – SP, UNICAMP, 2005.

LUCCA, Nívea Naves Garcia. *A variação tu / você na fala brasiliense*. 2005. Dissertação (Mestrado em Linguística). Instituto de Letras / Universidade de Brasília.

MALHOTRA, N. K. *Pesquisa de marketing*. Porto Alegre: Bookman, 2012.

MARTINS, Germano Ferreira. *A alternância tu/você/senhor no município de Tefé – Estado do Amazonas*. 2010. Dissertação (Mestrado em Linguística). Instituto de Letras / Universidade de Brasília.

MENDES, Ronald Beline. *“Percepção e performance de masculinidades*. 2018. Tese (Livre Docência) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2018.

MENDES, Ronald Beline. *A terceira onda da Sociolinguística*. In: FIORIN, José Luiz. *Novos Caminhos da Linguística*. São Paulo: Contexto, p. 103-123, 2017.

MENDES, Ronald Beline. Nonstandard plural noun phrase agreement as an index of masculinity. *Language, Sexuality and Power. Studies in Intersectional Sociolinguistics*, v. 1, p. 105-129, 2016.

MENDES, Ronald Beline; OUSHIRO, Livia. '*Percepções Sociolinguísticas sobre as Variantes Tepe e Retroflexa na cidade de São Paulo*'. In: Hora, Dermeval da; Negrão, Esmeralda Vailate. *Estudos da Linguagem: casamento entre temas e perspectivas*. João Pessoa: Ideia/Editora Universitária. P. 229-245, 2011.

MEYERHOFF, Miriam. *Introducing sociolinguistics*. London and New York: Routledge, Taylor & Francis e-Library, 2006.

MIRANDA, A. L. A. *Crenças, atitudes e usos variáveis da concordância verbal com o pronome tu*. Rio de Janeiro: 2014. 157 f. Teses (Doutorado em Linguística), Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2014.

MODESTO, Artaxerxes. Formas de tratamento no Português Brasileiro: a alternância tu/você na cidade de Santos-SP. *Revista Letra Magna*. n. 07, p. 1-27, 2º semestre, 2007.

MODESTO, Artaxerxes. *Formas de tratamento no Português Brasileiro: a alternância tu/você na cidade de Santos-SP*. 2006, 152 f. Dissertação (Mestrado em Letras) – Universidade de São Paulo, Faculdade de Filosofia, Letras e Ciências Humanas, Programa de Pós-Graduação em Filologia e Língua Portuguesa, São Paulo, 2006.

MODESTO, Artaxerxes. Formas de tratamento no Português Brasileiro: a alternância tu/você na cidade de Santos-SP. *Revista Letra Magna*. n. 07, p. 1-27, 2º semestre, 2007.

MONTEIRO, J. L. *Pronomes pessoais*. João Pessoa: Edições UFC, 1994.

MOTA, Maria Alice. *A variação dos pronomes “tu” e “você” no português oral de São João da Ponte (MG)*. 2008. 125 f Dissertação (Mestrado em Linguística) – Universidade Federal de Minas Gerais, Faculdade de Letras, Programa de Pós-Graduação em Estudos Linguísticos, Belo Horizonte, 2008.

NASCENTES, Antenor. *Bases para a elaboração do Atlas Lingüístico do Brasil*. Rio de Janeiro: MEC: Casa de Rui Barbosa, 1958. v. I.

NASCENTES, Antenor. *O tratamento de ‘você’ no Brasil*. Letras, nº 5/6, p. 114-122, 1956.

OUSHIRO, Livia (2017). *Introdução à Estatística para Linguistas*, v.0.3(mai/2017). Cópia disponível com autora: oushiro@iel.unicamp.br.

OUSHIRO, Livia. Avaliações e percepções sociolinguísticas. *Estudos Linguísticos (São Paulo. 1978)*, v. 50, n. 1, p. 318-336, 2021.

OUSHIRO, Livia. *Identidade na Pluralidade: Avaliação, produção e percepção linguística na cidade de São Paulo*. Tese de Doutorado. USP, 2015. 394f.

PODESVA, Robert J. Phonation type as a stylistic variable: The use of falsetto in constructing a persona 1. *Journal of sociolinguistics*, v. 11, n. 4, p. 478-504, 2007.

R TEAM, Development Core. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2023. URL: <http://www.R-project.org/>.

RONCARATI, C.; ABRAÇADO, J. (eds.). *Português Brasileiro – contato lingüístico, heterogeneidade, história*. Rio de Janeiro: 7 Letras, 160-169.

RUMEU, M. C. B. *Para uma História do Português no Brasil: Formas Pronominais e Nominais de Tratamento em Cartas Setecentistas e Oitocentistas*. Volumes I e II. Dissertação (Mestrado em Língua Portuguesa) – Curso de Pós-graduação em Letras Vernáculas. Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2004.

RUMEU, M. C. B. Vestígios da pronominalização de Vossa Mercê > Você em missivas cariocas e mineiras: uma incursão pelo português brasileiro escrito nos séculos XIX e XX. *Veredas On-line*, v. 2, p. 36-55. 2012.

SACCONI, L. A. *Nossa gramática: teoria e prática*. São Paulo: Atual, 1986.

SANTOS, Viviane Maia dos. “*Tu vai para onde? ... Você vai para onde?*”: manifestações da segunda pessoa na fala carioca. 2012, 137 f. Dissertação (Mestrado em Letras Vernáculas: Língua Portuguesa). Faculdade de Letras. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

SANTOS, W. S. dos. *Percepções sociolinguísticas acerca da variação subjuntivo/indicativo em São Luís e São Paulo*. 2020. Tese (Doutorado em Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2020.

SCHERRE, M. M. P. et al. Por onde transitam o tu e o você no Nordeste? *Revista de Letras*, v. 40, n. 1, p.164-197, jan./jun. 2021. Disponível em: . Acesso em: 09 de setembro de 2023.

SCHERRE, M. M. P. et al. *Redesenhando o mapa dos pronomes tu/você/cê/ocê no português brasileiro falado*. In: WITCHES, P. H.; VIEIRAMACHADO, L. M. da C.; FURLAN, C. K. J.; NOGUEIRA, M. de O. (org.). *Conquistas e desafios dos estudos linguísticos na contemporaneidade*. Porto Alegre: Fi.org, 2020. p. 270-276. Disponível em: . Acesso em: 09 de setembro de 2023.

SCHERRE, M. M. P. et al. *Variação dos pronomes “tu” e “você”*. In: MARTINS, M. A.; ABRAÇADO, J. (org.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015. p. 133-172.

SENE, M. G. *A percepção sociolinguística de gêneros e sexualidade: efeitos da duração de /s/ e do pitch médio*. 2022. 220 f. Tese (Doutorado em Linguística e Língua Portuguesa). Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara, 2022

SENE, M. G. Relatório estatístico sobre o experimento de percepção sociolinguística da variável duração de /s/ em posição de coda. *Open Science Framework*. 2021. DOI: 10.17605/OSF.IO/V2GJY; Disponível em: <https://osf.io/v2gjy/>

SERRA, Flávia Pereira et al. *“Eu não digo ‘não’ duas vezes não”*: usos e percepções avaliativas sobre a dupla negação no português falado no maranhão. 2018. Dissertação – Universidade Federal do Maranhão, São Luís, 2018.

SOARES, I. C. R.; LEAL, M. da G. F. Do senhor ao tu: uma conjugação em mudança. Moara. *Revista do curso de mestrado (UFPA)*, Belém, n. 1, p. 27-64, mar/set 1993.

SORIANO, Larissa Grasiela Mendes. *Percepções sociofonéticas do (-R) em São Paulo*. 2016. Tese de Doutorado. Universidade de São Paulo. 2016.

SOUSA, Layane K. P. *A dupla negação do português no Maranhão: o que mostram os dados do Projeto ALiMA*. Relatório (Projeto de Iniciação Científica) — Universidade Federal do Maranhão, 2016.

APÊNDICES

Formulário de Percepção

Pesquisa de Linguística

Olá, meu nome é Rayane de Andrade Rodrigues. Sou aluna do mestrado em Letras da Universidade Federal do Maranhão - UFMA.

O meu objetivo é verificar como as pessoas percebem a fala uma das outras.

Antes de passar algumas instruções sobre como responder o questionário, gostaria de lhe agradecer a por participar, com sua contribuição. Muito obrigada!

Dê preferências pelo uso de um fone quando for ouvir os áudios. Ouça esses áudios quantas vezes quiser, antes de responder às perguntas sobre as vozes dos áudios. Depois de ouvir quantas vezes quiser cada um dos áudios, deverá responder algumas perguntas a respeito desses áudios.

Em seguida, dará sua opinião sobre informações sociais desses ouvintes.

Por fim, responderá a algumas perguntas sobre você, mas sem precisar se identificar.

Lembramos que não tem respostas certas ou erradas. O que levaremos em consideração é a sua impressão sobre as vozes que ouvir.

Fique à vontade para entrar em contato comigo caso tenha qualquer dúvidas em relação ao experimento.

Email: rayane.andrade@discente.ufma.br

Clique no vídeo abaixo e responda o questionário a seguir:



Essa pessoa soa como: *

	1	2	3	4	5	6	
Nada escolarizado	☐	☐	☐	☐	☐	☐	Muito escolarizado

Essa pessoa soa como: *

	1	2	3	4	5	6	
Nada inteligente	☐	☐	☐	☐	☐	☐	Muito inteligente

Essa pessoa soa como: *

	1	2	3	4	5	6	
Nada gentil	☐	☐	☐	☐	☐	☐	Muito gentil

Essa pessoa soa como: *

	1	2	3	4	5	6	
Nada formal	☐	☐	☐	☐	☐	☐	Muito formal

Essa pessoa soa como: *

	1	2	3	4	5	6	
Nada caxiense	☐	☐	☐	☐	☐	☐	Muito caxiense

Você acha que essa pessoa pertence a qual escolaridade? *

- ☐ Ensino fundamental
- ☐ Ensino médio
- ☐ Ensino superior

Você acha que essa pessoa pertence a qual faixa etária? *

- ☐ 20 a 30 anos
- ☐ 30 a 40 anos
- ☐ 40 anos ou mais

Você acha que essa pessoa pertence a qual classe social? *

- ☐ Classe alta
- ☐ Classe média
- ☐ Classe baixa

Você acha que essa pessoa é de qual lugar do Brasil? *

- ☐ Centro-Oeste (Goiás, Mato Grosso e Mato Grosso do Sul)
- ☐ Nordeste (Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe)
- ☐ Norte (Acre, Amapá, Amazonas, Pará, Rondônia, Roraima e Tocantins)
- ☐ Sudeste (Espírito Santo, Minas Gerais, Rio de Janeiro e São Paulo)
- ☐ Sul (Paraná, Rio Grande do Sul e Santa Catarina)

Informações sociodemográficas

Quantos anos você tem? *

.....

Qual o seu sexo ? *

- ☐ Masculino
- ☐ Feminino

Qual a sua escolaridade? *

- ☐ Ensino Fundamental
- ☐ Ensino Médio
- ☐ Ensino Superior

Você é de Caxias/MA? *

- ☐ Não
- ☐ Sim