



UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS

PALOMA BATISTA CARDOSO

**EFEITOS LINGUÍSTICOS E PARALINGUÍSTICOS NA INFERÊNCIA DOS
SENTIDOS INDICADOS POR (EU) ACHO QUE EM ENTREVISTAS
SOCIOLINGUÍSTICAS**

SÃO CRISTÓVÃO/SE
2021

UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS

PALOMA BATISTA CARDOSO

**EFEITOS LINGUÍSTICOS E PARALINGUÍSTICOS NA INFERÊNCIA DOS
SENTIDOS INDICADOS POR (EU) ACHO QUE EM ENTREVISTAS
SOCIOLINGUÍSTICAS**

Dissertação apresentada ao Programa de Pós-graduação em Letras da Universidade Federal de Sergipe como requisito parcial para defesa do título de Mestre em Letras. Área de concentração: Estudos Linguísticos. Linha de pesquisa: Descrição, Análise e Usos Linguísticos.

Profa. Dra. Raquel Meister Ko Freitag (Orientadora)
Prof. Dr. Hector Julian Tejada Herrera (Coorientador)

SÃO CRISTÓVÃO/SE
2021

PALOMA BATISTA CARDOSO

EFEITOS LINGUÍSTICOS E PARALINGUÍSTICOS NA INFERÊNCIA DOS SENTIDOS
INDICADOS POR (EU) ACHO QUE EM ENTREVISTAS SOCIOLINGUÍSTICAS

Dissertação aprovada em 18/02/2021

Dissertação apresentada ao Programa de Pós-graduação em Letras da Universidade Federal de Sergipe como requisito parcial para defesa do título de Mestre em Letras. Área de concentração: Estudos Linguísticos. Linha de pesquisa: Descrição, Análise e Usos Linguísticos.

Avaliação: _____

Data da defesa: _____

BANCA EXAMINADORA:

Profa. Dra. Raquel Meister Ko Freitag
(Orientadora)

Prof. Dr. Hector Julian Tejada Herrera
(Coorientador)

Prof. Dr. Rene Alain Santana de Almeida
(Membro Interno)

Profa. Dra. Mahayana Cristina Godoy
(Membro Externo)

AGRADECIMENTOS

A conclusão desta etapa da minha formação não seria possível sem o apoio e o trabalho em conjunto com professores e colegas de pesquisa, especialmente durante o ano de 2020, marcado pelas dificuldades de uma pandemia, Por isso, agradeço a minha orientadora, profa. Dra. Raquel Meister Ko. Freitag, por ter despertado em mim o interesse pela pesquisa desde a graduação (a primeira aula que assisti no curso de Letras foi dela), pelos conselhos que fizeram com que este trabalho tomasse o rumo apropriado, por ser meu exemplo de mulher na ciência e também pelos momentos de descontração enquanto estávamos no laboratório, regados a café e receitas deliciosas.

Ao meu coorientador, prof. Dr. Hector Julian Tejada Herrera, pela disponibilidade e pela paciência em lidar com todos os percalços que surgiram ao longo da execução desta pesquisa. Ao prof. Dr. Rene Alain Santana de Almeida, por todos os ensinamentos sobre prosódia, indicações de referências e pela constante disposição para contribuir com este trabalho; às professoras Dras. Leilane Ramos e Mahayana Godoy, pelas tão importantes contribuições no exame de qualificação.

A CAPES, pelo subsídio financeiro.

À equipe do GELINS pelo trabalho em equipe sem o qual não seria possível concluir a coleta dos dados que analisei, em especial a Manoel, Viviane, Fernanda e Vanessa, com quem compartilhei muitos cafezinhos, bolos e a cansativa mas imensamente satisfatória rotina da pós-graduação.

A Bruno, Lucas e Vitor, pela parceria na pesquisa desde a iniciação científica, pelas viagens inesquecíveis que renderam embates e muitos momentos nos quais rimos mesmo estando um tanto desesperados. A Diego, pelo constante apoio e incentivo que eu só poderia obter por meio de uma amizade verdadeira, forjada pela literatura; e a Gabriel, que desde o início desta pesquisa ouviu minhas dúvidas, incertezas e, em muitos momentos, me disse que tudo daria certo. A ele, que partiu prematuramente, dedico meus agradecimentos e também este texto.

A minha mãe, Lucila, por nunca ter medido esforços para que eu estudasse e por sempre dizer que posso fazer o que eu quiser. A força que me move vem dela!

À Crisevenyn, minha irmã de outra mãe e amor da minha vida, pelo apoio, pelo amor, por me ouvir e por vibrar comigo em cada etapa deste percurso. Também por me proporcionar as melhores risadas e festinhas (só nossas) quando precisei.

À Francielly, que também é minha irmã de outra mãe e amor da minha vida, pelo constante incentivo, por estar presente mesmo estando a milhares de km de distância e por ser a melhor companhia para falar tanto da vida acadêmica quanto das milhares de outras coisas pelas quais nos interessamos. O que Alicia Florrick uniu, nada separa.

À Graziela, *mein schatz*, por estar de mãos dadas comigo há tantos anos, pelo apoio e pela disposição para me ajudar e ouvir, mesmo às 3 da madrugada (risos). Obrigada pelas trocas, pela partilha de estrelas e sonhos, pois como bem sabemos:

“There’s no life without a dream, no power strong enough to hold it back”

Calling from the wild, Tarja Turunen.

RESUMO

(Eu) acho que é uma construção gramaticalizada no português brasileiro como um modalizador epistêmico polissêmico, que pode indicar sentidos de certeza, dúvida e incerteza (GALVÃO, 1999; FREITAG, 2003; GONÇALVES, 2003). Estes sentidos são inferidos pelos falantes no contexto de interação, e os estudos descritivos apontam que existem pistas linguísticas associadas a cada um destes sentidos, como o tipo de complemento introduzido por *(eu) acho que*, tópico discursivo e o envolvimento do falante (direto, indireto) com o que se fala. No entanto, pistas paralinguísticas também podem auxiliar na inferência destes sentidos, como as pistas acústicas e de expressões faciais. Do ponto de vista acústico, dúvida e incerteza são caracterizados por uma alta média de frequência fundamental, maiores valores de intensidade, duração e presença de pausas silenciosas e preenchidas, ao contrário de certeza (SILVA, 2008; OLIVEIRA, 2011; ANTUNES; AUBERGÉ; SASA, 2014; FERNANDES; ANTUNES, 2017). Há também diferenças gestuais: Swerts et al (2003) sugeriu que certeza é caracterizado pela expressão facial neutra enquanto incerteza, pela contração da linha das sobrancelhas e boca. Ao contrair os músculos da face, os falantes demonstram expressões que segundo Ekman (2000) são importantes para a manutenção da interação. O objetivo deste trabalho é investigar o efeito de variáveis linguísticas, acústicas e expressões faciais para a inferência dos sentidos de *(eu) acho que* na amostra Deslocamentos 2020, composta por 30 entrevistas sociolinguísticas gravadas em áudio e vídeo, realizadas com estudantes de graduação na Universidade Federal de Sergipe. Identificamos 1038 ocorrências de *(eu) acho que*, que foram codificadas quanto ao sentido e analisadas quanto a variáveis linguísticas (ocorrência de *acho que* ou *eu acho que*, escopo, presença de modalizador, tópico discursivo, experiência do falante e polaridade), prosódicas (médias da frequência fundamental, intensidade e duração de *(eu) acho que*, duração de pausas silenciosas e preenchidas) e expressões faciais (considerando os movimentos dos músculos que indicam raiva, deboche, nojo, medo, felicidade, neutra, tristeza e surpresa, captados por um script em linguagem Python). Os sentidos de *(eu) acho que* têm associação com o escopo, presença de outros modalizadores, tópico discursivo e experiência do falante, corroborando os resultados de outros estudos descritivos. Para as pistas acústicas, análises de variância entre o sentido de *(eu) acho que* e as variáveis prosódicas sugeriram efeito somente da intensidade e da duração. Os sentidos foram diferenciados pela força do som e pelo prolongamento de *(eu) acho que*: ocorrências dessa construção com menor intensidade, mais fracas e mais longas foram caracterizadas como dúvida e incerteza, respectivamente. A análise das expressões faciais evidencia movimentos de contração da linha das sobrancelhas e da boca quando os sentidos inferidos para *(eu) acho que* eram dúvida e incerteza, diferente de quando o sentido inferido era certeza. Os resultados das análises sugerem que além de variáveis linguísticas, recursos paralinguísticos também são relevantes para a inferência dos sentidos de *(eu) acho que*, corroborando perspectivas a sistematicidade da língua para além do nível linguístico.

Palavras-chave: Modalizadores. Recursos paralinguísticos. Variáveis linguísticas. Variáveis prosódicas. Expressões faciais.

ABSTRACT

In Brazilian Portuguese, *I think that* is a construction grammaticalized as a polysemous epistemic modalizer which may indicate meanings of certainty, doubt, and uncertainty (GALVÃO, 1999; FREITAG, 2003; GONÇALVES, 2003). These meanings are inferred by the speakers in the context of interaction. According to descriptive studies, there are linguistic clues associated with each one of them such as the kind of complement introduced by *I think that*, the discursive topic, and the involvement of the speaker (direct, indirect) with what is said. However, paralinguistic clues, e.g., acoustic clues and facial expressions, may also help in the inference of those meanings. From the acoustic point of view, doubt and uncertainty are characterized by a high mean of fundamental frequency, higher values of intensity, duration, and presence of silent and filled pauses, which is the opposite for certainty (SILVA, 2008; OLIVEIRA, 2011; ANTUNES; AUBERGÉ; SASA, 2014; FERNANDES; ANTUNES, 2017). There are also gestural differences: Swerts et al (2003) suggested that certainty is characterized by a neutral facial expression, while uncertainty is characterized by the contraction in the lines of the eyebrows and of the mouth. When contracting facial muscles, speakers demonstrate expressions which, according to Ekman (2000), are important to keep the interaction. The aim of this study is to investigate the effect of linguistic, acoustic, and facial expression variables for the inference of the meanings of *I think that* in the sample “Deslocamentos 2020”, which is composed of 30 sociolinguistic interviews recorded on both audio and video, carried out with undergraduate students in the Federal University of Sergipe. We identified 1038 occurrences of *I think that*, which were codified in relation to meaning and analyzed in relation to linguistic (occurrence of *I think that*, scope, presence of modalizer, discursive topic, speakers’ experiences, and polarity), prosodic (means of the fundamental frequency, intensity, and duration of *I think that*, duration of silent and filled pauses) and facial expression variables (considering the movements of muscles that indicate anger, mockery, disgust, fear, happiness, neutral, sadness, and surprise, captured through a script in Python language). The meanings of *I think that* have association with the scope, presence of other modalizers, discursive topic, and speakers’ experiences, corroborating the results of other descriptive studies. For acoustic clues, variance analysis between the meaning of *I think that* and the prosodic variables suggested an effect only in the intensity and the duration. The meanings were differentiated by the strength of the sound and the extension of *I think that*: weaker and longer occurrences of this construction were characterized as doubt and uncertainty, respectively. The analysis of facial expressions evidences movements of contraction in the line of the eyebrows and of the mouth when the inferred meanings for *I think that* were of doubt and uncertainty, in contrast to the inferred meaning of certainty. The results of the analyses suggest that besides linguistic variables, paralinguistic resources are also relevant to the inference of the meanings of *I think that*, corroborating the perspective of language systematicity beyond the linguistic level.

Keywords: Modalizers. Paralinguistic resources. Linguistic variables. Prosodic variables. Facial expressions.

LISTA DE ABREVIATURAS

GELINS – Grupo de Estudos em Linguagem, Interação e Sociedade

LAMID – Laboratórios Multiusuários de Informática e Documentação

NURC – Projeto da Norma Linguística Urbana Culta

F0 – Frequência fundamental

FACS – Facial action units

AUs – Action units

LISTA DE FIGURAS

Figura 1: expressão de dúvida.....	14
Figura 2: representação do movimento final de f0 nas expressões de certeza e incerteza das mulheres	36
Figura 3: representação do movimento final de f0 nas expressões de certeza e incerteza dos homens	37
Figura 4: registro do falante ao indicar incerteza (<i>funny face</i>).....	40
Figura 5: adaptação das emoções básicas propostas por Ekman e Friesen	44
Figura 6: facial action units (movimentos dos músculos do rosto).....	45
Figura 7: dataset brasileiro/colombiano	47
Figura 8: localização da Universidade Federal de Sergipe	50
Figura 9: localização de (eu) acho que no ELAN	51
Figura 10: alinhamento entre áudio e vídeo no ELAN	52
Figura 11: codificação prosódica de acho que (certeza)	61
Figura 12: codificação prosódica de acho que (incerteza)	61
Figura 13: processamento das expressões faciais no OpenFace	63
Figura 14: facial action units (movimentos dos músculos do rosto).....	81
Figura 15: action units de DAN2MF (certeza).....	82
Figura 16: action units de DAN2MF (dúvida)	83
Figura 17: action units de DAN2MF (incerteza)	84
Figura 18: action units de FRA3FF (certeza).....	85
Figura 19: action units de FRA3FF (dúvida)	86
Figura 20: action units de FRA3FF (incerteza).....	87

LISTA DE QUADROS

Quadro 1: trabalhos que utilizaram a amostra Deslocamentos 2020 como corpus.....	18
Quadro 2: tipos de achar.....	21
Quadro 3: desenvolvimento da gramaticalização de <i>achar</i>	23
Quadro 4: modelo de gramaticalização de <i>acho</i> e <i>parece</i>	24
Quadro 5: características gerais de certeza, incerteza e certeza de não saber	34
Quadro 6: tipologia das emoções básicas.....	43
Quadro 7: número de imagens que compõem os <i>datasets</i> brasileiro e colombiano	48
Quadro 8: perfil dos pesquisadores que conduziram as entrevistas sociolinguísticas	54

LISTA DE TABELAS

Tabela 1: valores de frequência fundamental inicial para certeza e dúvida.....	28
Tabela 2: valores de frequência fundamental final para certeza e dúvida	29
Tabela 3: valores de intensidade para certeza e dúvida	29
Tabela 4: duração dos enunciados.....	30
Tabela 5: frequência fundamental inicial para dúvida, incredulidade e certeza	31
Tabela 6: frequência fundamental final para dúvida, incredulidade e certeza.....	31
Tabela 7: intensidade máxima dos enunciados para dúvida, incredulidade e certeza.....	32
Tabela 8: duração dos enunciados para dúvida, incredulidade e certeza	33
Tabela 9: velocidade de fala e duração final das sílabas, em segundos	35
Tabela 10: ocorrências de pausas, self-talk e fillers em certezas e incertezas	37
Tabela 11: média e desvio padrão das taxas de articulação e elocução	38
Tabela 12: características acústicas e gestuais do SEP e PtD	41
Tabela 13: movimentos observados em vídeos.....	42
Tabela 14: associação entre função de <i>(eu) acho que</i> e tipo de ocorrência	68
Tabela 15: associação entre sentido de <i>(eu) acho que</i> e escopo.....	69
Tabela 16: associação entre sentido de <i>(eu) acho que</i> e modalizadores	71
Tabela 17: associação entre sentido de <i>(eu) acho que</i> e tópico discursivo	72
Tabela 18: associação entre sentido de <i>(eu) acho que</i> e a experiência do falante.....	74
Tabela 19: associação entre o sentido de <i>(eu) acho que</i> e a polaridade	75
Tabela 20: valores descritivos da média de f0 em relação ao sentido de <i>(eu) acho que</i>	76
Tabela 21: valores descritivos da média da intensidade em relação ao sentido de <i>(eu) acho que</i>	77
Tabela 22: valores descritivos da média da duração do modalizador em relação ao sentido de <i>(eu) acho que</i>	78
Tabela 23: valores descritivos da duração das pausas silenciosas em relação ao sentido de <i>(eu)</i> <i>acho que</i>	79
Tabela 24: valores descritivos da duração das pausas preenchidas em relação ao sentido de <i>(eu)</i> <i>acho que</i>	80

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVOS E HIPÓTESES DE PESQUISA	16
2	CERTEZA, DÚVIDA E INCERTEZA	20
2.1	(EU) ACHO QUE NO PORTUGUÊS BRASILEIRO	20
3	PROSÓDIA ACÚSTICA, GESTUAL E PARALINGUAGEM.....	27
3.1	CERTEZA, DÚVIDA E INCERTEZA EM ESTUDOS DE PROSÓDIA ACÚSTICA	27
3.2	CERTEZA, DÚVIDA E INCERTEZA EM ESTUDOS DE PROSÓDIA GESTUAL.....	38
3.3	EXPRESSÕES FACIAIS, CERTEZA, DÚVIDA E INCERTEZA.....	42
3.4	PROPOSTA DE ANÁLISE.....	48
4.	MÉTODO	50
4.1	CONSTITUIÇÃO E TRATAMENTO DA AMOSTRA.....	50
4.2	INSTRUMENTOS DE COLETA	54
4.2.1	<i>Roteiro de entrevista.....</i>	<i>54</i>
4.2.2	<i>Ficha social e termo de consentimento livre e esclarecido.....</i>	<i>55</i>
4.3	VARIÁVEIS CONTROLADAS	55
4.4	VARIÁVEIS LINGUÍSTICAS	55
4.4.1	<i>Ocorrência.....</i>	<i>55</i>
4.4.2	<i>Ocorrências excluídas</i>	<i>56</i>
4.4.3	<i>Escopo</i>	<i>57</i>
4.4.4	<i>Presença de modalizador</i>	<i>57</i>
4.4.5	<i>Tópico discursivo.....</i>	<i>58</i>
4.4.6	<i>Experiência do falante.....</i>	<i>59</i>
4.4.7	<i>Polaridade</i>	<i>59</i>
4.5	VARIÁVEIS ACÚSTICAS	60
4.5.1	<i>Medidas de frequência fundamental.....</i>	<i>61</i>
4.5.2	<i>Medidas de duração</i>	<i>62</i>
4.6	VARIÁVEIS PARALINGUÍSTICAS: EXPRESSÕES FACIAIS	62
4.7	TRATAMENTO ESTATÍSTICO	65
5	RESULTADOS E DISCUSSÕES	67
5.1	VARIÁVEIS LINGUÍSTICAS.....	67
5.1.1	<i>Ocorrência.....</i>	<i>67</i>
5.1.2	<i>Escopo</i>	<i>69</i>
5.1.3	<i>Presença de modalizador</i>	<i>70</i>
5.1.4	<i>Tópico discursivo.....</i>	<i>72</i>
5.1.5	<i>Experiência do falante.....</i>	<i>73</i>
5.1.6	<i>Polaridade</i>	<i>75</i>
5.2	VARIÁVEIS ACÚSTICAS.....	76
5.2.1	<i>Frequência fundamental de (eu) acho que</i>	<i>76</i>
5.2.2	<i>Intensidade de (eu) acho que.....</i>	<i>77</i>
5.2.3	<i>Duração de (eu) acho que</i>	<i>78</i>

5.2.4	<i>Duração das pausas silenciosas</i>	79
5.2.5	<i>Duração das pausas preenchidas</i>	79
5.3	VARIÁVEIS PARALINGUÍSTICAS: EXPRESSÕES FACIAIS	80
6	CONSIDERAÇÕES FINAIS	89
	REFERÊNCIAS	91
6.1	ANEXO 1.....	95
6.2	ANEXO 2.....	98
6.3	ANEXO 3.....	100

1 INTRODUÇÃO

Em uma conversa, o modo como uma pessoa fala influencia nossa compreensão do que foi dito. Considere a seguinte situação hipotética: você deseja reformar a sala de sua casa e, ao perguntar ao pedreiro se é seguro demolir uma parede, ele responde, pausadamente, de modo não assertivo e contraindo as sobrancelhas (mais ou menos como na figura 1), o seguinte:

(1) *Acho que* (pausa) não tem problema



Figura 1: expressão de dúvida

Fonte: <https://br.freepik.com/fotos-vetores-gratis/duvida>

Você confiaria nesse profissional? A partir dessa resposta, consideraria que ele tem segurança em relação ao que disse?

A sentença em (1) pode ser interpretada como início de certeza, dúvida e incerteza, respectivamente:

- a) Com certeza, não há problema em demolir a parede;
- b) Não sei se é possível demolir a parede;
- c) Talvez não seja possível demolir a parede.

Essas diferentes possibilidades de sentido decorrem do caráter polissêmico de *acho que*, formada pelo verbo “achar” cristalizado no presente do indicativo mais a conjunção “que”. Juntos, esses dois itens formam uma construção que emparelha forma e, conforme exemplificamos, diferentes significados (BYBEE, 2016). Na vida cotidiana, é comum passarmos por situações parecidas. Como diferenciamos cada um dos sentidos indicado por *acho que*?

Do ponto de vista linguístico, *acho que* (que pode ocorrer com ou sem a presença do sintagma nominal “eu” – *eu acho que*) é formada a partir do processo de gramaticalização e mudança semântica do verbo “achar”, que na sua forma latina (*afflare*) era um vocábulo pertencente à linguagem de caçadores. Ao longo dos séculos, o uso e os sentidos desse verbo mudaram: ele deixou de funcionar somente de modo pleno, selecionando um objeto direto (*eu achei um lápis*) e, junto a “que” passou a introduzir complementos oracionais, como exemplificado em (1). Desse modo, (*eu*) *acho que* passou a demonstrar a percepção do falante quanto a algo (CUNHA; OLIVEIRA; VOTRE, 1999), funcionando como construção modalizadora parentética epistêmica.

O uso de (*eu*) *acho que* para indicar certeza, dúvida e incerteza foi amplamente documentado e descrito em contextos formais e informais por meio de estudos de caráter funcionalista (ROSA, 1992; GALVÃO, 1999; FREITAG, 2003; GONÇALVES, 2003) os quais evidenciam a multiplicidade de usos e sentidos que essa construção indica, inferidos pelos pesquisadores a partir de sua sensibilidade de falante, por meio de variáveis estruturais e subjetivas: tipo de complemento oracional por ela introduzida, experiência direta ou indireta do falante em relação ao que é dito, tópico discursivo, etc.

Assim como na situação hipotética que descrevemos, na qual há presença de elementos que não são captados pelo código linguístico: o caráter não-assertivo da voz, a presença de pausas e a contração das sobrançelas, expressão facial que pode ser compreendida como indício de dúvida e de incerteza, na interação, para compreendermos o sentido de construções polissêmicas como (*eu*) *acho que*, recorreremos às informações que captamos na fala (se a pessoa fala sobre algo a partir das suas próprias experiências) à propriedades acústicas e gestuais.

O som da fala e as expressões faciais são elementos utilizados pelos falantes de maneira sistemática: enquanto falantes nativos do português brasileiro, sabemos que, grosso modo, o “tom” da nossa voz é tão importante quanto os itens lexicais que mobilizamos na fala. O “tom”, ou melhor, o som que produzimos quando falamos – assertivo ou não assertivo – é estudado, em seus parâmetros físicos e perceptuais, pela prosódia acústica. Do mesmo modo que o português brasileiro apresenta uma ordem de palavras que todos os falantes, sem exceção, vão obedecer: por exemplo, ninguém utilizará o determinante antes do nome, como em “casas as”; nosso idioma também apresenta parâmetros acústicos que serão igualmente respeitados: os sons da voz ao demonstrar certeza, dúvida e incerteza não são os mesmos. Do ponto de vista acústico, sentenças que indicam dúvida e incerteza são caracterizadas por uma curva de frequência

fundamental descendente, com maiores valores de duração e presença de pausas, ao contrário da certeza (SILVA, 2008; OLIVEIRA, 2011; ANTUNES; AUBERGÉ; SASA, 2014; FERNANDES; ANTUNES, 2017), diferenças percebidas pelos falantes desde a mais tenra idade.

Ao interagirmos com alguém, o que vemos também é relevante para a inferência de sentidos: antes do excerto (1), descrevemos (e ilustramos na figura 1) um falante de sobrancelhas contraídas ao proferir *acho que*. Por isso, é bastante provável que você tenha entendido que ali havia expressão de dúvida. Por meio da contração dos músculos do rosto, podemos demonstrar medo, raiva, tristeza, felicidade, nojo e surpresa (EKMAN; FRIESEN, 1975), emoções básicas entendidas como reações a estímulos externos que podem ser demonstradas por meio das expressões faciais. Além de indicar reações, essas mesmas expressões atuam na manutenção da interação (EKMAN, 2000), por isso elas são amplamente investigadas tanto na psicologia quanto em estudos de prosódia gestual que investigam se aspectos visuais, junto aos acústicos, são relevantes para a expressão e inferência dos sentidos de certeza, dúvida e incerteza. As evidências disponíveis sugerem que sim, tanto entre falantes do português brasileiro quanto do português europeu e holandês: quando não têm certeza sobre o que dizem, os falantes tendem a fazer pausas, falar mais devagar e a contrair a linha das sobrancelhas e a boca (SWERTS ET AL, 2003; CRUZ; SWERTS; FROTA, 2017; ANTUNES; AUBERGÉ; SASA, 2014), movimentos musculares presentes na expressão de nojo mas que, na interação, indicam outros sentidos, percebidos pelos falantes.

As evidências acústicas e gestuais que caracterizam certeza, dúvida e incerteza, em sua maioria, não foram obtidas a partir de estudos que investigam, de modo integrado, o funcionamento de parâmetros linguísticos, acústicos e gestuais na inferência de certeza, dúvida e incerteza. Além disso, a expressão desses sentidos é investigada tradicionalmente por meio de sentenças. A partir desta pesquisa, objetivamos contribuir com os estudos que descreveram os usos e sentidos indicados por *(eu) acho que* ampliando o modelo de análise ao investigar se variáveis linguísticas, acústicas e expressões faciais atuam juntas na inferência dos sentidos de certeza, dúvida e incerteza indicados por um item específico.

1.1 OBJETIVOS E HIPÓTESES DE PESQUISA

A partir do que expusemos, nesta pesquisa temos como objetivo geral investigar se variáveis linguísticas junto a parâmetros acústicos e expressões faciais, contribuem para a

inferência dos sentidos de certeza, dúvida e incerteza indicados por *(eu) acho que*. Especificamente, objetivamos:

1. Verificar a influência de variáveis linguísticas (ocorrência de *(eu) acho que*, escopo, presença de outros modalizadores, tópico discursivo, experiência do falante e polaridade) na inferência de certeza, dúvida e incerteza;
2. Verificar a influência de variáveis acústicas (média de frequência fundamental, intensidade, duração, presença de pausas silenciosas e preenchidas) na inferência de certeza, dúvida e incerteza;
3. Verificar a influência de expressões faciais na inferência de certeza, dúvida e incerteza;

A partir das evidências apresentadas por estudos de prosódia acústica e gestual (ANTUNES; AUBERGÉ; SASA, 2014; SWERTS ET AL, 2003) que sugerem que sentenças indicadoras de certeza, dúvida e incerteza são diferenciadas pelos parâmetros de frequência fundamental, intensidade, duração, presença de pausas silenciosas e preenchidas, mais a contração das sobrancelhas e boca, tomamos como hipótese que esses mesmos fatores diferenciarão os sentidos indicados por *(eu) acho que* na amostra que analisamos. Desse modo, a partir dos estudos já realizados sobre essa construção no português brasileiro pautados na gramaticalização (GALVÃO, 1999; FREITAG, 2003; GONÇALVES, 2003), propomos uma ampliação da análise da inferência de sentidos.

Para análise das variáveis acústicas, adotamos o modelo utilizado por Antunes, Aubergé (2015) e Fernandes, Antunes (2017) que, baseados em abordagens amplas de prosódia (CRYSTAL, 1969; CRUTTENDEN, 1986), sugerem que frequência fundamental, intensidade, duração, ocorrência de pausas preenchidas e silenciosas diferenciam sentidos. Para descrição e análise dos movimentos dos músculos faciais, seguimos as evidências apresentadas por Swerts et al (2003); Antunes; Aubergé e Sasa (2014); Cruz, Swerts e Frota (2017), que sugeriram que expressões faciais podem ser indicadores de certeza, dúvida e incerteza. Diferente destes autores, que descreveram os movimentos dos músculos faciais manualmente, neste trabalho utilizamos o protocolo automático proposto por Tejada et al (2021), executado por meio de *scripts* em linguagem Python que reconhecem e classificam as expressões faciais dos falantes. Como corpus de análise, utilizamos parte da amostra Deslocamentos 2020, que integra o banco de dados Falares Sergipanos (edital 02/2015 SENACON/MJ; edital CAPES/FAPITEC/SE 10/2016 PROMOB), aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de

Sergipe no processo CAAE: 0386.0.107.000-11. A amostra Deslocamentos 2020 é constituída por 100 entrevistas, realizadas com alunos de graduação da Universidade Federal de Sergipe, *campus* Professor José Aloísio de Campos. Desse total, utilizamos 30 neste estudo, gravadas simultaneamente em áudio e vídeo. A referida amostra foi constituída em parceria com outros 9 pesquisadores, bolsistas de iniciação científica (IC) e alunos de pós-graduação, membros do Grupo de Estudos em Linguagem, Interação e Sociedade – GELINS. Todos utilizaram esse mesmo corpus para a execução dos seguintes trabalhos:

PESQUISADOR	NÍVEL	TRABALHO
Bruno Felipe Marques Pinheiro	Mestrado	Pistas linguísticas e paralinguísticas para os sentidos dos diminutivos
Viviane Silva de Novais	Mestrado	Variação na concordância verbal de terceira pessoa do plural na fala de universitários sergipanos
José Manoel Siqueira da Silva	Mestrado	Variação no preenchimento da posição determinante antes de possessivos pré-nominais: padrões dialetais e contatos
Fernanda Gabrielle Costa Rodrigues	Mestrado	Variação na regência de complementos locativos de verbos de movimento na fala de universitários da UFS
Lucas Santos Silva	Mestrado	Gradiência da palatalização /t/ e /d/ em Sergipe
Victor Rene Andrade Souza	IC	Transposição de processos fonológicos variáveis da fala para a leitura oralizada
Vitória Laís Santos Silva	IC	Taxa de produtividade a partir dos processos fonológicos que passam da fala para leitura oralizada
Mauro Monteiro de Araújo Júnior	IC	Ocorrência do futuro do pretérito e a colocação pronominal

Quadro 1: trabalhos que utilizaram a amostra Deslocamentos 2020 como corpus

Fonte: elaborado pela autora

Além de possibilitar a descrição e análise linguística por meio da documentação em áudio, os dados em vídeos poderão ser utilizados em trabalhos posteriores que tenham como objeto de estudo a análise de expressões faciais¹.

O presente texto está dividido da seguinte forma: no capítulo 2, descrevemos estudos sobre *(eu) acho que* pautados na gramaticalização, a partir dos quais propomos expandir a análise dos sentidos indicados por *(eu) acho que*. Depois, conceituamos e caracterizamos certeza, dúvida e incerteza. No capítulo 3, apresentamos evidências acústicas e gestuais que sustentam as hipóteses que norteiam esta pesquisa. No capítulo 4, descrevemos o método a partir do qual este trabalho foi executado. No capítulo 5, reportamos os resultados obtidos e, por fim, no capítulo 6 expomos nossas considerações finais.

¹ Entre 2019 e 2020, a professora Dra. Raquel Meister Ko. Freitag e o professor Dr. Hector Julian Tejada Herrera desenvolveram um banco de dados de expressões faciais que categoriza as expressões de raiva, nojo, medo, felicidade, deboche, neutra, tristeza e surpresa de falantes brasileiros e colombianos, a fim de investigar também a influência dos movimentos dos músculos do rosto na interação e, por conseguinte, na inferência dos sentidos expressos por construções linguísticas.

2 CERTEZA, DÚVIDA E INCERTEZA

Neste capítulo, inicialmente descrevemos trabalhos descritivos que apresentaram evidências do comportamento de *(eu) acho que* como construção modalizadora parenética epistêmica (ROSA, 1992; GALVÃO, 1999; GONÇALVES, 2003; FREITAG, 2003) a fim de justificar a ampliação do seu modelo de análise. A partir de Traugott e Dasher (2014); Cunha, Oliveira e Votre (1999) falamos sobre os processos de mudança semântica e gramaticalização que tornaram *(eu) acho que* polissêmica, podendo indicar certeza, dúvida e incerteza, sentidos definidos a partir de Silva (2008) e Oliveira (2011) e, conforme hipotetizamos, diferenciados por parâmetros estruturais, acústicos e por expressões faciais.

2.1 (EU) ACHO QUE NO PORTUGUÊS BRASILEIRO

A investigação sobre os sentidos indicados por *(eu) acho que* no português brasileiro tem sido executada a partir de dados coletados em diferentes situações de interação, entre falantes com diferentes graus de escolaridade. Rosa (1992), valendo-se de dados coletados em elocuições formais e entrevistas sociolinguísticas realizadas com estudantes e professores universitários, estudou a ocorrência de *acho que* como marcador de atenuação. Para a referida autora, esse item é utilizado para preservação da face do locutor e seu uso está relacionado ao grau de comprometimento com o que é dito, às experiências do falante e aos seus objetivos em uma dada situação comunicativa. De acordo com Rosa (1992, p. 91), *(eu) acho que* é um marcador modalizador que tem como finalidade demonstrar a opinião de quem fala e indica maior ou menor grau de adesão/comprometimento com o que é dito. A adesão é demonstrada também por elementos que estão além da estrutura linguística, nos quais a autora inclui risos, pausas e hesitações, que apesar de relevantes não foram analisados no trabalho dela.

O uso de *(eu) acho que* para indicar o grau de comprometimento do falante também pode ser observado a partir de dados do NURC (Projeto da Norma Linguística Urbana Culta) e de dicionários que datam da década de 1960: Galvão (1999) a partir dessas fontes e sob o prisma da gramaticalização, com foco na mudança categorial, descreveu cinco tipos de *achar*:

PROPRIEDADES	ACHAR 1	ACHAR 2 (APRECIAÇÃO)	ACHAR 2'	ACHAR 3 (PALPITE)	ACHAR 4
Paráfrase	Encontrar; procurar; descobrir	Considerar; pensar; afirmar	Considerar	Supor; é possível	Talvez; provavelmente
Variabilidade temporal	Presente; perfeito; imperfeito	Presente; perfeito	Presente; perfeito	Presente; perfeito	Presente
Variabilidade de modo	Indicativo; imperfeito; subjuntivo	Indicativo; subjuntivo	Indicativo; subjuntivo	Indicativo	Indicativo
Pessoas do verbo	1 ^a , 2 ^a , 3 ^a S/P	1 ^a , 2 ^a , 3 ^a S/P	1 ^a , 2 ^a , 3 ^a S/P	1 ^a , S	1 ^a , S
Presença e tipo do argumento interno	SN que pode vir ou não realizado foneticamente, com o traço [+concreto]	Oracional	SN que pode vir ou não realizado foneticamente, é predicado por um qualificador [+abstrato]	Oracional	Perda do argumento interno
Tipo de sentença em que aparece	Absoluta; hipotética; núcleo; encaixada	Principal	Absoluta; núcleo; principal	Principal	Depois de uma emissão completa de raciocínio e pausa (fora da oração)
Caracterização	Pressupõe o encontro entre duas entidades do mundo real; [-controle], incerteza	Performativo; modalizador; avaliativo depende de fatores inerentes ao falante	Performativo; modalizador; avaliativo; maior grau de certeza	Modalizador; fatores externos ao falante fazem com que a incerteza passe para o nível das possibilidades	Incerteza, dúvida, probabilidade

Quadro 2: tipos de achar
Fonte: Galvão, 1999, p. 98

Achar 1 [Sentença nominal – *achar* – sentença nominal]: o verbo atua como verbo pleno e estrutura toda a sentença, selecionando um argumento externo, ou seja, um agente e um

objeto. Há duas possibilidades semânticas para achar em sua forma plena: ação e processo, respectivamente, como nos exemplos que seguem, retirados de Galvão, 1999, p. 73:

a) *Acha* o café pra mim, *acha* o açúcar (NS)

b) Valdinei é o meu mais velho...Valdinei num tá estudando tá trabalhando...quero vê si *acho* uma vaga...vai volta a istudá (RD F, 40^a, 0)

Achar 2 [Sintagma nominal – *achar* – que]: o verbo² seleciona dois argumentos e um complemento oracional para introduzir considerações/apreciações, a partir das quais o falante emite julgamentos. Para a referida autora, achar 2 é uma extensão de achar 1 processo: “a partir de certos pontos de natureza concreta fornecidos pelo item fonte, o falante, por associação, explora o entendimento de conceitos menos concretos”, como em:

c) *Acho que* a economia é mais forte do que a lei

Achar 2’ [Sintagma nominal – *achar* – Sintagma adjetival]: o verbo introduz um sintagma adjetival que indica opinião/certeza e pode ou não ser precedido por preposição:

d) *eu acho* a vida da gente muito curta (Rondon, M, 32 a 2)

e) *eu acho* (que) a vida da gente (é) muito curta

Achar 3 [Sintagma nominal – *achar* – que]: neste caso, tem-se um modalizador epistêmico que põe a incerteza no nível da possibilidade e, por sua vez, pode ser permutado por *deve ser*. Achar 3 introduz um complemento oracional que, por indicar o julgamento do falante, é conjugado somente na primeira pessoa:

f) Eu *acho* que Rondon *deve ser* muito legal

Achar 4 [S – *achar*]: embora, do ponto de vista morfológico, seja verbo, se comporta como advérbio: pode estar no início, meio ou fim da sentença, sem ligação com nenhum argumento. Sua função é predicar todo o conteúdo, de modo a indicar certeza ou probabilidade:

g) Tristeza...*acho*...sei lá (RD, F, 30 A, 1)

Os tipos de *achar* descritos por Galvão (1999), presentes tanto em dados de fala quanto de escrita, formam um *continuum* que abrange *achar* como verbo pleno, *acho que* como

² Aqui, referimo-nos a *acho que* como verbo, pois esse é o termo utilizado por Galvão (1999), diferente da abordagem por nós assumida. Com base em Bybee (2016), entendemos que não é somente *achar* que é gramaticalizado, mas toda a construção (*eu*) *acho que*.

elemento que introduz complementos oracionais sem modificá-los e achar 4, que não introduz nenhum complemento e se comporta como advérbio. De modo semelhante, para Gonçalves (2003) *acho que* é uma construção que, em decorrência do processo de mudança semântica e gramaticalização, indica opinião/certeza e incerteza em relação a um conteúdo proposicional, com um desenvolvimento categorial parecido com o proposto por Galvão (1999):

Achar: (verbo pleno) > (v. encaixador de predicação) > (v. encaixador de proposição) > (auxiliar perifrástico) > (satélite)

Quadro 3: desenvolvimento da gramaticalização de *achar*
Fonte: adaptado de Gonçalves, 2003

Em uma abordagem semelhante, Freitag (2003) desenvolveu um estudo sincrônico no qual investigou o processo de gramaticalização de *acho que* e *parece que* em dados de entrevistas sociolinguísticas. A autora uniu duas perspectivas: a da gramaticalização, com foco na mudança semântica de *acho que* e *parece que* e a da variação, sob justificativa de que, em determinado ponto, essas duas construções variam. O processo de mudança semântica e gramaticalização de *acho que* e *parece que* tem como consequência a formação de traços de evidencialidade, ou seja, marcas que “indicam não só noções de origem de informação, mas também atitudes do falante e crenças sobre o seu conhecimento” (FREITAG, 2003, p. 11). Sendo assim, o uso de *acho que* ou *parece que* para indicar dúvida é condicionado pelo tipo de sequência discursiva (narrativas de experiência de vida e opinativas, por exemplo) e pelo tipo de experiência do falante com o tópico abordado: direta, quando o falante vivenciou a situação; mediana, quando a situação foi vivenciada por intermédio de pessoas próximas; e indireta, quando o falante soube da vivência/experiência por fontes que não lhes são próximas. Os resultados da pesquisa apontaram que

na função *marcador de dúvida*, a forma *acho* tende a ocorrer em contextos semântico discursivos de maior envolvimento do falante com o assunto discorrido, quando está sob o escopo de ações/situações imperfectivas, em sequências discursivas argumentativas, e em assuntos cuja complexidade foi considerada menor; para a forma *parece*, o ambiente semântico discursivo mais propício para a sua ocorrência no desempenho da função *marcador de dúvida* é caracterizado pelo menor envolvimento do falante com o assunto discorrido, sob o escopo de ações/situações perfectivas, em sequências discursivas narrativas e em assuntos cuja complexidade foi considerada maior (FREITAG, 2003, p. 102, grifos da autora).

O processo sofrido por *acho que* e *parece que* tem como consequência a reanálise dessas construções: elas passam a funcionar como um único item que seleciona complementos mais ou menos integrados, cuja estrutura é representada no quadro 4:

	ACHO	PARECE
V. pleno	“encontrar”	“assemelhar-se”
V. principal + complemento oracional	Acho que: marcador de opinião	Parece que: marcador de percepção
V. principal + complemento Integrados	Acho que: marcador de dúvida	Parece que: marcador de dúvida
Verbo + predicativo	Acho + adjetivo	Parece + adjetivo

Quadro 4: modelo de gramaticalização de *acho* e *parece*

Fonte: adaptado de Freitag, 2003, p. 86

A autora exemplificou as estruturas esboçadas acima para *acho que* com os seguintes exemplos (FREITAG, 2003, p. 46):

Verbo principal + complemento oracional (opinião):

35. Quando não era corrida de bastão era pata cega, aquele negócio. É porque na época eu acredito que na época, os professores de educação física não eram formados realmente em educação física. Eu *acho que* eles eram realmente improvisados.

Verbo principal + complemento integrado (dúvida):

41. E: A senhora tem filhos?
 F: Tenho, tenho cinco filhos.
 E: É? E qual a idade deles?
 F: Olha, eu tenho uma filha com- ela vai fazer é eu *acho que* ela vai fazer trinta e oito anos. E tenho um filho com trinta e seis, outra filha com trinta e quatro, outro filho com trinta, eu já disse quatro, né? e o mais moço vai fazer vinte e nove anos. SC FLP

No modelo aplicado por Freitag (2003), verbo mais complemento oracional e verbo mais complemento integrado correspondem a achar 2' e achar 3, propostos por Gonçalves (1999). Além do tipo de complemento introduzido pela construção modalizadora, em estudo posterior, Freitag (2007b), sugeriu que no processamento das sentenças um fator que influencia a inferência de sentidos é a polaridade positiva ou negativa, caracterizada, estruturalmente, pela presença ou ausência da negação expressa por “não”, como em 35 e 41. Nesses dois recortes de fala, *acho que* indica sentidos que, na interação, são inferidos pelos participantes tanto pelos contextos quanto pela estrutura das sentenças introduzidas por essa construção.

Nos estudos descritivos que listamos, ao indicar opinião, certeza, dúvida ou incerteza (a depender dos rótulos adotados pelas pesquisadoras) quanto a um complemento oracional, (*eu*)

acho que demonstra diferentes graus de asserção e situa o falante quanto a uma proposição (FREITAG, 2007a). Neste caso, essa construção atua como parentética epistêmica, pois “não complementa ou modifica um outro elemento dentro da frase. Na realidade, interage com a força assertiva da frase em que ocorre, pois os parentéticos epistêmicos possuem propriedades modais-epistêmicas, relacionadas à codificação da atitude do falante.” (FREITAG, 2007a, p. 84).

Galvão (1999); Freitag (2003); Gonçalves (2003) são unânimes: *(eu) acho que* é resultado da gramaticalização e dão indícios do conhecimento do falante sobre algo. Para Traugott e Dasher (2004), esse processo é motivado pela mudança semântica, que acontece a partir de dois processos cognitivos: metáfora e metonímia, que impulsionam a expansão conceitual sofrida por um item cujo sentido vai do mais concreto ao mais abstrato. A expansão ocorre do ponto de vista do falante e por isso uma das suas consequências é o desenvolvimento de marcas de subjetivação.

Segundo Cunha, Oliveira e Votre (1999), a mudança semântica de *achar* segue a escala proposta por Heine et al (1991): corpo > objeto > processo > espaço > tempo > qualificação³. Em sua forma latina, *achar* (*afflare*), que significava soprar, era um elemento da linguagem de caçadores. Ao longo das mudanças semânticas que sofreu, de soprar passou a significar sentir a proximidade da caça pelo odor, arejar e, posteriormente, descobrir ou encontrar a caça. Nesse percurso, de verbo intransitivo, *achar* tornou-se transitivo e passou a selecionar um objeto direto, tornando-se um verbo de cognição/percepção. Seguindo o processo, deixou de indicar algo concreto, ou seja, a caça. Junto ao complementizador *que*, formou *(eu) acho que*, construção polissêmica que indica certeza, dúvida e incerteza, sentidos formados por processos cognitivos complexos.

De acordo com Silva (2008, p. 36), o uso dos termos dúvida e incerteza não é aprofundado nas produções científicas. Por isso, ela discutiu a diferença entre os dois nas abordagens pragmática, da prosódia acústica e modalização epistêmica. Na abordagem pragmática de Searle (1995), ter ou não certeza diz respeito à intencionalidade do falante em demonstrar algo ao seu ouvinte, o que implica diferenças no modo como se fala. Do ponto de vista acústico, certeza é caracterizado por uma curva de frequência fundamental descendente e dúvida e incerteza por uma curva ascendente, diferenças inferidas pelos falantes conforme

³ No processo de mudança semântica, *achar* não passa pela etapa de objeto. De espaço físico (achar algo em algum lugar) esse item passa a indicar processo.

indicado pelos estudos de Silva (2008), Oliveira (2011), Antunes; Aubergé; Sasa (2014), Fernandes; Antunes (2017). Do ponto de vista gestual, certeza é caracterizado por uma expressão facial neutra, enquanto dúvida e incerteza são caracterizados pela contração da linhas das sobrancelhas e da boca (SWERTS ET AL, 2003; ANTUNES; AUBERGÉ; SASA, 2014)

De acordo com Silva (2008), a incerteza

permite uma gradação de acordo com o nível de conhecimento sobre um assunto. (...) Incerteza pressupõe uma gradação até a certeza definitiva. Ela aproxima-se ainda da indecisão, da hesitação. Já a expressão de dúvida não. A dúvida aproxima-se da pergunta feita para nós mesmos, sem pedir precisamente uma resposta. Ela aparece e não permite gradação. Não tem como ter mais ou menos dúvida: tem-se a dúvida. Podemos ainda pensar na dúvida como ausência de certeza. (SILVA, 2008, p. 39-40).

Concepção semelhante foi adotada por Oliveira (2011), que investigou os aspectos acústicos que caracterizam orações que indicam dúvida, incredulidade e incerteza. Essa autora definiu a dúvida como

a alternância de um sim e um não, e a incerteza como uma asserção fraca. Considerando tais atitudes em um contínuo, sugerimos que a dúvida e a certeza estejam em dois extremos opostos, e a incerteza na transição. Em relação à incredulidade adotaremos o conceito de quebra de expectativa, informação contrária ao que se esperava (OLIVEIRA, 2011, p. 46).

Galvão (1999) e Freitag (2003) sinalizaram que *(eu) acho que* pode indicar tanto certeza quanto incerteza, sentidos inferidos pelo grau de experiência do falante em relação a algo. A partir do que Oliveira (2011) e Silva (2008) propuseram, entendemos que dúvida, por se aproximar de uma pergunta a nós mesmos, está associado a experiências pessoais. Incerteza, por permitir gradação, está associado às experiências indiretas. Os sentidos de certeza, dúvida e incerteza também são indicados parâmetros acústicos e por elementos paralinguísticos, ou seja, por expressões faciais que apesar de estarem além do código linguístico elas contribuem para a inferência de sentidos (TRAUNMÜLLER, 2000).

3 PROSÓDIA ACÚSTICA, GESTUAL E PARALINGUAGEM

Neste capítulo, descrevemos estudos sobre prosódia acústica (SILVA, 2008; OLIVEIRA, 2011; ANTUNES; AUBERGÉ; SASA, 2014; ANTUNES; AUBERGÉ, 2015; FERNANDES; ANTUNES, 2017) e gestual (SWERTS ET AL, 2003; CRUZ; SWERTS; FROTA, 2017) que apontam evidências de que sentenças que indicam certeza, dúvida e incerteza possuem parâmetros de expressão distintos e que, conforme tomamos como hipótese, atuam na inferência dos sentidos indicados por *(eu) acho que* nas entrevistas sociolinguísticas que constituem a amostra Deslocamentos 2020.

3.1 CERTEZA, DÚVIDA E INCERTEZA EM ESTUDOS DE PROSÓDIA ACÚSTICA

A prosódia acústica diz respeito ao modo como falamos. Os sons emitidos na fala são produzidos por três subsistemas:

- a) Respiratório, composto pelos pulmões, laringe e véu palatino;
- b) Laríngeo, composto por cartilagens, ligamentos, músculos e tecido membranoso;
- c) Supralaríngeo, composto pela língua, bochechas, maxilar, palato (duro e véu) e a cavidade nasal (BARBOSA; MADUREIRA, 2015).

O fluxo de ar que adentra a cavidade torácica na inspiração fornece energia para a produção de ondas sonoras, originadas pelas vibrações das pregas vocais. Tais vibrações geram a frequência fundamental (f_0) que, junto às medidas de duração (temporal) e intensidade, que expressa quão forte é um som, constituem os correlatos físicos da prosódia. Do ponto de vista perceptivo, a prosódia possui três correlatos: *pitch*, que diz respeito “à percepção de um som como grave ou agudo em um trecho sonoro correspondente a determinada unidade linguística ou enunciado” (BARBOSA, 2019, p. 30); duração percebida, ou seja, “a sensação que permite avaliar se uma unidade é longa ou curta em relação a outra” (BARBOSA, 2019, p. 31); e volume, a sensação que se dá na escala forte/fraco, determinado por uma relação logarítmica com a intensidade.

Nos estudos sobre o português brasileiro, os sentidos de certeza, dúvida e incerteza são caracterizados por parâmetros prosódicos e entoacionais próprios, captados por gravadores e por microfones de alta qualidade capazes de registrar os aspectos dinâmicos da fala com precisão (OLIVEIRA JR, 2014). Na literatura, a discussão sobre prosódia e entonação é bastante ampla (HIST; DI CRISTO, 1998). Em abordagens restritas (HALLIDAY, 1970;

COUPER-KUHLEN, 1986), entonação é uma propriedade independente, caracterizada pela frequência fundamental, já em abordagens amplas (CRYSTAL, 1969; CRUTTENDEN, 1986), frequência fundamental, intensidade, duração e pausas constituem as propriedades acústicas do enunciado e diferenciam seus sentidos.

Na interação, a fala tem como objetivo não apenas transmitir informação, mas também a atitude ou intenção do locutor. Foi a partir dessa afirmação que Silva (2008) investigou, em uma abordagem ampla, quais parâmetros acústicos diferenciam certeza e dúvida, mas não antes de problematizar a diferença entre dúvida e incerteza. A referida pesquisadora assumiu a proposta de Machado (1996, p. 2), para quem “a dúvida é feita da coexistência antagônica do sim e do não, e a incerteza uma alternância entre um sim e um não”. Silva (2008) não pontuou qual fator diferenciaria dúvida e incerteza. A princípio, para fins de análise, ela considerou os dois sentidos como sinônimos e descreveu a produção de certeza e dúvida a partir de dados gravados por 10 informantes do sexo masculino, alunos do curso de teatro que, em laboratório, interpretaram sentenças e executaram uma tarefa de leitura. Ao todo, foram produzidas 270 expressões: 90 de certeza, 90 de dúvida e 90 de leitura, a partir das quais a autora controlou os parâmetros de frequência fundamental inicial e final do centro da sílaba de cada sentença, amplitude melódica da sílaba proeminente, tessitura, duração, intensidade e pausas silenciosas. Aqui, reportamos apenas os resultados da frequência fundamental, intensidade, duração e pausas.

Os resultados das análises de Silva (2008) demonstraram que a frequência fundamental, inicial e final, diferenciou certeza de dúvida. Este segundo sentido apresentou média maior do que o primeiro, diferença estatisticamente significativa, conforme demonstrado pelos valores reportados nas tabelas 1 e 2, respectivamente:

MÉDIA			DESVIO PADRÃO			SIGNIFICÂNCIA		
Certeza	Dúvida	Leitura	Certeza	Dúvida	Leitura	C-D	C-L	D-L
127,93	141,45	119,09	23,4001	34,1495	18,4359	0,0002	0,0272	0,0000

Tabela 1: valores de frequência fundamental inicial para certeza e dúvida
Fonte: adaptado de Silva, 2008, p. 66, com grifos da autora

MÉDIA			DESVIO PADRÃO			SIGNIFICÂNCIA		
Certeza	Dúvida	Leitura	Certeza	Dúvida	Leitura	C-D	C-L	D-L
100,02	105,48	96,06	28,7119	20,6582	19,6936	0,1364	0,3518	0,0029

Tabela 2: valores de frequência fundamental final para certeza e dúvida

Fonte: adaptado de Silva, 2008, p. 69, com grifos da autora

Intensidade também diferenciou certeza de dúvida. Para Silva (2008), que assumiu uma abordagem pragmática pautada em Searle (1995), essas diferenças refletem intenções distintas do falante para com o ouvinte. Na tabela 3 reportamos a média, desvio padrão e os valores de significância desse parâmetro:

MÉDIA			DESVIO PADRÃO			SIGNIFICÂNCIA		
Certeza	Dúvida	Leitura	Certeza	Dúvida	Leitura	C-D	C-L	D-L
65,52	59,90	62,99	4,05328	4,66652	3,82755	0,0000	0,0000	0,0000

Tabela 3: valores de intensidade para certeza e dúvida

Fonte: adaptado de Silva, 2008, p. 100, com grifos da autora

Silva (2008) controlou a duração das sentenças e das sílabas proeminentes com suas duas sucessoras e antecessoras. Todas essas demarcações apresentaram resultados estatisticamente significativos e indicaram diferença entre dúvida, sempre com maiores médias em comparação à certeza e às realizações de leitura. Na tabela 4 reportamos a média, desvio padrão e os valores de significância da duração de certeza e dúvida:

MÉDIA			DESVIO PADRÃO			SIGNIFICÂNCIA			
Sujeitos	Certeza	Dúvida	Leitura	Certeza	Dúvida	Leitura	Certeza	Dúvida	Leitura
1	1,07	1,36	1,23	0,117622	0,283653	0,234511	0,0217	0,2907	0,3982
2	0,97	1,24	1,09	0,09727	0,168811	0,094465	0,0001	0,0949	0,0318
3	1,09	1,22	1,18	0,096011	0,134505	0,15556	0,0711	0,2756	0,7408
4	1,18	1,60	1,52	0,174225	0,097506	0,216768	0,0000	0,0003	0,5374
5	1,03	1,35	1,18	0,112866	0,112198	0,151845	0,0000	0,0299	0,0207
6	1,06	1,97	1,08	0,127733	0,330008	0,12896	0,0000	0,9712	0,0000
7	1,06	1,60	1,11	0,133879	0,459939	0,197869	0,0011	0,9227	0,0030
8	0,94	1,03	0,97	0,096238	0,132029	0,104816	0,1676	0,7254	0,5227
9	0,88	1,49	0,90	0,101995	0,274524	0,13976	0,0000	0,9700	0,0000

Tabela 4: duração dos enunciados

Fonte: adaptado de Silva, 2008, p. 93, com grifos da autora

Depois de apresentar os dados de produção, Silva (2008) retomou a discussão sobre o que poderia diferenciar dúvida de incerteza. Ela assumiu que a dúvida diz respeito a um questionamento individual, enquanto incerteza é demonstrado quando respondemos a uma solicitação (SILVA, 2008, p. 118). A partir dessa afirmação, a referida pesquisadora passou a categorizar os dados gravados pelos atores como indicativos de incerteza, o que faz sentido quando consideramos o tipo de dado analisado: orações pré-produzidas, inseridas em parágrafos cujo objetivo era contextualizar a expressão de cada sentido.

Em uma abordagem semelhante, Oliveira (2011), a fim de investigar os parâmetros prosódicos que caracterizam dúvida, incerteza e incredulidade em enunciados e em dados de leitura partiu do princípio de que esses três sentidos indicam diferentes graus de força ilocucionária e por isso não são sinônimos. Ela considerou

a definição de dúvida como a alternância de um sim e um não, e a incerteza como uma asserção fraca. Considerando tais atitudes em um contínuo, sugerimos que a dúvida e a certeza estejam em dois extremos opostos, e a incerteza na transição. Em relação à incredulidade adotaremos o conceito de quebra de expectativa, informação contrária ao que se esperava (OLIVEIRA, 2011, p. 46).

Para a execução do estudo que propôs, Oliveira (2011) controlou a variação da frequência fundamental (valores iniciais, finais, mínimos e máximos), intensidade, presença de pausas e duração. A amostra analisada por essa autora foi constituída por 16 informantes: 8 do sexo feminino e 8 do sexo masculino. Cada um produziu 10 enunciados de cada sentido. O

mesmo foi feito para a obtenção dos dados de leitura. A comparação dos valores de frequência fundamental inicial dos enunciados indicou diferenças significativas entre dúvida e incerteza:

SEXO	ATTITUDES	MÉDIA	SIGNIFICÂNCIA (P-VALOR)		
			Duv x Incred	Duv x Incert	Incert x Incred
Masculino		Média (DP)			
	Dúvida	7,74 (5,60)	0,000*	0,000*	0,959
	Incredulidade	4,85 (4,53)			
Incerteza	4,88 (3,80)				
Feminino	Dúvida	17,07 (4,30)	0,000*	0,000*	0,664
	Incredulidade	14,49 (3,94)			
	Incerteza	14,20 (4,38)			

Tabela 5: frequência fundamental inicial para dúvida, incredulidade e certeza

Fonte: Oliveira, 2011, p. 90, com grifos da autora

A frequência fundamental final dos enunciados também foi estatisticamente significativa, de acordo com os resultados descritos na tabela 6:

SEXO	ATTITUDES	MÉDIA	SIGNIFICÂNCIA (P-VALOR)		
			Duv x Incred	Duv x Incert	Incert x Incred
Masculino		Média (DP)			
	Dúvida	7,74 (5,60)	0,144	0,003*	0,079
	Incredulidade	4,85 (4,53)			
Incerteza	4,88 (3,80)				
Feminino	Dúvida	17,07 (4,30)	0,017*	0,486	0,114
	Incredulidade	14,49 (3,94)			
	Incerteza	14,20 (4,38)			

Tabela 6: frequência fundamental final para dúvida, incredulidade e certeza

Fonte: Oliveira, 2011, p. 93, com grifos da autora

Dentre os valores máximo e mínimo da frequência fundamental, apenas o mínimo apresentou resultados estatisticamente significativos entre falantes do sexo feminino e do masculino. O padrão de resultados apresentados por Oliveira (2011) indica um gradiente que vai dos menores aos maiores valores de frequência fundamental para dúvida, incredulidade e certeza. A diferença entre esses três sentidos também foi evidenciada pelos valores de intensidade máxima do enunciado, conforme reportado na tabela 7:

SEXO	ATITUDES	MÉDIA	SIGNIFICÂNCIA (P-VALOR)		
			Duv x Incred	Duv x Incert	Incert x Incred
Masculino		Média (DP)			
	Dúvida	65,77 (4,91)	0,000*	0,997	0,000*
	Incredulidade	70,65 (6,66)			
Incerteza	65,76 (7,57)				
Feminino	Dúvida	65,42 (4,91)	0,000*	0,010*	0,000*
	Incredulidade	69,84 (5,44)			
	Incerteza	63, 54 (4,22)			

Tabela 7: intensidade máxima dos enunciados para dúvida, incredulidade e certeza

Fonte: Oliveira, 2011, p. 142, com grifos da autora

Os resultados de duração apresentados por Oliveira (2011) também se mostraram estatisticamente significativos, reforçando a diferença acústica entre dúvida, incredulidade e incerteza, de acordo com o exposto na tabela 8:

SEXO	ATITUDES	MÉDIA	SIGNIFICÂNCIA (P-VALOR)		
		Média (DP)	Duv x Incred	Duv x Incert	Incert x Incred
Masculino	Dúvida	1178 (333)	0,000*	0,000*	0,024*
	Incredulidade	1554 (412)			
	Incerteza	1805 (867)			
Feminino	Dúvida	1118 (194)	0,000*	0,000*	0,054
	Incredulidade	1605 (318)			
	Incerteza	1503 (350)			

Tabela 8: duração dos enunciados para dúvida, incredulidade e certeza
Fonte: Oliveira, 2011, p. 151, com grifos da autora

Apesar da diferença entre os valores de duração, Oliveira (2011) destacou a possível influência da diferença entre o tamanho dos enunciados. Além disso, na realização deles houve ocorrência de pausas principalmente na expressão de incerteza, embora essa variável não tenha sido estatisticamente significativa. Destacamos a caracterização que essa autora fez sobre incerteza e dúvida, gradientes que, de acordo com a hipótese que assumimos baseados nos estudos descritivos sobre *(eu) acho que* (ROSA, 1992; GALVÃO, 1999; FREITAG, 2003; GONÇALVES, 2003), evidenciam diferentes graus de asserção pautados no tipo de experiência do falante.

Também em contexto de laboratório, Antunes, Aubergé e Sasa (2014) elaboraram um experimento com duas etapas, no qual 9 estudantes brasileiros (dois homens e duas mulheres) que estavam na Universidade de Grenoble (França) acreditaram que teriam suas performances em uma entrevista de emprego avaliadas por um robô. Na primeira etapa do experimento, os 9 participantes foram avisados de que antes da entrevista propriamente dita eles responderiam 12 perguntas factuais para que o robô reconhecesse suas vozes. Essas perguntas, que deveriam proporcionar expressões de certeza e incerteza, foram selecionadas por 10 estudantes universitários brasileiros e categorizadas em níveis de dificuldade (4 fáceis, 4 médias e 4 difíceis). Já na segunda etapa do experimento, as pesquisadoras elaboraram outras 8 perguntas, também categorizadas em níveis de dificuldade (fáceis, médias e difíceis), respondidas pelos

informantes enquanto eles eram gravados. Para analisar as propriedades acústicas de certeza e incerteza, Antunes, Aubergé e Sasa (2014) consideraram:

1. Tempo de latência (tempo transcorrido entre o final da pergunta e o início da resposta);
2. Presença de pausas preenchidas;
3. Movimento final da frequência fundamental;
4. Inversão da ordem ou presença de itens lexicais que demarcaram atitude;
5. Duração (taxa de sílabas por segundo, duração das últimas sílabas átona e tônica).

Antunes, Aubergé e Sasa (2014) caracterizaram o tempo de latência (delay) como longo e curto, quantificaram o total de ocorrência (N) de pausas preenchidas e classificaram o movimento da frequência fundamental como ascendente ou descendente. De acordo com as informações expostas no quadro 5, a expressão de certeza (C) apresentou menor tempo de latência: quando certos de algo, os falantes responderam as perguntas que lhes foram feitas mais rapidamente. Na análise dos dados, as autoras incluíram ainda a certeza de não saber (CNS), que apresentou tempo de latência maior. A certeza apresentou movimento de frequência fundamental descendente e menor ocorrência de pausas, ao contrário da incerteza (I):

SEXO	TEMPO DE LATÊNCIA (L – LONGO OU C – CURTO)			P. DE PAUSAS PREENCHIDAS			MOV. FINAL DA F0 (A – ASCENDENTE OU D – DESCENDENTE)		
	C	CNS	I	C	CNS	I	C	CNS	I
FEMININO	29L 84C	24L 14C	34L 13C	2	24	21	54 ^a 59D	10 ^a 28D	32A 15D
MASCULINO	4L 10C	11L 7C	5L 7C	0	10	7	5 ^a 9D	5A 13D	8 ^a 4D
TOTAL (N)	33L 94C (127)	35L 21C (56)	39L 20S (59)	2 (127)	34 (56)	28 (59)	59 ^a 68D (127)	15A 41D (56)	40A 19D (59)

Quadro 5: características gerais de certeza, incerteza e certeza de não saber
Fonte: adaptado de Antunes, Aubergé e Sasa, 2014, p. 113

De acordo com as autoras, as diferentes direções de frequência fundamental, se comparadas com os movimentos ascendentes observados para incerteza e descendentes para certeza e certeza de não saber, são significativamente diferentes ($p < 0,05$, CI 95%). Quanto aos valores de duração, estratificados em relação ao gênero, os resultados das mulheres mostraram que os homens, para expressar incerteza, falaram mais rápido do que as mulheres:

ATITUDE	MASCULINO			FEMININO		
	Sílaba/s	Tônica	Pre-tonica	Sílaba/s	Tônica	Pré-tônica
Certeza	4.48	296.5	179.9	4.58	310.2	192.8
Certeza de não saber	5.01	293.7	158.1	4.65	274.3	183.8
Incerteza	5.99	252.5	158.3	4.64	284.9	163.2

Tabela 9: velocidade de fala e duração final das sílabas, em segundos

Fonte: adaptado de Antunes, Aubergé e Sasa, 2014, p. 113

No estudo de Antunes, Aubergé e Sasa (2014, p. 113) a duração final das sílabas, embora seja elencada pelos estudos de produção como um importante indicativo de incerteza no português brasileiro, não apresentou grande diferença entre os sentidos analisados. Posteriormente, Antunes e Aubergé (2015), a partir das evidências reportadas em outras pesquisas, sugeriram que a expressão de certeza e incerteza apresenta as seguintes características:

movimento melódico final ascendente alinhado à última tônica da sentença na expressão da incerteza ou dúvida (AZEVEDO, 2007; SILVA, 2008; CELESTE, 2010; ANTUNES; AUBERGÉ; SASA, 2014), embora esse movimento não ocorra em todas as sentenças com essas atitudes; ii) um alongamento dos enunciados com incerteza ou dúvida, principalmente devido à presença de pausas (silenciosas ou preenchidas) nesses enunciados, bem como uma velocidade de fala menor na incerteza ou dúvida (SILVA, 2008; OLIVEIRA, 2011; ANTUNES; AUBERGÉ; SASA, 2014); iii) maiores valores de f_0 nos enunciados com incerteza e dúvida, quando comparados à certeza ou ao neutro (SILVA, 2008; MORAES et al., 2010; OLIVEIRA, 2011); iv) maior intensidade nos enunciados com certeza ou neutro, quando comparados à dúvida ou incerteza (MORAES et al., 2010; OLIVEIRA, 2011); mais variação de f_0 (*pitch range*) nos enunciados com dúvida, quando comparados ao neutro (RILLIARD et al., 2012) (ANTUNES, AUBERGÉ, 2015, p. 216).

Fora do contexto de laboratório e em uma situação não-atuada e real, Fernandes e Antunes (2017) estudaram os sentidos de certeza e incerteza, entendidos por elas como atitudes proposicionais, em 80 sentenças (10 em cada cidade, cinco de cada tipo, 5 proferidas por homens e 5 por mulheres) em respostas ao questionário do AliB aplicado em Recife, Belém, Porto Alegre e Curitiba. As respostas analisadas foram especificamente as dos questionários fonético-fonológico e semântico lexical, que possibilitaram respostas curtas, como estas: “*pergunta*: quando se compra uma televisão ela vem dentro de que?” “*resposta*: caixa”; “*pergunta*: como se chama aquela fruta amarela, parecida com a laranja que se descasca com a mão?” “*respostas esperadas*: tangerina, mexerica, etc (FERNANDES, ANTUNES, 2017, p. 60). Como parâmetro de análise, as autoras consideraram valores de frequência fundamental e duração:

1. Tempo de latência;
2. Pontos da frequência fundamental: inicial, final e a média desses valores;
3. Movimento final de F0 da sentença, alinhado à última sílaba tônica do enunciado;
5. Pausas silenciosas e preenchidas;
6. Taxas de articulação e de elocução;
7. Duração da última sílaba tônica e pretônica de cada enunciado;
8. *Self-talk*, ou seja, a repetição total ou parcial da pergunta.

Das variáveis elencadas, somente curva de frequência fundamental, tempo de latência e ocorrência de *self-talk* apresentaram resultados significativos. Os padrões de movimentos finais de frequência fundamental demonstraram que as expressões de certeza e incerteza em entrevistas sociolinguísticas não diferem de situações de testes em laboratório, tanto entre falantes do sexo feminino quanto do sexo masculino:

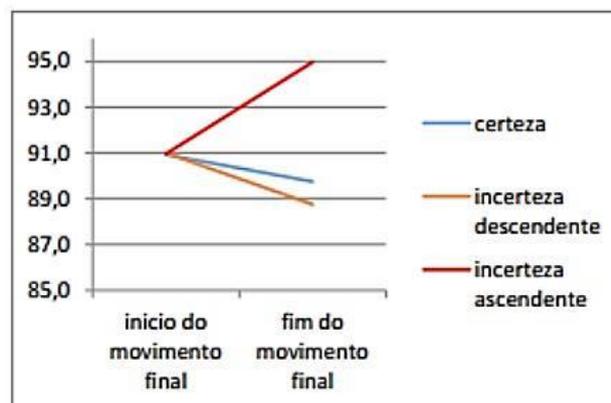


Figura 2: representação do movimento final de f0 nas expressões de certeza e incerteza das mulheres.

Fonte: Fernandes e Antunes, 2017, p. 61

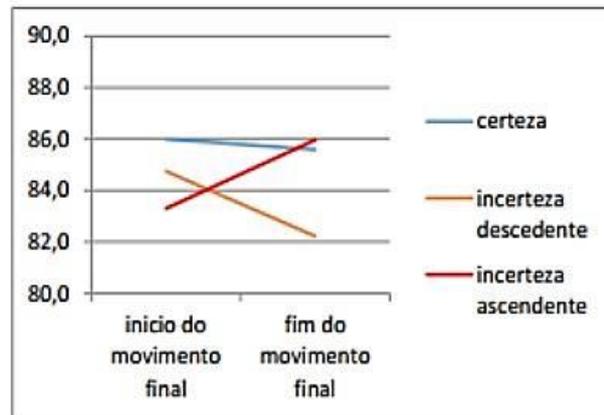


Figura 3: representação do movimento final de f0 nas expressões de certeza e incerteza dos homens.
Fonte: Fernandes e Antunes, 2017, p. 61

Os resultados confirmaram um padrão semelhante ao mencionado por Antunes e Aubergé (2015): os valores de frequência fundamental para incerteza são ascendentes e podem também ser descendentes, enquanto para certeza são sempre descendentes. Além disso, quando certos sobre o que dizem, os informantes se mostram mais ágeis para responder às perguntas feitas pelo documentador. Junto aos valores de frequência fundamental e à diferença do tempo de latência, Fernandes e Antunes (2017) consideraram em sua análise o *self-talk*. Em algumas situações, o informante repetiu a pergunta feita pelo documentador antes de respondê-la. Segundo as autoras, esses recursos são utilizados a fim de construir sentido (FERNANDES; ANTUNES, 2017, p. 61). Conforme demonstrado na tabela 10, pausas e *self-talk* ocorreram somente nos contextos em que houve expressão de incerteza:

	CERTEZA	INCERTEZA
Pausas	0%	17,5%
<i>Self-talk</i>	0%	5%
Fillers	0%	17,5%

Tabela 10: ocorrências de pausas, *self-talk* e fillers em certezas e incertezas
Fonte: adaptado de Fernandes e Antunes, 2017, p. 61

O tempo de latência, a ocorrência de pausas e *self-talk* e *fillers* (falas consigo mesmo e pausas preenchidas, respectivamente) não apresentaram, do ponto de vista estatístico, relevância sobre as taxas de articulação e de elocução dos falantes: a diferença entre média e desvio padrão (entre parênteses) entre as respostas que indicaram certeza e incerteza não foi significativa, de acordo com os dados apresentados na tabela 11:

	MASCULINO		FEMININO	
	Certeza	Incerteza	Certeza	Incerteza
Taxa de articulação	5,31	4,30	4,59	3,73
	(1,08)	(1,91)	(1,30)	(2,68)
Taxa de elocução	5,31	5,18	4,59	4,31
	(1,08)	(1,17)	(1,30)	(0,93)

Tabela 11: média e desvio padrão das taxas de articulação e elocução, em sílabas por segundo
Fonte: adaptado de Fernandes e Antunes, 2017, p. 61

O estudo de Silva (2008) foi executado com orações, gravadas em laboratório, assim como o de Oliveira (2011) e o de Antunes, Aubergé e Sasa (2014). O de Fernandes e Antunes (2017) foi realizado com respostas curtas dos questionários fonético-fonológico e semântico lexical do AliB. Apesar das diferenças entre as amostras e os parâmetros de controle, todos apresentaram padrões acústicos semelhantes para as sentenças que indicam os sentidos certeza, dúvida e incerteza. Parâmetros que, segundo hipotetizamos, atuam na inferência dos mesmos sentidos indicados pela construção modalizadora parentética epistêmica (*eu acho que*).

3.2 CERTEZA, DÚVIDA E INCERTEZA EM ESTUDOS DE PROSÓDIA GESTUAL

Do ponto de vista linguístico, conforme Rosa (1992), Galvão (1999), Gonçalves (2003) e Freitag (2003) apontaram, (*eu acho que* indica certeza, dúvida e incerteza sobre um conteúdo proposicional. Enquanto falantes nativos, para inferirmos qualquer um desses sentidos utilizamos elementos que estão além do código linguístico, mas que ainda assim são importantes para a interação. Os elementos que não fazem parte do código linguístico constituem a paralinguagem, cuja definição é problemática porque não há consenso sobre quais os seus componentes e limites (SCHÖTZ, 2002). Alguns estudos (LAVER, 1980; ROACH ET AL, 1998) consideram como parte da paralinguagem tudo que acontece durante a interação, mas que não é marcado pelo código: ruídos (espirro, tosse, pigarro), aspectos relativos à idade e ao sexo biológico do falante, percebidos pelo timbre da voz. Outro tipo de abordagem (TRAUNMÜLLER, 2000) considera como parte da paralinguagem tudo que é útil para a inferência de sentidos: pausas e hesitações na fala, sorrisos e movimentos corporais, além de medidas de frequência fundamental. É esta perspectiva que assumimos, com o objetivo de investigar se, na amostra Deslocamentos 2020, parâmetros acústicos (medidas de frequência

fundamental, intensidade, duração, ocorrência de pausas silenciosas e preenchidas), junto às expressões faciais, que entendemos como pistas paralinguísticas, contribuem para a inferência dos sentidos de certeza, dúvida e incerteza indicados por *(eu) acho que*.

A importância dos elementos paralinguísticos tem sido focalizada especialmente na área da computação para síntese e reconhecimento de voz (VIANA, 2001; MARTINS, 2001). Embora seja comum utilizar prosódia para fazer referência aos parâmetros acústicos, esse termo não se refere somente às medidas de frequência fundamental, intensidade, duração, pausas silenciosas e preenchidas. Ela “pode ser definida como toda a gama de recursos que determinam não somente *o que* as pessoas dizem, mas *como* elas dizem.”⁴ (SWERTS ET AL, 2003, p. 25, grifos do autor). Nesse caso, tem-se a prosódia gestual, que inclui movimentos corporais que contribuem para a expressão de sentidos não captados pelo léxico. Para Swerts et al (2003), a codificação dos sentidos que não são captados pelo léxico é importante especialmente para máquinas que usam tecnologia de reconhecimento de fala, imagens e gestos. No presente estudo, assumimos que a importância da contração dos músculos do rosto está relacionada às inferências que eles suscitam na interação e, por isso, devem ser incluídos na descrição linguística, a fim de proporcionar uma compreensão acurada do processo comunicativo.

A fim de observar a relação entre respostas que indicam certeza, incerteza e gestos corporais, Swerts et al (2003) executou o seguinte experimento: 20 falantes nativos de holandês foram convidados a responder perguntas sobre literatura, esportes, história, dentre outros temas, organizadas de acordo com nível de dificuldade e distribuídas aleatoriamente. Enquanto respondiam as perguntas, os informantes foram filmados. As seguintes pistas gestuais foram elencadas e descritas perceptualmente, por falantes que não participaram do experimento:

1. Desvio do olhar: olhar direcionado para baixo enquanto o informante responde o que foi perguntado;
2. Olhar fixo: olhar direcionado para frente enquanto o informante responde o que foi perguntado;
3. *Funny face*: expressão facial marcada, caracterizada pela compressão das sobrancelhas e da linha da boca, movimentos que se opõem à expressão neutra, na qual os músculos da face estão em repouso.

⁴ In a broad sense, prosody can be defined as the whole amount of features that do not determine *what* people are saying, but rather *how* they are saying it.”

Os resultados apontaram que, quando incertos sobre uma resposta, os falantes desviaram o olhar, comprimiram a boca e demonstraram *funny face* (figura 7). Quando certos, mantiveram expressão de neutralidade na qual os músculos do rosto se encontram em estado de repouso. Assim como os falantes diferenciam certeza e incerteza por meio de propriedades acústicas, do ponto de vista gestual esses mesmos sentidos são diferenciados pelos movimentos do rosto, exemplificados na figura 4:



Figura 4: registro do falante ao indicar incerteza (*funny face*)
Fonte: Swerts et al, 2003, p. 26

Como parte do mesmo estudo, 60 outros participantes foram convidados a julgar se, ao enunciar cada resposta, aquele falante estava ou não certo do que disse. Os resultados comprovaram a importância das pistas acústicas e visuais: os estímulos em que havia ocorrência de pausas preenchidas, desvio de olhar e *funny face* foram reconhecidos como expressão de incerteza, o que sugere que além de parâmetros acústicos, pistas gestuais são importantes para a inferência sentidos. Além de diferenciar sentidos, os gestos corporais também caracterizam grupos de falantes e sentidos pragmáticos, o que sugere sistematicidade entre esses movimentos.

Valendo-se do *the interactive atlas of the prosody of portuguese webplataform*, banco de dados cujo objetivo é documentar variação acústica e gestual a partir de uma notação baseada no FACS – *Facial Action Coding System*⁵, Cruz, Swerts e Frota (2017) realizaram um experimento com 40 falantes das variedades padrão (SEP) e a de Ponta Delgada (PtD) (20 de cada) do português europeu. O experimento, com três condições: apenas áudio, apenas vídeo,

⁵ A notação FACS cataloga os movimentos dos músculos do rosto e caracteriza as expressões faciais de raiva, nojo, medo, alegria, tristeza, surpresa. Para mais informações: <https://www.paulekman.com/facial-action-coding-system/>

áudio mais vídeo (com estímulos originais e manipulados), durou em média 15 minutos. Os 40 falantes, sentados em frente a um computador, tiveram como tarefa classificar os estímulos em mais ou menos declarativos ou interrogativos a partir de uma escala Likert de 5 pontos. O objetivo era investigar o papel da entonação (altura do *pitch*: H – high; L – low, alto e baixo, respectivamente) e de pistas visuais, especificamente o movimento das sobrancelhas, na percepção de sentenças neutras, afirmativas e negativas:

AFIRMAÇÕES NEUTRAS			NEUTRAS AFIRMATIVAS E NEGATIVAS	
Variedade	Tonal	Visual	Tonal	Visual
SEP	H+L*	Cabeça levantada-baixa	H+L*	Cabeça levantada-baixa + sobrancelhas arqueadas
	L%	Posição neutra	LH%	Posição neutra
PtD	(H+)L*	Cabeça levantada-baixa	(H+)L*	Cabeça levantada-baixa + sobrancelhas arqueadas
	L%	Posição neutra	L%	Posição neutra

Tabela 12: características acústicas e gestuais do SEP e PtD
Fonte: adaptado de Cruz, Swerts e Frota, 2017, p. 2

Do ponto de vista gestual, há uma diferença entre as variedades SEP e PtD: os falantes de PtD arqueiam as sobrancelhas ao proferirem os três tipos de sentença. Por isso, Cruz, Swerts e Frota (2017) levantaram sete hipóteses, das quais duas foram descartadas e duas apresentamos aqui, devido à proximidade com os objetivos do presente trabalho:

1. Falantes da SEP vão utilizar pistas visuais para identificar sentenças neutras afirmativas e negativas produzidas pelos falantes da PtD, quando essas pistas estiverem disponíveis;
2. As mesmas sentenças vão ser identificadas como afirmações neutras quando as pistas visuais não estiverem disponíveis.

Os resultados mostraram que há uma hierarquia entre as pistas utilizadas pelos falantes para diferenciar os três tipos de sentenças. Cruz, Swerts e Frota (2017) concluíram que a

percepção dos falantes da SEP e PtD depende do tipo de sentença, da variedade de fala e do tipo de condição (apenas áudio, apenas vídeo, áudio e vídeo). De modo geral, os falantes das duas variedades se basearam, principalmente, na entonação para diferenciar os três tipos de sentenças, ao contrário do que foi hipotetizado (CRUZ; SWERTS; FROTA, 2017, p. 2). É importante destacar que na condição apenas vídeo, os falantes flutuaram no julgamento. Para os autores, esse resultado sugere que apesar da diferença entre as duas variedades sob análise, sozinhas, as pistas visuais não são suficientes para diferenciar afirmações neutras, neutras afirmativas e negativas. Sua relevância está no funcionamento coordenado.

No Brasil, a partir do experimento executado por Swerts et al (2003), Antunes, Aubergué e Sasa (2014) investigaram a relação entre prosódia acústica, movimentos corporais e a expressão de certeza e incerteza, entendidas por esses autores como atitudes. Para isso, utilizaram respostas espontâneas de 9 falantes dadas a um robô, gravadas a fim de verificar movimentos da sobrancelha, dos olhos e a ocorrência de *funny face*, além da direção do olhar (direto ou desviado) e movimentos executados com a cabeça (afirmativos e negativos). Os resultados sobre a expressão de certeza e incerteza no português brasileiro foram semelhantes aos do holandês: grande diferença no percentual de movimento dos olhos, boca e cabeça quando os falantes demonstram esses sentidos, entendidos pelos autores como atitudes:

ATITUDE	MOV. DAS SOBRANCELHAS	DESVIO DO OLHAR	MOV. DA BOCA	MOV. DA CABEÇA
Certeza (127)	7 (5.5%)	26 (20.5%)	2 (1.6%)	4 (3.1%)
Certeza ou não (56)	16 (28.6%)	29 (51.8%)	20 (35.7%)	17 (30.4%)
Incerteza (59)	21 (35.6%)	52 (88.1%)	25 (42.4%)	25 (42.4%)

Tabela 13: movimentos observados em vídeos
Fonte: Adaptado de Antunes, Aubergé, Sasa, 2014, p. 127

3.3 EXPRESSÕES FACIAIS, CERTEZA, DÚVIDA E INCERTEZA

Os movimentos dos músculos do rosto que demonstram os sentidos de certeza, dúvida e incerteza no holandês, no português europeu (variedades padrão e de ponta delgada) e no português brasileiro podem são expressões faciais que podem indicar respostas emocionais.

Segundo Ekman (2000), as emoções indicam respostas a estímulos externos e também contribuem para a manutenção das relações interpessoais.

Na psicologia, o debate sobre expressões faciais e emoções envolve sua definição, o que elas indicam e qual sua finalidade tanto do ponto de vista individual quanto social. Essa discussão tem sido feita à luz de diferentes correntes e abordagens, dentre as quais destacamos a psicoevolucionista e a cognitivista. Para a primeira, as emoções são reflexo do processo evolutivo pelo qual todas as espécies passaram. O desenvolvimento do medo na espécie humana, por exemplo, permitiu sua sobrevivência e perpetuação no planeta, o que sugere que as emoções possuem caráter básico, tendo evoluído

de um conjunto finito de estados emocionais, sendo que cada um deles possuía sua funcionalidade adaptativa e expressão típica (Ekman, 2003; Gazzaniga, Irvy e Mangun, 2006; Lundqvist e Ohman, 2005; Markham e Wang, 1996; Plutchick, 2002)” (MIGUEL, 2015, p. 154).

O caráter básico das emoções tem gerado inúmeras discussões nas áreas por elas interessadas. A nomenclatura *básica* pode fazer referência a emoções contrárias (como, por exemplo, alegria e tristeza), a união de duas ou mais emoções para a formação de outra mais complexa ou aquelas presentes em todas as espécies como reflexo do processo evolutivo (EKMAN, 2000, p. 45). Outro ponto de discussão é a definição de quantas e quais são as emoções básicas:

REFERÊNCIA	TIPOLOGIA
Ekman; Friesen (1975)	Raiva, nojo, medo, alegria, tristeza, surpresa
McDougall (1926)	Raiva, nojo, alegria, medo, sujeição, concurso, emoção, maravilha
Plutchick (1982)	Aceitação, raiva, antecipação, nojo, alegria, medo, tristeza, surpresa
Tomkins (1984)	Raiva, interesse, desprezo, repulsa, angústia, medo, alegria, vergonha, surpresa

Quadro 6: tipologia das emoções básicas

Fonte: elaborada pela autora

Há múltiplas tipologias acerca de quantas e quais são as emoções comuns a toda a espécie humana. As destacadas no quadro 6 foram elencadas como respostas rápidas a estímulos

externos pelos pesquisadores citados. Dentre as categorizações descritas acima, destacamos a proposta por Ekman e Friesen (1975), que exerce grande influência nos estudos sobre emoção. Nela, as emoções básicas, salvo particularidades socioculturais, são demonstradas por meio de um padrão fisiológico de contração dos músculos da face:

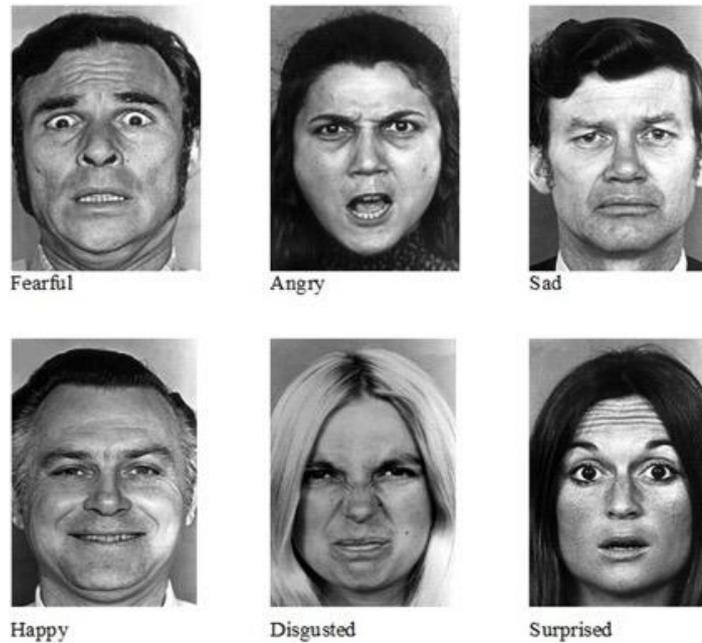


Figura 5: adaptação das emoções básicas propostas por Ekman e Friesen
Fonte: Ekman; Friesen, 1975, p. 173-202

A expressão facial de medo (*fearful*) é caracterizada pela abertura da mandíbula e das pálpebras superiores, ocasionando maior abertura dos olhos. Essa expressão constitui-se como uma resposta a ameaças físicas ou psicológicas (EKMAN; CORDARO, 2011, p. 365). A partir do momento em que o medo é registrado pelo cérebro, o indivíduo tenta se proteger de qualquer tipo de perigo. A raiva (*angry*) é uma reação hostil a algo que causa frustração. Ela pode propiciar ataque a algo ou alguém, a fim de dirimir a insatisfação sentida pelo indivíduo. Essa expressão é caracterizada pela tensão das sobrancelhas e abertura da boca. A tristeza (*sad*) é desencadeada pela sensação de perda de algo valioso, sendo representada pelo abaixamento dos lábios, elevação das bochechas e o franzir das sobrancelhas (EKMAN; FRIESEN, 2003). Já na expressão de felicidade (*happy*), desencadeada pela satisfação por algo, há a contração do músculo zigomático maior, que vai das bochechas até o canto dos lábios, formando um sorriso (EKMAN; FRIESEN, 2003, p. 221). O nojo (*disgusted*), por sua vez, é uma reação a algo repulsivo, desagradável. Na sua expressão, o nariz e a boca são franzidos, enquanto as sobrancelhas são rebaixadas. A surpresa (*surprised*) é a expressão mais rápida de todas (EKMAN; CORDARO, 2011, p. 365), pois é gerada bruscamente, em uma situação inesperada.

A expressão facial de surpresa é semelhante à de medo, caracterizada pelo levantamento das pálpebras inferiores e por uma leve abertura da mandíbula. A descrição das características de cada uma das seis expressões demonstradas na figura 5 originou um manual que descreve *Facial Action Units (FACS)*, movimentos das partes superior e inferior do rosto que podem indicar medo, raiva, tristeza, felicidade, nojo e surpresa:

Unidades de Ação da Face Superior					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
 Sobrancelha Interna Levantada	 Sobrancelha Externa Levantada	 Sobrancelhas Baixas	 Pálpebra Superior Levantada	 Bochechas Levantadas	 Pálpebras Apertadas
*AU 41	*AU 42	*AU 43	AU 44	AU 45	AU 46
 Abaixamento das Pálpebras	 Contração Retinal	 Olhos Fechados	 Olhos Semicerrados	 Piscada dos olhos	 Piscada de um olho
Unidades de Ação da Face Inferior					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
 Nariz Enrugado	 Lábio Superior Levantado	 Nasolabial Aprofundado	 Cantos dos Lábios Puxados	 Bochecha Inchada	 Fazendo Covinhas
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
 Cantos dos Lábios Pressionados	 Lábio Inferiore Pressionado	 Queixo Levantado	 Lábios Enrugados	 Lábios Esticados	 Lábios Afunilados
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
 Lábios Endurecidos	 Lábios Pressionados	 Lábios Separados	 Mandíbula Caída	 Apertando a Boca	 Sucção dos Lábios

Figura 6: facial action units (movimentos dos músculos do rosto)

Fonte: adaptado de <https://infograph.venngage.com/p/8769/microexpressions>

É importante destacarmos que a contração da linha das sobrancelhas e boca, por exemplo, não necessariamente indica tristeza. A figura 4 é um exemplo de *funny face* que, para Swerts et al (2003), dá pistas de incerteza. Os movimentos ali demonstrados são descritos pelas AUs 4, 15 e 17, expostas na figura 6. Naquele contexto específico em que os falantes holandeses estavam inseridos, esses movimentos indicaram incerteza, não tristeza.

O manual publicado por Ekman e Friesen (1975) teve como objetivo treinar indivíduos para o reconhecimento e análise manual de expressões faciais. Todavia, com o avanço da

tecnologia é possível executar essas tarefas por meio de *scripts* em linguagem de programação e softwares computacionais que, a partir de *datasets* que categorizam cada expressão facial, captam a face dos falantes, mapeiam os movimentos dos seus músculos e indicam a intensidade das expressões demonstradas. O reconhecimento e análise das expressões faciais é relevante tanto para estudos sobre emoções humanas e interação entre sujeitos quanto para o desenvolvimento de produtos tecnológicos que funcionam por meio do reconhecimento da face. Por isso, há inúmeros *datasets* que permitem seu reconhecimento e classificação, feitas a partir do rótulo das expressões faciais universais, que variam a depender da abordagem teórica assumida. Tejada et al (2021)⁶ destacaram que grande parte dos *datasets* de expressões faciais disponíveis no mercado são constituídos por faces americanas e europeias, o que pode prejudicar a classificação de rostos de outros povos. Por isso, esses autores desenvolveram um estudo cujo objetivo era construir um banco de dados de expressões brasileiras e colombianas.

Para desenvolver o banco de dados de expressões faciais, Tejada et al (2021) realizaram um estudo experimental com 110 pessoas (não atores, 50 em Bogotá, capital da Colômbia, e 60 em Aracaju, Brasil), 54 mulheres e 56 homens que tiveram como objetivo replicar as expressões de raiva, nojo, medo, felicidade, tristeza, surpresa e deboche apresentadas em um computador no laboratório, enquanto eram gravados por uma webcam Logitech C90 HD pro que gerou 30 frames por segundo⁷. Em relação ao deboche, resposta negativa caracterizada por um sorriso sutil que indica julgamento negativo ou discordância em relação a algo, os autores destacaram que essa expressão aparenta demonstrar particularidades específicas entre falantes brasileiros, por isso ela foi incluída.

Ao todo, 662,172 frames (350x350, em escala cinza) foram gerados pela câmera e posteriormente processados por *scripts* em linguagem Python que reconheceram os movimentos dos músculos dos rostos dos participantes a partir de 68 pontos distribuídos nas

⁶ Hector Julian Tejada Herrera, Raquel Meister Ko. Freitag, Paloma Batista Cardoso, Bruno Felipe Marques Pinheiro, Lucas Santos Silva e Vitor Rene Andrade Souza elaboraram um banco de dados de expressões faciais colombianas e brasileiras, que utilizamos na análise dos vídeos da amostra Deslocamentos 2020. Esse banco foi elaborado por causa da limitação metodológica de um estudo que investigou se expressões faciais dão pistas do julgamento negativo do rotacismo (FREITAG ET AL, 2020). O *script* utilizado, baseado em faces europeias, reconheceu a expressão de alegria, o que os autores julgaram incoerente. Eles assumiram que era mais provável que os falantes brasileiros demonstrassem deboche. Por isso, essa expressão foi categorizada no banco de dados de expressões colombianas e brasileiras.

⁷ O banco de imagens elaborado por esses autores bem como os *scripts* para reconhecimento e classificação das expressões faciais estão disponíveis em: https://github.com/julian-tejada/EmotionRecognition/blob/master/Script_EmotionClassification.py e podem ser baixados gratuitamente. O arquivo que reporta todos os procedimentos e resultados está em tramitação para publicação no corrente ano (2021).

linhas da sobrancelha, nariz e boca. Os autores utilizaram o banco de dados KDF para classificar as expressões mais comuns: raiva, nojo, medo, surpresa, felicidade, tristeza e surpresa. Depois, os resultados dessa classificação foram validados por 10 professores universitários que tiveram como tarefa, em uma *survey* online, avaliar cada expressão. Deboche não estava na classificação inicial. Para categorizar essa expressão, os juizes tiveram como critério a maior probabilidade (acima de 0.7) das expressões de nojo e surpresa. As imagens categorizadas pelos 10 juizes que atenderam ao critério de probabilidade foram selecionadas para compor o banco de dados, que passou por um novo processo de validação pelos mesmos professores universitários. Do total, apenas uma imagem de cada participante para cada expressão foi selecionada:

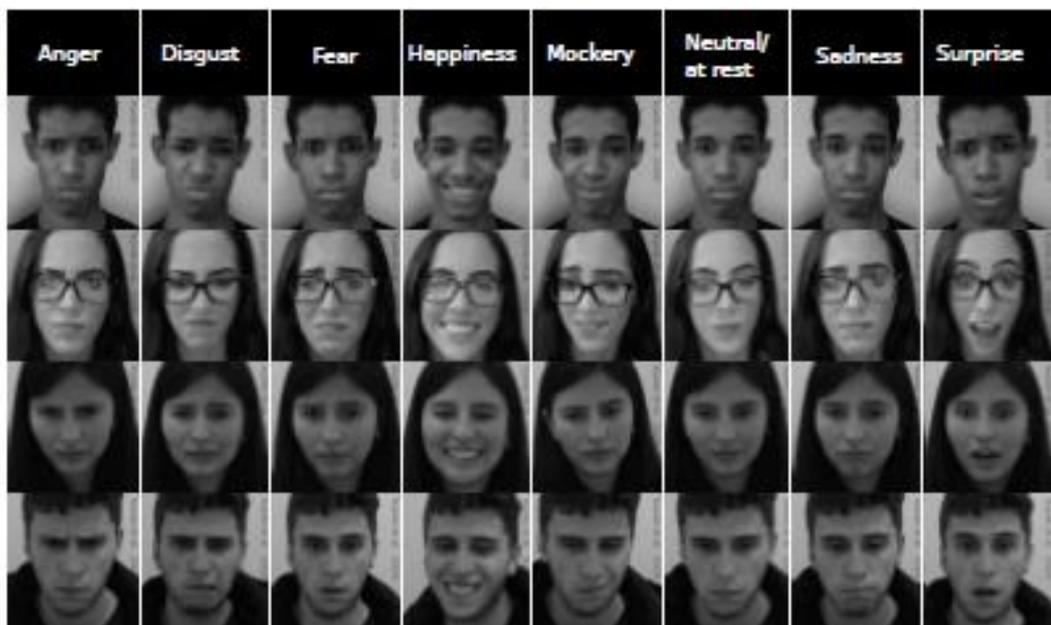


Figura 7: Dataset brasileiro/colombiano
Fonte: Tejada et al (2021)

Por fim, o último banco de dados avaliado pelos juizes foi comparado (brasileiros vs. colombianos) com outro banco, o KDEF. Para validar os dados finais, o banco de imagens brasileiro e o colombiano foi dividido em duas partes, com 75% e 25% das imagens para executar um procedimento de *learning machine* e observar a matriz de confusão que possibilitou a validação de cada uma das expressões de raiva, nojo, medo, felicidade, deboche, neutra, tristeza e surpresa. A amostra de falantes brasileiros e colombianos, composta por mulheres (M) e homens (H) teve a seguinte classificação:

EMOÇÃO	SUBAMOSTRA BRASILEIRA	SUBAMOSTRA COLOMBIANA
Raiva	37 (21M, 16H)	36 (15M, 21H)
Deboche	55 (32M, 25H)	44 (14M, 25H)
Nojo	42 (24M, 18H)	45 (16M, 29H)
Medo	34 (21M, 13H)	31 (12M, 19H)
Felicidade	57 (35M, 22H)	47 (21M, 26H)
Tristeza	39 (21M, 18H)	30 (10M, 20H)
Surpresa	50 (28M, 22H)	42 (19M, 23H)
Neutra/Descanso	44 (25M, 18H)	34 (13M, 21H)

Quadro 7: número de imagens que compõem os *datasets* brasileiro e colombiano

Fonte: Tejada et al, 2021

As expressões faciais indicam processos cognitivos em níveis distintos. São respostas multi-níveis (TEASDALE, 1999) que na interação, integradas a elementos acústicos podem também contribuir para a inferência de sentidos. Os trabalhos realizados por Swerts et al (2003), Antunes; Aubergué e Sasa (2014); Cruz, Swerts e Frota (2017) sinalizaram a importância dos movimentos dos músculos faciais para a inferência de sentidos. Valendo-nos das evidências dos trabalhos que descrevemos e de um banco de dados que atende às especificidades antropomórficas dos falantes que constituem a amostra Deslocamentos 2020, objetivamos contribuir com a descrição da inferência dos sentidos indicados por *(eu) acho que* e com o desenvolvimento do estudo das relações entre língua e movimentos dos músculos faciais. A partir do exposto, tomamos como hipótese que os sentidos de certeza, dúvida e incerteza são caracterizados por parâmetros estruturais, subjetivos e também acústicos e gestuais. Os parâmetros gestuais, especificamente as expressões faciais, atuam como pistas paralinguísticas que, na interação, contribuem para a inferência dos sentidos indicados por *(eu) acho que*.

3.4 PROPOSTA DE ANÁLISE

Certeza, dúvida e incerteza, em estudos de descrição linguística e acústica, são definidos como atitudes ou afetos sociais, conforme mencionado nos estudos que reportamos em 3.2 e 3.3. Neste trabalho, não temos como objetivo propor nem discutir as distinções entre esses rótulos. Por isso, atemo-nos à noção de sentidos, tradicionalmente descritos por meio de

sentenças. Contudo, nesta pesquisa, objetivamos contribuir com os estudos que descreveram os usos de *(eu) acho que*.

A partir das evidências mobilizadas, tomamos como hipótese que certeza, dúvida e incerteza quando indicados pela construção destacada são inferidos por variáveis linguísticas (tipo de ocorrência, escopo, presença de modalizador, tópico discursivo, experiência do falante, polaridade); acústicas (média da frequência fundamental, intensidade, duração, pausas silenciosas e preenchidas); e paralinguísticas (expressões faciais). Desse modo, propomos a testagem da associação entre os parâmetros elencados para observarmos se eles dão pistas dos sentidos indicados por *(eu) acho que* nas entrevistas sociolinguísticas que compõem a amostra Deslocamentos 2020.

4. MÉTODO

Neste capítulo, apresentamos os procedimentos metodológicos para a descrição da inferência dos sentidos de certeza, dúvida e incerteza indicados por (*eu acho que*) por meio de variáveis linguísticas, acústicas e paralinguísticas (expressões faciais). Em 4.1, descrevemos como a amostra analisada foi constituída, caracterizamos o perfil dos informantes e dos pesquisadores entrevistadores. Posteriormente, em 4.2 apresentamos os instrumentos de coleta, as variáveis que controlamos para fins de análise e, por último, descrevemos o tratamento estatístico dos dados coletados.

4.1 CONSTITUIÇÃO E TRATAMENTO DA AMOSTRA

Os dados utilizados para investigarmos a inferência dos sentidos indicados (*eu acho que*) foram retirados de 30 entrevistas sociolinguísticas coletadas entre 2019 e 2020⁸, na cabine acústica do Condomínio de Laboratórios Multiusuários de Informática e Documentação – LAMID, na Universidade Federal de Sergipe, *campus* Professor José Aloísio de Campos:

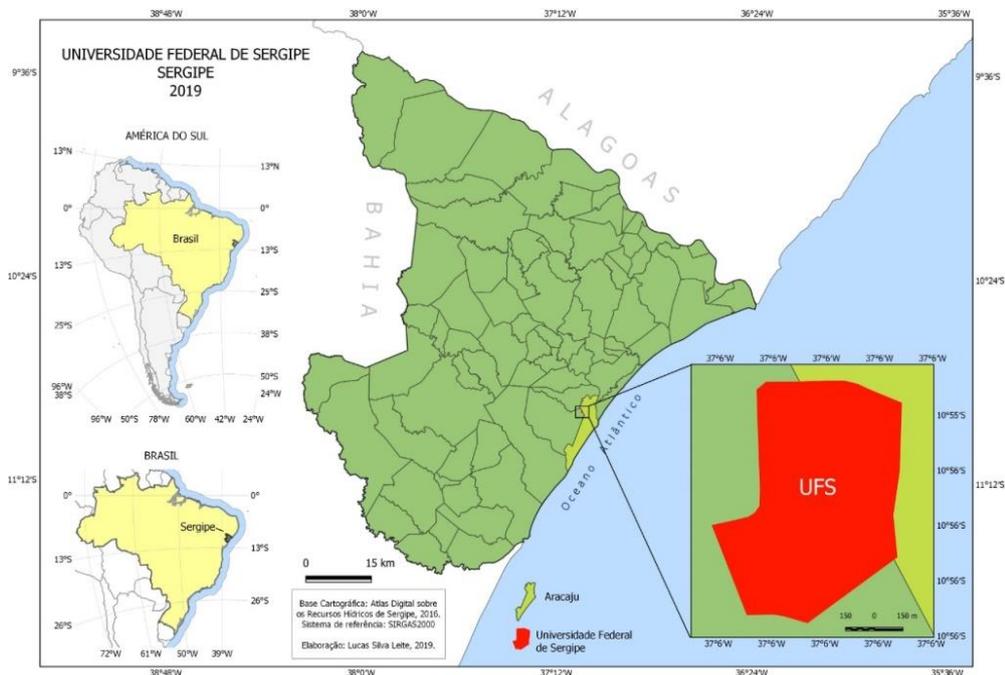


Figura 8: Localização da Universidade Federal de Sergipe

Fonte: Pinheiro, 2021, p. 42

⁸ Nesse período, realizamos 100 entrevistas sociolinguísticas. 80 foram transcritas no software ELAN para uso dos pesquisadores do Grupo de Estudos em Linguagem, Interação e Sociedade. Desse total, 30 foram gravadas em áudio e vídeo para subsidiar a análise das expressões faciais. Nem todas as entrevistas da amostra Deslocamentos 2020 foram gravadas em vídeo pois essa foi a primeira coleta de dados de imagem. Essa decisão foi tomada para evitar possíveis problemas quanto à viabilidade da amostra.

Essas 30 entrevistas, gravadas em áudio e vídeo, fazem parte da amostra Deslocamentos 2020, constituída por 100 entrevistas ao todo. A documentação dos dados linguísticos e das expressões faciais foi feita em áudio, com o gravador de mesa Maratntz, modelo PMD661, e vídeo, com duas webcams Logitech C920 HD. Entrevistado e entrevistador foram posicionados frente a frente, nos limites de uma cabine acústica revestida por painéis brancos e iluminada verticalmente. Os arquivos de áudio foram salvos no formato *.wave* para fins de transcrição no software ELAN (BRUGMAN, H.; RUSSEL, 2004), cuja escolha foi motivada pela sua funcionalidade e pelas vantagens que oferece, como

a sincronização entre o arquivo de mídia e a transcrição/anotação, o que facilita enormemente a análise linguística dos dados (por exemplo, para codificação de variantes de variáveis fonéticas); a possibilidade de criação de múltiplas trilhas, que propicia não só a separação da fala de diferentes participantes, mas também a anotação detalhada de outros aspectos linguísticos e contextuais, bem como a representação de ações simultâneas (por exemplo, sobreposição de vozes, ações gestuais concomitantes às verbais); ferramentas mais sofisticadas de buscas dentro de um corpus (por exemplo, para encontrar todas as variantes de uma variável). (OUSHIRO, 2014, p. 46)

A transcrição das entrevistas, conforme argumenta Oushiro (2014), teve como finalidade permitir a localização de todas as realizações de *(eu) acho que* para posterior tabulação no excel, de acordo com as variáveis que controlamos.

Search eaf files

Annotations

Search [Define search domain](#)

regular expression case sensitive [Export](#)

1-50 31-63

Nr	File	Tier	Before	Annotation	After	Parent
1	ISA2FF	DOCBP	é mulher brasil...	eu acho que eu nem considero mais assédio né	ai já passa por ...	
2	ISA2FF	DOCBP	ah entendi ente...	mas eu acho que sua que sua atração é só por homens mesmo e justamente porque	eh	
3	ISA2FF	ISA2FF	não	porque eu acho que	quando a gente...	
4	ISA2FF	ISA2FF	ai não sei eu tip...	então eu acho que é só a questão mesmo de alternativas do que fazer porque	basicamente se...	
5	ISA2FF	ISA2FF	o Rio eu não g...	porque eu acho que o Rio só é bom em partes pra quem tem muito dinheiro pra gastar	tipo	
6	ISA2FF	ISA2FF	tipo	e a natureza também eu acho que o que deixa o Rio bonito é a natureza mas eu não	achei legal né ...	
7	ISA2FF	ISA2FF	com os pais vo...	eu acho que	é mais divertido...	
8	ISA2FF	ISA2FF	((RUÍDO))	porque eu não acho que seja uma coisa excludente eu não acho que exclua quem	quem não vive ...	
9	ISA2FF	ISA2FF	pra quem não t...	então eu acho que é muito interessante isso	é porque aí já v...	
10	ISA2FF	ISA2FF	é realmente inv...	mas eu acho que nunca é demais né ((RISOS)) investir na educação	porque se você...	
11	ISA2FF	ISA2FF	e por exemplo ...	eu acho que a UFS ela realmente tenta enten- tenta suprir	essa necessida...	
12	ISA2FF	ISA2FF	eu acho que a ...	essa necessidade só que também eu acho que vai muito da questão da falta de cuidado	então mas assim	
13	ISA2FF	ISA2FF	conseguiu	atender minhas expectativas eu não acho que foi	uma experiênci...	
14	ISA2FF	ISA2FF	tipo assim é mu...	eh eu acho que eu vou tentar entrar na indústria	eu não	
15	ISA2FF	ISA2FF	eu não	até agora eu ainda não tive a pretensão de fazer mestrado e nem voltar pra cá mas eu	não	
16	ISA2FF	ISA2FF	então hoje em ...	eu acho que tem realmente casos que	quando a pess...	
17	ISA2FF	ISA2FF	não vai dar res...	eu acho que	passou a ser	
18	ISA2FF	ISA2FF	passou a ser	nem nem é a questão de ser modinha eu acho que passou	a ser usado de ...	
19	ISA2FF	ISA2FF	ela estar usand...	eu não me sinto empoderada por ela eu acho que esse não é um discurso que você	que deve ser fe...	
20	ISA2FF	ISA2FF	se portar da for...	muitas vezes eu não acho que é um que é muito um discurso de empoderamento	empoderamento	
21	ISA2FF	ISA2FF	o meu cabelo s...	meu cabelo crespo e comecei a usa-lo dessa forma eu acho que	começou a ser	
22	ISA2FF	ISA2FF	começou a ser	a ser um termo usado em situações muito erradas sabe eu acho que muitas situações	tipo você não s...	
23	ISA2FF	ISA2FF	tão	eu acho que é muito pior a questão pras mulheres	((RISOS))	
24	ISA2FF	ISA2FF	eu acho	eu acho que as motivações	da agressão sã...	
25	ISA2FF	ISA2FF	que são agredi...	eu acho que é a mes a questão do feminismo	começou	

Figura 9: localização de *(eu) acho que* no ELAN
Fonte: elaborada pela autora

Depois, as sentenças em que houve ocorrência dessa construção foram recortadas no software Audacity a fim de permitir a transcrição dos excertos no PRAAT (BOERSMA; WEENINK, 2017), o que possibilitou a extração dos parâmetros de frequência fundamental, intensidade, e duração de *(eu) acho que*, das pausas silenciosas e preenchidas com o script *analyse tyer* (HIRST, 2012).

A documentação em vídeo⁹ teve como objetivo registrar a face dos informantes para investigar a influência, na entrevista sociolinguística, dos movimentos dos músculos do rosto na inferência de sentidos. Segundo Ekman (2000), salvo particularidades culturais, as emoções de alegria, tristeza, surpresa, medo, raiva e nojo são expressas pelos movimentos dos músculos faciais, e são importantes para as relações interpessoais. Considerando a proposta deste trabalho, pautada em uma análise que tem como objetivo observar associação entre variáveis linguísticas, acústicas e expressões faciais, no ato de transcrição das entrevistas alinhamos os arquivos de áudio e vídeo do informante, conforme demonstrado na figura 10:

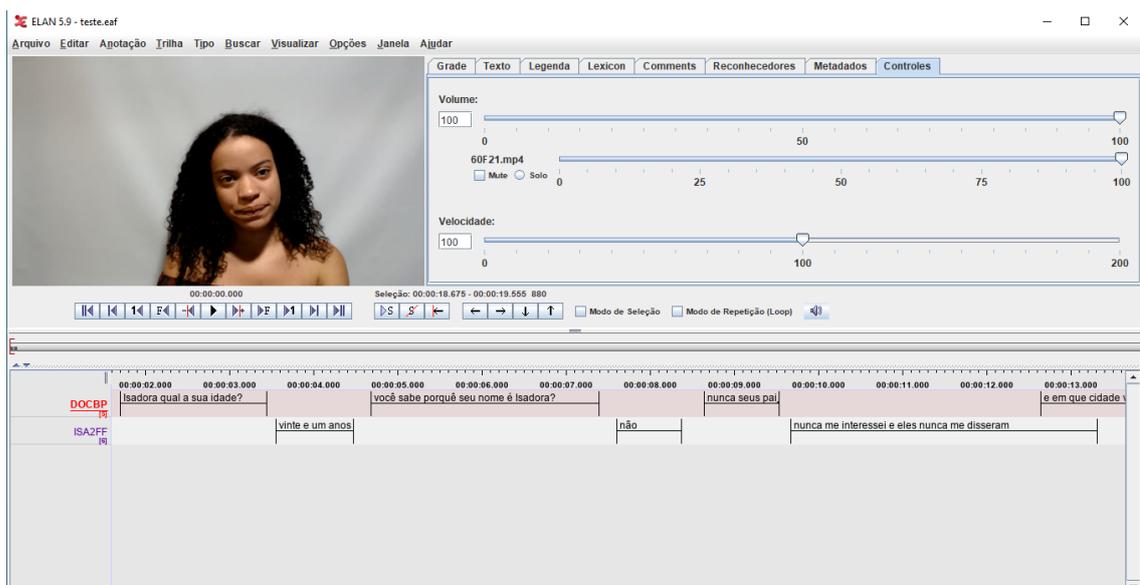


Figura 10: alinhamento entre áudio e vídeo no ELAN
Fonte: Pinheiro, 2021, p. 45

Os arquivos das entrevistas que constituem a amostra Deslocamentos 2020 foram nomeados de acordo com o perfil de constituição da amostra, considerando as três primeiras letras dos nomes dos informantes, para manter o sigilo de suas identidades, sua região de

⁹ Os procedimentos para o tratamento dos vídeos das entrevistas, captação dos rostos e expressões faciais é semelhante ao utilizado por Pinheiro (2021), que analisou o mesmo corpus a fim de observar a influência das expressões faciais na inferência dos sentidos indicados por diminutivos.

origem¹⁰, gênero (F para feminino, M para masculino)¹¹ e período de curso (I para início, F para fim). O áudio da informante na figura 10, por exemplo, foi salvo como ISA2FF, de acordo com as diretrizes adotadas pelo Grupo de Estudos em Linguagem, Interação e Sociedade – GELINS. Neste trabalho, assumimos os pressupostos teórico-metodológicos adotados para a constituição do banco de dados Falares Sergipanos, cujo objetivo consiste em

subsidiar a investigação de variedades linguísticas do português em seus diferentes níveis (do morfofonológico ao discursivo) e com diferentes propósitos (dos descritivos aos moldes sociolinguísticos às aplicações educacionais, subsidiando programas de ensino de língua materna). (FREITAG, 2013, p. 156-157).

A metodologia adotada para a constituição do banco de dados Falares Sergipanos é baseada nos princípios da sociolinguística variacionista. Nessa perspectiva, a língua é um sistema caracteristicamente mutável, que pode ser sistematizado, descrito, associado e correlacionado a fatores de ordem linguística, social (LABOV, 2008) e interacional. Sendo assim, as relações entre língua e sociedade podem ser evidenciadas a partir do padrão dos usos linguísticos, dos significados e julgamentos sociais realizados pelos falantes na/pela língua.

As 30 entrevistas que compõem o *corpus* analisado neste trabalho possuem um formato semi-conduzido. Durante o período de coleta de dados, todos os entrevistadores incentivaram os informantes a discorrerem sobre experiências pessoais e temáticas de interesse social, selecionadas a partir do Sistema de Indicadores de Percepção Social, do Instituto de Pesquisa Econômica Aplicada – Ipea, que mede a avaliação dos brasileiros quanto às políticas públicas sobre gênero, saúde pública e desigualdade social. A coleta das entrevistas que constituem a amostra Deslocamentos 2020 foi feita em parceria com outros nove integrantes¹² do GELINS, pesquisadores de pós-graduação e de iniciação científica, assim nomeados:

¹⁰ A amostra Deslocamentos 2020 não foi usada somente neste trabalho. No GELINS, há pesquisadores que investigam a variação de fenômenos fonológicos e morfossintáticos na perspectiva dos estudos de contato, como Siqueira (2021) e Silva (2021). Por isso, os dados coletados foram divididos em quatro deslocamentos, nomeados 1, 2, 3 e 4, Algarismos presentes nos nomes dos arquivos. Eles fazem referência a falantes nascidos e que vivem na capital sergipana, no interior do estado ou que vieram dos estados da Bahia e Alagoas.

¹¹ Na ficha social preenchida antes da entrevista, o preenchimento da categoria gênero foi feito de modo aberto. Reportamos os rótulos feminino ou masculino pois foram esses que os falantes utilizaram.

¹² Na introdução deste texto, citamos os pesquisadores e seus respectivos trabalhos, também realizados com a amostra Deslocamentos 2020.

DOCUMENTADOR
DOCBP
DOCFR
DOCLS
DOCMS
DOCMJ
DOCPC
DOCVL
DOCVR
DOCVN

Quadro 8: perfil dos pesquisadores que conduziram as entrevistas sociolinguísticas

Fonte: elaborado pela autora

Para fechar a amostra e atender às necessidades de todos os pesquisadores do GELINS, convidamos pessoalmente, no entorno do Condomínio de Laboratórios Multiusuários de Informática e Documentação – LAMID e por redes sociais (Facebook e Instagram) alunos do *campus* Professor José Aloísio de Campos para participarem de uma entrevista. Alguns dos estudantes que aceitaram o convite deixaram o contato de potenciais informantes que, posteriormente, foram contactados. Em muitos casos, a pessoa indicada concordou em nos conceder entrevista. A princípio, os alunos sabiam que enquanto estivéssemos na cabine acústica conversaríamos sobre experiências pessoais, memórias de infância, vivência acadêmica e políticas públicas enquanto éramos gravados em áudio e vídeo. Em nenhum momento mencionamos que os dados coletados seriam utilizados para pesquisas de análise linguística.

4.2 INSTRUMENTOS DE COLETA

4.2.1 Roteiro de entrevista

As entrevistas que compõem o *corpus* analisado neste trabalho possuem um formato baseado no modelo proposto por Labov (2008). O modelo que adotamos (ANEXO 1) bem como a seleção dos tópicos abordados teve como objetivo possibilitar contextos de fala nos quais os

informantes discorressem sobre experiências pessoais e emitissem opiniões sobre assuntos que impactam a vida em sociedade.

4.2.2 *Ficha social e termo de consentimento livre e esclarecido*

Para delinear o perfil de cada estudante universitário, antes da entrevista os alunos foram convidados a preencher uma ficha social (ANEXO 3) e um termo de consentimento livre e esclarecido (ANEXO 2), por meio do qual autorizaram o uso dos dados gerados para fins de pesquisa acadêmica. A ficha social contém informações sobre cada informante: nome, curso de graduação, cidade de origem, nível de escolaridade dos pais, local onde mora, tipo de transporte utilizado para ir à Universidade Federal de Sergipe e onde almoça, pois essas informações são relevantes para outros trabalhos desenvolvidos no âmbito do GELINS.

4.3 VARIÁVEIS CONTROLADAS

Nesta seção, apresentamos as variáveis linguísticas, acústicas e de expressões faciais que controlamos para a execução do presente estudo. Dentre as variáveis linguísticas, consideramos como dependente o sentido indicado por *(eu) acho que*: certeza, dúvida, incerteza. Como independentes, consideramos o tipo de ocorrência do modalizador (*acho que, eu acho que*), escopo (oração, sintagma, turno), presença de outro modalizador (presença, ausência), tópico discursivo (perguntas de infância, moradia, lazer, educação, segurança pública, políticas de igualdade de gênero, saúde pública, atitudes linguísticas), experiência do falante com o que foi dito (direta, indireta) e polaridade (positiva, negativa). Dentre as variáveis prosódicas, consideramos as médias de frequência fundamental, intensidade, duração do modalizador, das pausas preenchidas e silenciosas para observarmos suas características em função do sentido indicado por *(eu) acho que*. Para analisarmos a relevância dos movimentos dos músculos do rosto na inferência de certeza, dúvida e incerteza controlamos as expressões de raiva, deboche, nojo, medo, felicidade, neutra, tristeza e surpresa, conforme a categorização proposta por Tejada et al (2021).

4.4 VARIÁVEIS LINGUÍSTICAS

4.4.1 *Ocorrência*

Achar pode ser realizado como verbo pleno que introduz sintagmas nominais e organiza uma estrutura argumental ou pode anteceder a conjunção *que*, dando forma a *(eu) acho que*,

podendo ou não incluir o sintagma nominal *eu*, como nos exemplos (2) e (3), na fala de um mesmo informante:

(2) MAI3MI: não *acho que* a escola pública a escola a educação deve ser regida pelo governo sim.

(3) MAI3MI: e uma um uma carga de conteúdo muito elevado *eu acho que* não dá pra conciliar

Devido ao elevado número de (*eu*) *acho que*, para a execução deste estudo computamos somente as 60 primeiras ocorrências dessa construção.

4.4.2 Ocorrências excluídas

O roteiro de entrevista que norteou a coleta da amostra Deslocamentos 2020 teve como objetivo instigar os falantes a discorrerem sobre aspectos pessoais de suas vidas e a opinarem sobre assuntos que afetam a sociedade como um todo. Ao longo das entrevistas, houve situações em que o documentador perguntou “qual a sua opinião sobre a política de cotas nas universidades públicas?” ou “você acha importante o debate sobre empoderamento feminino ou acha que isso é mimimi?”. No segundo caso, é possível que a presença de “você acha” na pergunta tenha influenciado a resposta do informante, como demonstrado em (4) e (5). Por isso, respostas com *acho que* imediatamente antecedidas “você acha” foram descartadas.

(4) DOCLS: você acha que a quantidade de médicos e especialistas no SUS é suficiente para atender a demanda da população?

(5) ANAIFF: *acho que* não porque se não num teria problemas de pessoas sem encontrar vagas nem médicos sem serem atendidas na fila de hospitais porque não tem (hes) quem atender então eu acho que precisaria de mais médicos.

Na amostra analisada, houve ocorrência de *acho* seguido de predicativo, como em (6):

(6) ANAIFF: Então eu não gosto muito das praias de Aracaju é tanto quando eu vou na praia geralmente eu vou nas da Barra Coque-Barra dos coqueiros que eu *acho* melhor.

Esse tipo de ocorrência foi excluído da análise, uma vez que nosso objeto de estudo foi *acho* mais o complemento *que*. Houve também ocorrência de *acho* no final de sentenças, acompanhado ou não de pronome:

(7) DOCB: tem pontos positivos e negativos (no SUS)?
VAG2FF: você precisa fazer uma cirurgia cê passa dias e dias dentro de um hospital s- eh ali sofrendo naquele ambiente super

contaminado de bactérias porque não tem vaga pra você fazer uma cirurgia simples que é colocar um osso no lugar no braço isso é uma falha muito grande *eu acho*.

Também excluimos esse tipo de ocorrência por causa da especificidade de objeto de análise desta pesquisa. Excluimos também realizações em que houve sobreposição de falas e risos sob justificativa de que não seria possível obter valores confiáveis dos parâmetros acústicos que caracterizam os sentidos de certeza, dúvida e incerteza.

4.4.3 Escopo

Como consequência dos processos de mudança semântica e gramaticalização, *(eu) acho que* não atua como verbo pleno. Essa construção seleciona como escopo orações, sintagmas e turnos, como exemplificado em (8), (9) e (10), respectivamente:

(8) DOCLS: por que teus pais colocaram o nome só por causa disso
CAR3FF: (...) que é o nome da minha irmã Carliane e botou Cariane *eu acho que* ela não encontrou em lugar nenhum ela não sabe o significado não tem

(9) DOCLS: essas pessoas podem identificar de onde você é sua origem apenas pela maneira como você fala?
CAR3FF: *acho que* sim

(10) ANT4FF: *acho que* talvez por isso tenta tido mais espaço

Em (8) *eu acho que* tem como escopo a oração “ela não encontrou em lugar nenhum ela não sabe o significado. Já em (9) a construção destacada tem como escopo apenas o sintagma nominal “sim”. Em (10) *acho que*, todo o turno de fala da informante, que indica incerteza sobre o espaço que pessoas LGBTQIA+ tem ganhado na mídia. Galvão (1999) e Freitag (2003) sugeriram que o sentido da construção que analisamos neste trabalho têm relação com seu escopo. Por isso, incluímos essa variável em nossas análises.

4.4.4 Presença de modalizador

(Eu) acho que indica os sentidos de certeza, dúvida e incerteza em relação ao seu escopo. Conforme argumentamos, essa construção não organiza uma estrutura argumental. Por isso, pode ser sucedida por outros modalizadores, como em (11) e (12):

(11) LAV4FF: *eu acho que* **infelizmente** isso coloca piadas junto com ações.

(12) LAI1FI: foi muito importante porque eu acho que **talvez** sem esse tipo de preparação eu não conseguiria.

Por entendermos que a presença de outros modalizadores depois de *(eu) acho que se* relaciona com o sentido dessa construção, incluímos essa variável em nossas análises.

4.4.5 Tópico discursivo

Ainda que a coleta de dados para a constituição da amostra Deslocamentos 2020 tenha sido guiada por um roteiro, as entrevistas sociolinguísticas não constituem um questionário. Inicialmente, os informantes responderam perguntas de checagem que incluíram informações pessoais: nome, cidade de origem e cidade de residência. Depois, os entrevistadores conduziram a conversa entre os tópicos discursivos abaixo.

1. Perguntas de infância;
2. Moradia;
3. Lazer;
4. Educação;
5. Segurança pública;
6. Políticas de igualdade de gênero;
7. Saúde pública;
8. Atitudes linguísticas.

Os temas sobre os quais os participantes da pesquisa foram convidados a discorrer são de caráter pessoal, profissional e social, como nos excertos (13) e (14) em que o informante compara sua infância com a das crianças de 2020 e depois fala sobre a importância da educação pública:

(13) ENI4MI: eu não vejo tantas crianças assim na rua brincando como eu brincava antigamente aí *eu acho que* mudou bastante essa interação.

(14) ENI4MI: então a educação pública tem que ser assegurada pelo governo para dar mais acessibilidade às pessoas *eu acho que* essa questão é imprescindível pra isso.

Essa estrutura teve como objetivo possibilitar que o falante transitasse entre diferentes tópicos discursivos, com os quais tinha ou não experiência o que, de acordo com o que Freitag (2003) sugeriu, influencia os sentidos indicados por *(eu) acho que*.

4.4.6 Experiência do falante

Freitag (2003) apresentou indícios de que a experiência do falante em relação ao tópico discursivo é um fator relevante para os sentidos indicados por *(eu) acho que*. Por isso, neste trabalho controlamos dois tipos de experiência, inferidas por meio das informações que o informante forneceu ao longo de sua fala:

1. Direta: o falante vivenciou o fato narrado e indica sentido suas experiências pessoais:

(15) DOCPC: foi produtivo assim pra você estudar de manhã e de noite era cansativo?

(16) LAIIFI: era cansativo mas foi muito importante porque *eu acho que* talvez sem esse tipo de preparação eu não conseguiria.

2. Indireta: o falante soube de algo por meio de outrem ou um julgamento de opinião a partir de suas suposições:

(17) LAIIFI: porque eu acho que o conhecimento eh eh como eu posso dizer ah não sei ele liberta as pessoas sabe.

4.4.7 Polaridade

(Eu) acho que indica certeza, dúvida e incerteza, (FREITAG, 2007a, p. 84) com a presença ou ausência de negação. Estruturalmente, isso é demarcado pela presença ou ausência do “não”, como nos exemplos (18) e (19):

(18) LAIIFI: *acho que* teve **sim** a durante o período histórico esse um acesso menor a educação essas coisas que geral igual a escola pública que geram uma certa eh como posso dizer uma dificuldade mais pra atingir o conhecimento.

(19) DOCMS: eh você considera importante as políticas para garantir o acesso e a permanência dos alunos na educação básica como transporte escolar livro didático merenda?

FRA3FF: *acho que* muita gente acaba **não** tendo como os pais não tem como pagar ajudar a manter.

De acordo com Freitag (2007b), a presença ou ausência de negação influencia o processamento das sentenças e, por conseguinte, a inferência de certeza, dúvida e incerteza. Por isso, incluímos essa variável na codificação dos dados a fim de observar se ela foi relevante para o sentido indicado por *(eu) acho que* na amostra Deslocamentos 2020.

4.5 VARIÁVEIS ACÚSTICAS

Nos estudos sobre entonação e prosódia há divergências quanto à definição e aos componentes dessas propriedades. No presente trabalho, adotamos uma abordagem ampla de prosódia (CRYSTAL, 1969; CRUTTENDEN, 1986) e nos baseamos na proposta de Antunes; Aubergé e Sasa (2014), Antunes e Aubergé (2015), Fernandes e Antunes (2017) para a análise que executamos. Por isso, consideramos como parâmetros prosódicos¹³:

1. Média da frequência fundamental de *(eu) acho que*;
2. Duração de *(eu) acho que*, de pausas silenciosas com mais de 150 milissegundos (KOWAL; WIESE; O'CONNELL, 1983) e preenchidas.

Oliveira (2011), Antunes; Aubergé e Sasa (2014), Fernandes e Antunes (2017) sugeriram que esses parâmetros apresentaram contornos significativos para diferenciar certeza de dúvida e incerteza. Tradicionalmente, estudos de prosódia tomam como objeto de análise sentenças completas, com início e fim bem delimitados, como nos resultados que apresentamos em 3.1. Contudo, o banco de dados analisado neste trabalho é constituído por entrevistas sociolinguísticas. Nelas, os informantes foram instigados a falar naturalmente o máximo possível. Houve ocorrências de *(eu) acho que* em sentenças curtas, como “*acho que sim*”, bem como em sentenças longas. Essa diferença poderia enviesar todos os parâmetros prosódicos, por isso tomamos como objeto de análise somente *(eu) acho que*. As variáveis prosódicas elencadas foram extraídas por meio do software PRAAT (BOERSMA; WEENINK, 2017), ferramenta gratuita que gera os dados numéricos de frequência fundamental, intensidade, duração, pausas silenciosas e preenchidas por meio do script *analyse tyer* (HIRST, 2012). Para isso, especificamos no PRAAT (BOERSMA; WEENINK, 2017) quatro tiras:

1. Oração, para todo o contexto de realização de *(eu) acho que*;
2. Mod, para *(eu) acho que*. As ocorrências dessa construção foram marcadas como MOD seguidas de algarismos que indicam a ordem de ocorrência do modalizador que indicou certeza (C), dúvida (D) e incerteza (I), conforme ilustrado nas figuras 11 e 12;
3. PS, seguidas de algarismos que também indicam a ordem de ocorrência das pausas silenciosas. As ocorrências foram marcadas como PSC, PSD, PSI quando a pausa ocorreu em contexto de certeza, dúvida e incerteza.

¹³ Para a documentação em áudio das entrevistas utilizamos o gravador Maratntz, modelo PMD661. Apesar de ele não ser o ideal para obtenção de dados para análise prosódica em decorrência dos ruídos internos do equipamento, optamos por manter a análise dos parâmetros de frequência fundamental e intensidade devido a ausência de estudos acústicos sobre certeza, dúvida e incerteza com foco na construção *(eu) acho que*.

4. PP, seguidas de algarismos que indicam igualmente a ordem de ocorrência das pausas preenchidas. As ocorrências foram marcadas como PPC, PPD, PPI quando a pausa ocorreu em contexto de certeza, dúvida e incerteza:

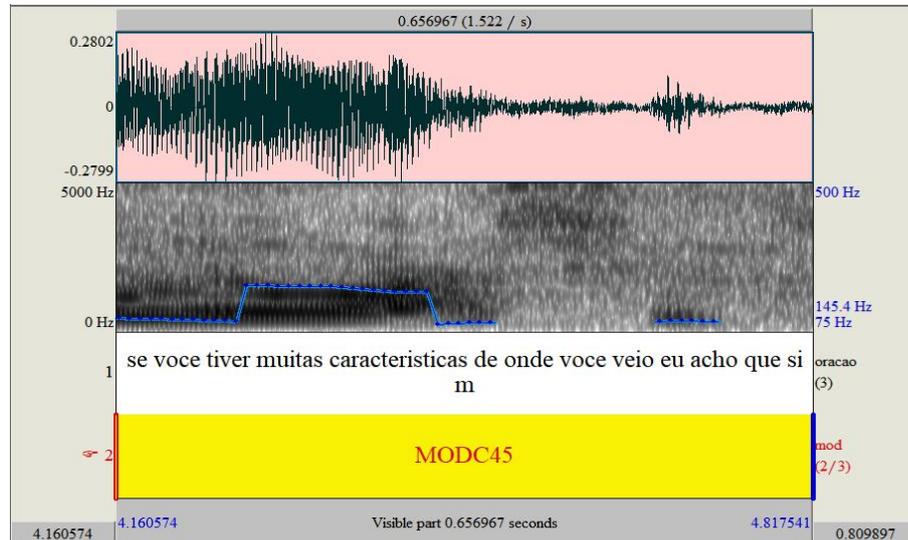


Figura 11: codificação prosódica de acho que (certeza)
Fonte: elaborada pela autora

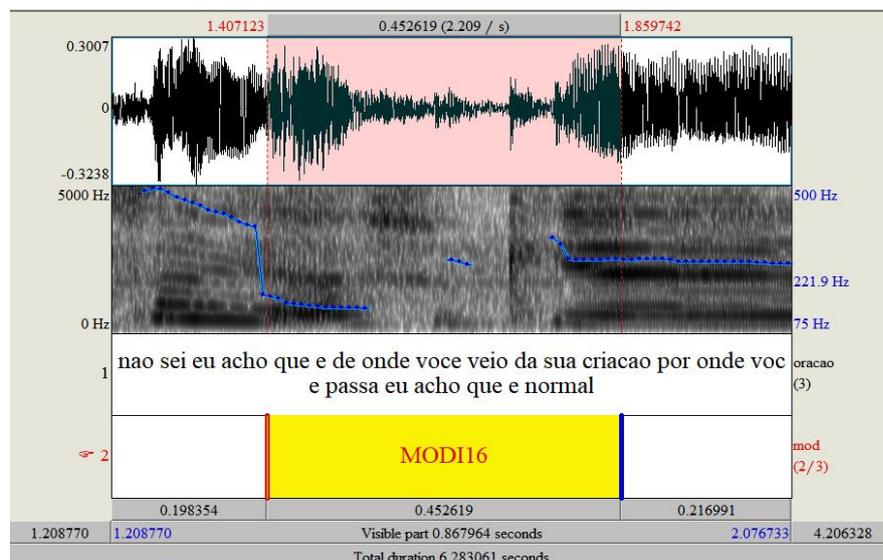


Figura 12: codificação prosódica de acho que (incerteza)
Fonte: elaborada pela autora

4.5.1 Medidas de frequência fundamental

A partir da segmentação de *(eu) acho que* na tira “mod” obtivemos, no PRAAT (BOERSMA; WEENINK, 2017), a média da frequência fundamental dessa construção. Os valores foram medidos em semitons, com medida de *pitch* em 100 Hertz a fim de captar a frequência mais perceptível pela audição dos participantes da interação e balancear as

diferenças fisiológicas entre falantes do sexo feminino e masculino, que de acordo com o estudo de Oliveira (2011) constituem diferenças significativas entre os sentidos. Baseamo-nos no método executado por Antunes, Aubergé e Sasa (2014); Antunes e Aubergé (2015); Fernandes e Antunes (2017), pois conforme sugerem os resultados dos estudos desses autores, os falantes produzem padrões de frequência fundamental com valores elevados como indício de dúvida e incerteza, ao contrário de certeza.

4.5.2 Medidas de duração

Medidas de duração também são relevantes para produção e inferência de dúvida, incerteza e certeza (OLIVEIRA, 2011; ANTUNES, AUBERGÉ, 2015; FERNANDES, ANTUNES, 2017). Por isso, elas foram incluídas na análise prosódica. Consideramos a duração de *(eu) acho que*, das pausas silenciosas e preenchidas com mais de 150 milissegundos (KOWAL; WIESE; O'CONNELL, 1983), marcadas nas transcrições ortográficas das entrevistas com *ah, eh, ih*, como em (20):

(20) DOCFR: o que é eu você acha? Você acha que realmente bandido bom é bandido morto?

JAD3M1: não *eu acho que* realmente **eh** tem que fazer tem que prender né.

A duração foi a única medida das pausas silenciosas e preenchidas que controlamos.

4.6 VARIÁVEIS PARALINGUÍSTICAS: EXPRESSÕES FACIAIS

As entrevistas que compõem a amostra Deslocamentos 2020 foram gravadas em vídeo com duas webcams Logitech C920 HD que geraram 24 frames por segundo do entrevistado e do entrevistador. Em nossas análises, consideramos somente os dados dos entrevistados. O procedimento para o reconhecimento facial dos movimentos dos rostos foi executado pelo OpenFace, de acordo com os parâmetros programados por Tadas Batrusaitis¹⁴, que possibilitou a extração das características das faces dos estudantes para que alimentássemos, conforme a categorização proposta por Tejada et al (2021), o classificador de expressões faciais.

¹⁴ Disponível em: github.com/TadasBaltrusaitis/OpenFace

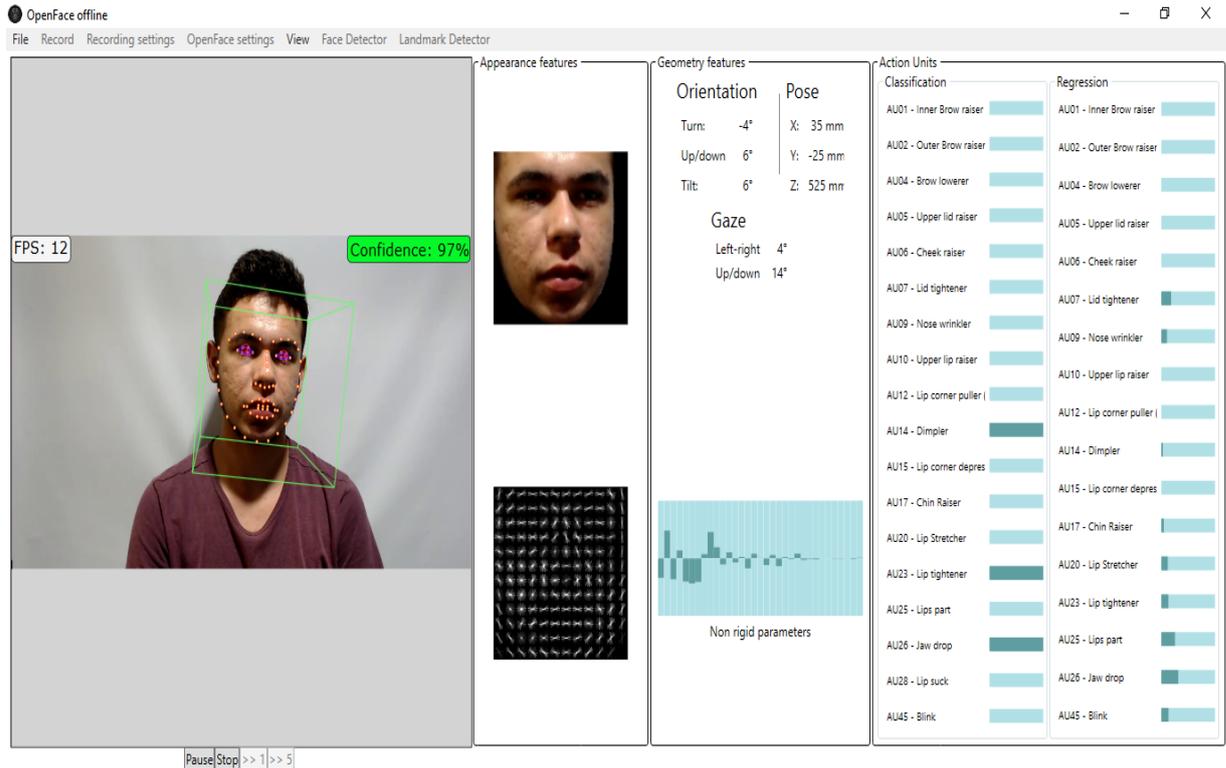


Figura 13: processamento das expressões faciais no OpenFace

Fonte: elaborada pela autora

Para o reconhecimento e classificação das expressões faciais utilizamos como referência imagens de informantes brasileiros e colombianos, classificadas por professores universitários da área das ciências humanas por meio de uma *survey* online (TEJADA ET AL, 2021) e separadas nos arquivos “Train_Brazil.npy” e “Train_Colombia.npy”. Cada arquivo teve um correspondente para o procedimento de *machine learning* nomeado “Train_Brazil_labels.npy” e “Train_Colombia_labels.npy”, com imagens na seguinte ordem:

0. raiva;
1. deboche;
2. nojo;
3. medo;
4. felicidade;
5. neutra;
6. tristeza;
7. surpresa.

A fim de classificar cada expressão facial, utilizamos um algoritmo que passou por diversos testes para que soubéssemos:

1. Qual recurso ou combinação de recursos fornecidos pelo OpenFace produziria os melhores resultados em termos de classificação precisa das expressões;
2. Quais valores de parâmetros deveríamos utilizar para o classificador;

Após validarmos esse algoritmo, elaboramos um classificador com a biblioteca Python de *machine learning* sk-learn (do sci-kit), no modelo do Support Vector Classification (SVC). Para treinarmos esse classificador submetemos o banco de dados brasileiro ao OpenFace, que gerou 8 arquivos csv, um para expressão de cada emoção. Esses mesmos arquivos continham parâmetros de marcos faciais, informações bi e tridimensionais da cabeça do informante, estimativas da direção do olhar e as FACS. Todos os arquivos csv foram convertidos em *pandas dataframes* para agruparmos os dados nas seguintes categorias:

ATTR	INTERVAL
gaze:	gaze_0_x to gaze_angle_y
eye landmarks 2D:	eye_lmk_x_0 to eye_lmk_y_55
eye landmarks 3D:	eye_lmk_X_0 to eye_lmk_Z_55
head pose (no rotation):	pose_Tx to pose_Tz
head pose (rotation):	pose_Rx to pose_Rz
face landmarks 2D:	x_0 to y_67
face landmarks 3D:	X_0 to Z_67
poses 1:	p_scale to p_ty
poses 2:	p_0 a p_33
facets:	AU01_r to AU45_c

A partir dessas informações, criamos um dataframe com os rótulos raiva, deboche, nojo, medo, felicidade, neutra, tristeza e surpresa. Se o arquivo "raiva.csv" teve 50 entradas, criamos 50 entradas no dataframe "rótulos" com valor 0 (que é o valor correspondente à raiva no modelo adotado). Esse procedimento permitiu que tivéssemos um número x de 1's, onde x foi o número de entradas do próximo arquivo csv. Depois, dividimos os resultados em um conjunto de treinamento e outro de teste, com uma proporção de 80/20, respectivamente. Essa proporção é uma prática comum em *machine learning*. Para evitarmos distorções no algoritmo, a amostragem de quais entradas foram para os conjuntos foi aleatória. Com a organização dos rótulos das emoções, inicializamos um classificador SVC e o alimentamos com os conjuntos de dados mencionados anteriormente. Criamos uma matriz de confusão para registrar e medir os resultados do treinamento e os plotamos em um mapa de calor usando matplotlib, que indicou

que o parâmetro que produziu melhores resultados para indicar expressões faciais foi a categoria "FACS" (colunas "AU01_r" a "AU45_c" nos arquivos csv criados pelo OpenFace). Por isso, somente ela foi utilizada na classificação dos frames das entrevistas sociolinguísticas que compõem a amostra Deslocamentos 2020.

4.7 TRATAMENTO ESTATÍSTICO

Após a codificação dos dados, realizamos análises descritiva e inferencial para identificarmos a distribuição da variável dependente (sentidos de certeza, dúvida, incerteza indicados por *acho que* e *eu acho que*) quanto às independentes (tipo de ocorrência, escopo, presença de modalizador, tópico discursivo, experiência do falante e polaridade). Elaboramos testes de associação com pacote library (sjplot), com a função sjt.xtab, no R (CORE TEAM, 2018). O teste estatístico de associação que utilizamos foi o *qui-quadrado* de *Pearson* com estatística de quadrado (χ^2). Adotamos 5% ($p < 0.05$) como valor de referência para alfa. Para medir a força de associação entre variável dependente e independente utilizamos V^2 de *Cramer*, definido por três níveis:

- i) 1 representa o nível mais forte de associação entre as variáveis;
- ii) 0 representa o nível mais fraco de associação entre as variáveis.

Para este teste, consideramos duas hipóteses:

- i) H_0 : não há associação entre as variáveis;
- ii) H_1 : há associação entre as variáveis.

Os resultados da análise estatística inferencial foram apresentados por meio de tabelas de contingência (FREITAG, 2020). Em seguida, procedemos à análise das variáveis prosódicas. Utilizamos o pacote library (rstatix) para extrairmos os valores das médias e desvio padrão das ocorrências de *(eu) acho que*, das pausas silenciosas e preenchidas e realizamos análises de variância para calcularmos se as médias de frequência fundamental, intensidade, duração do modalizador e dos dois tipos de pausas diferenciaram os sentidos de certeza, dúvida e incerteza. Por último, tratamos os dados de expressões faciais. Realizamos uma análise qualitativa¹⁵ de apenas 2 das 30 entrevistas da amostra Deslocamentos 2020: DAN2MF e FRA3FF, um falante

¹⁵ Até o prazo para a defesa desta dissertação, não conseguimos construir e testar um modelo de análise quantitativa para os dados de expressões faciais das 30 entrevistas gravadas em vídeo que utilizamos como amostra. Por isso, neste momento, foi possível somente a execução da análise qualitativa.

do sexo/gênero masculino e outro, feminino. A partir dos resultados das FACS, transformamos os valores dos picos e amplitudes dos segmentos em que houve ocorrência de *(eu) acho que* para gerarmos uma escala de 01 a 07 que corresponde às movimentações musculares dos rostos dos participantes:

- | | |
|--------------|--------------------|
| a) peaks_e-0 | i) max_amplitude_0 |
| b) peaks_e-1 | j) max_amplitude_1 |
| c) peaks_e-2 | k) max_amplitude_2 |
| d) peaks_e-3 | l) max_amplitude_3 |
| e) peaks_e-4 | m) max_amplitude_4 |
| f) peaks_e-5 | n) max_amplitude_5 |
| g) peaks_e-6 | o) max_amplitude_6 |
| h) peaks_e-7 | |

Com esses valores, desenvolvemos um *script* na plataforma R (TEAM CORE, 2018) para gerar gráficos com o pacote *ggplot2* com curvas para indicar a intensidade da ocorrência de cada movimento, demonstrado por meio de *action units* (AUs). Utilizamos esse procedimento para inferir, a partir do que sugeriram os estudos de prosódia gestual (SWERTS ET AL, 2003; ANTUNES; AUBERGÉ; SASA, 2014; CRUZ, SWERTS, FROTA, 2017) se os movimentos dos músculos do rosto deram pistas dos sentidos indicados por *(eu) acho que*. Os conjuntos de dados que analisamos (linguísticos, prosódicos e de expressões faciais) bem como todos os *scripts* utilizados estão integralmente disponíveis em: <https://osf.io/w89v4/>

5 RESULTADOS E DISCUSSÕES

Ao todo, computamos 1038 ocorrências de *(eu) acho que*. A média de realização dessa construção por falante foi de 35, com o mínimo de 7 e o máximo de 60 ($DP = 15$). Desse total, 34,4% (357) correspondeu à realização de *acho que* e 65,6% (681) à realização de *eu acho que* para indicar os sentidos de certeza, dúvida e incerteza. De 5.1.1 a 5.1.6 apresentamos resultados de associação entre as variáveis linguísticas independentes (ocorrência, escopo, presença do modalizador, tópico discursivo, experiência do falante e polaridade) em relação à dependente (sentidos de certeza, dúvida, incerteza indicados por *acho que* e *eu acho que*). Na seção 5.2, apresentamos os resultados de análises de variância (ANOVAs) multiníveis a partir das quais calculamos se os parâmetros acústicos (frequência fundamental, intensidade, duração, pausas silenciosas e preenchidas) diferenciaram os sentidos indicados por *(eu) acho que* nas 30 entrevistas da amostra Deslocamentos 2020 que analisamos. Por último, em 5.3 apresentamos os resultados da análise qualitativa das expressões dos informantes DAN2MF e FRA3FF.

5.1 VARIÁVEIS LINGUÍSTICAS

5.1.1 Ocorrência

Na amostra Deslocamentos 2020, os informantes indicaram os sentidos de certeza, dúvida e incerteza por meio de *acho que* e *eu acho que*. De acordo com Gonçalves (1999), essa construção pode ser caracterizada quanto ao tipo de ocorrência como:

- i) *Achar 2* [Sintagma nominal – *achar* – Sintagma adjetival], modalizador avaliativo que indica maior grau de certeza;
- ii) *Achar 3* [Sintagma nominal – *achar* – *que*], modalizador que indica fatores externos ao falante e fazem com que a incerteza passe para o nível das possibilidades. Dois exemplos de ocorrência dessas duas tipologias são:

(21) DOCMJ: cê acha que eh ah a infância das crianças hoje mudou muito em relação a sua infância lá?
 MAI3MI: eu creio que sim muito (...) tem crianças mesmo com dez anos mesmo já tendo esses aparelhos *eu acho que* está muito evoluído em relação ao meu.

(22) DOCFR: tirando a entrada tem outro problema de segurança que você enxerga aqui dentro da UFS?
 JAD3MI: recentemente *acho que* foi antes de ontem teve um assalto ali

perto do CODAP foi um cara de chegou dois cara de moto acho pegou bolsa que tinha notebook do aluno e tudo e aí saiu.

Em (21), depois de “eu creio que sim muito (...) tem criança mesmo com dez anos mesmo já tendo esses aparelhos” o informante utilizou *eu acho que* antes da sentença “está muito evoluído em relação ao meu” para indicar uma certeza pautada na comparação entre sua experiência de infância e a das crianças de 2020. Já em (22) o informante utilizou *acho que*, sem o sintagma nominal, antes de “foi antes de ontem teve um assalto ali perto do CODAP”, sentença que relata um acontecimento que ele soube por outrem e que dá indícios do sentimento de insegurança na universidade. Nos dois excertos que destacamos, os dois tipos de ocorrência da construção indicaram sentidos diferentes. Por isso, questionamos se existe associação entre os sentidos de certeza, dúvida, incerteza e ocorrência de *acho que* e *eu acho que*. As nossas hipóteses para o teste estatístico foram: H_0 não existe associação entre as variáveis sentido e ocorrência, H_1 : existe associação entre as variáveis. A partir da categorização proposta por Gonçalves (1999), controlamos a ocorrência de *acho que* com e sem “eu” para observarmos se a presença do sintagma nominal influenciou a inferência dos sentidos de certeza, dúvida e incerteza:

<i>Sentido</i>	<i>Ocorrência</i>		<i>Total</i>
	Acho que	Eu acho que	
Certeza	261 32 %	554 68 %	815 100 %
Dúvida	44 45.8 %	52 54.2 %	96 100 %
Incerteza	52 40.9 %	75 59.1 %	127 100 %
<i>Total</i>	357 34.4 %	681 65.6 %	1038 100 %

$$\chi^2=10.011 \cdot df=2 \cdot \text{Cramer's } V=0.098 \cdot p=0.007$$

Tabela 14: associação entre função de (*eu*) *acho que* e tipo de ocorrência

Fonte: elaborada pela autora

Os dados apresentados na tabela 14 apontam que *acho que* foi utilizado pelos falantes para indicar certeza em 32% (261/815) das ocorrências, dúvida em 45,8% (44/96) e incerteza em 40,9% (52/127); enquanto *eu acho que* foi utilizado para indicar certeza em 68% das

ocorrências (554/815), dúvida em 54,2% (52/96) e incerteza em 59,1% (75/127). Tais resultados sugerem que o conhecimento do falante foi demarcado pela presença de “eu”, utilizado pelos falantes para demonstrar maior assertividade. Os resultados do teste qui-quadrado sugeriram que houve associação entre o tipo de ocorrência (*acho que, eu acho que*) e seu sentido ($\chi^2 = 10.011 \cdot df = 2, p = 0.007$) com força fraca ($V^2 = 0.098$).

5.1.2 Escopo

Segundo Galvão (1999) e Gonçalves (2003), (*eu*) *acho que* não funciona como verbo pleno e não organiza nenhuma estrutura argumental por consequência dos processos de mudança semântica e gramaticalização, que ocorre em uma escala que vai do concreto ao abstrato: corpo > objeto > processo > espaço > tempo > qualificação (CUNHA, OLIVEIRA; VOTRE, 1999). No último estágio da escala, (*eu*) *acho que* indica os sentidos de certeza, dúvida e incerteza. De acordo com Freitag (2003), o tipo de complemento introduzido pelo modalizador tem relação com o sentido indicado. Para a análise dos dados que computamos, transformamos a variável escopo, codificada em três níveis (oração, sintagma, turno), em binária, com dois níveis (interno, externo) e questionamos se eles estavam associados aos sentidos de certeza, dúvida e incerteza indicados por (*eu*) *acho que*. As nossas hipóteses para o teste estatístico foram: H_0 não existe associação entre sentido de (*eu*) *acho que* e escopo, H_1 existe associação entre sentido de (*eu*) *acho que* e escopo. Os resultados da variável escopo são apresentados na tabela que segue:

<i>Sentido</i>	<i>Escopo</i>		<i>Total</i>
	Externo	Interno	
Certeza	49 6 %	766 94 %	815 100 %
Dúvida	5 5.2 %	91 94.8 %	96 100 %
Incerteza	18 14.2 %	109 85.8 %	127 100 %
<i>Total</i>	72 6.9 %	966 93.1 %	1038 100 %

$$\chi^2=11.826 \cdot df=2 \cdot \text{Cramer's } V=0.107 \cdot p=0.003$$

Tabela 15: associação entre sentido de (*eu*) *acho que* e escopo
Fonte: elaborada pela autora

Na amostra Deslocamentos 2020, observamos que o escopo (externo, interno) apresentou associação com o sentido de certeza, dúvida e incerteza. Houve maior distribuição da construção com escopo interno para indicar dúvida, com 94.8% (91/96) de ocorrência, certeza, com 94% (766/815) e incerteza, com 85.8% (109/127), e menor distribuição do escopo externo para indicar incerteza, com 14.2% (18/127) de ocorrência, dúvida, com 5.2% (5/96) e certeza, com 6% (49/815). Os resultados do teste de qui-quadrado sugeriram associação entre sentido de *(eu) acho que* ($\chi^2 = 11.826 \cdot df = 2, p = 0.003$) com força fraca ($V^2 = 0.107$), o que Freitag (2003) indicou ser consequência da gramaticalização. Essa construção funciona como parentética, ou seja, toma como escopo orações e sintagmas (FREITAG, 2007), como exemplificado em (22):

(22) DOCVL: você considera que esses programas (PROUNI, FIES) diminuem a desigualdade social entre os jovens?
ANT4FF: cara... *eu acho que* dá acesso a uma boa educação e a educação é uma ferramenta que você pode utilizar pra você melhorar a sua a sua condição de vida.

A informante ANTF44 produziu o marcador discursivo “cara...” para chamar a atenção do documentador para o início de sua resposta. Depois, utilizou *eu acho que* para definir seu posicionamento. A sentença “dá acesso a uma boa educação” está estruturada a partir do verbo dar (dá), que não possui dependência sintática com *eu acho que*. De acordo com a categorização utilizada por (FREITAG, 2003, p. 46), o que segue após a construção é um complemento integrado que indica certeza, demonstra a opinião do informante, expõe sua relação com fatos do mundo e o situa em um contexto específico (BYBEE; PERKINS; PAGLIUCA, 1994), como em (22).

5.1.3 Presença de modalizador

Enquanto construção modalizadora, *(eu) acho que* inclui força à asserção do informante (FREITAG, 2007a). Depois dela pode haver ocorrência de outro modalizador. Por isso, questionamos se a presença de outro modalizador na oração estava associada ao sentido de *(eu) acho que* para indicar certeza, dúvida e incerteza. As nossas hipóteses para o teste estatístico foram: H_0 : não há associação entre a presença de outro modalizador e o sentido indicado por *(eu) acho que*, H_1 há associação entre as variáveis. Em toda amostra Deslocamentos 2020 houve ocorrência de “basicamente”, “infelizmente”, “justamente”, “talvez”, “só”, “até” e “principalmente”, como exemplificado em (22) e (23):

(22) GON1MF: *acho que talvez* por isso você mesmo falou que “ah você não parece daqui” porque eu tenho alguns hábitos de fala que eu acabei pegando por convívio mesmo.

DOCVR: (...) e de modo geral você gosta de morar no Santa Maria?

(23) DAN1MF: eu gosto *eu acho que eh tipo* eu não sei explicar assim.

. Para melhor visualização dos resultados, codificamos essa variável em dois níveis: presença e ausência:

<i>Sentido</i>	<i>Modalizadores</i>		<i>Total</i>
	Ausência	Presença	
Certeza	802 98.4 %	13 1.6 %	815 100 %
Dúvida	93 96.9 %	3 3.1 %	96 100 %
Incerteza	118 92.9 %	9 7.1 %	127 100 %
Total	1013 97.6 %	25 2.4 %	1038 100 %

$$\chi^2=14.329 \cdot df=2 \cdot \text{Cramer's } V=0.117 \cdot \text{Fisher's } p=0.002$$

Tabela 16: associação entre sentido de (*eu*) *acho que* e modalizadores

Fonte: elaborada pela autora

Para certeza, com 98.4% (802/815) de ocorrência e dúvida, com 96.9% (93/96) de ocorrência e para incerteza, com 92.9% (118/127) o padrão foi a ausência de outros modalizadores depois de (*eu*) *acho que*. Os resultados do teste de qui-quadrado indicaram que houve associação entre o sentido dessa construção e a presença de outros modalizadores ($\chi^2 = 14.329$ $df = 2$, $p = 0.002$) com força fraca ($V^2 = 0.117$).

Nos excertos (22) e (23) observamos a ocorrência do que Galvão (1999) classificou como “Achar 3”, que põe a asserção do falante no nível das possibilidades, de modo a demonstrar uma suposição com menor força de asserção, como nos exemplos excertos dos falantes que destacamos, o que é corroborado pela maior frequência da presença de outros modalizadores após (*eu*) *acho que* para indicar dúvida e incerteza. Em (22), além de *acho que* há ocorrência de **talvez** antes de “por isso você mesmo falou que “ah você não parece daqui””, o que compreendemos como estratégia para reiterar a incerteza do informante quanto ao motivo

do comentário da documentadora. Em (23), na fala de DANIMF, depois de *eu acho que* há ocorrência de uma pausa preenchida (eh) mais a forma **típo** antes dele informar que não sabe explicar o que havia sido perguntado pelo documentador.

5.1.4 *Tópico discursivo*

As entrevistas sociolinguísticas que compõem a amostra Deslocamentos 2020 são semiestruturadas e abrangem tópicos de caráter pessoal e social, com o objetivo de instigar os informantes a expressarem, do modo mais natural possível, memórias, palpites e opiniões. Controlamos essa variável porque, de acordo com Freitag (2003), o tópico discursivo influencia o sentido de *(eu) acho que*. Por isso, questionamos se o sentido de *(eu) acho que* apresenta associação com o tópico discursivo. As nossas hipóteses para o teste estatístico foram H₀: não há associação entre sentido e tópico, H₁: há associação entre essas duas variáveis.

<i>Sentido</i>	<i>Tópico</i>								<i>Total</i>
	Atitudes ling.	Checagem	Educação	Lazer	Moradia	Quest. Gên.	Saúde	Segurança	
Certeza	134 16.4 %	35 4.3 %	275 33.7 %	42 5.2 %	83 10.2 %	148 18.2 %	32 3.9 %	66 8.1 %	815 100 %
Dúvida	13 13.5 %	12 12.5 %	21 21.9 %	13 13.5 %	30 31.2 %	2 2.1 %	3 3.1 %	2 2.1 %	96 100 %
Incerteza	33 26 %	10 7.9 %	33 26 %	4 3.1 %	11 8.7 %	16 12.6 %	3 2.4 %	17 13.4 %	127 100 %
Total	180 17.3 %	57 5.5 %	329 31.7 %	59 5.7 %	124 11.9 %	166 16 %	38 3.7 %	85 8.2 %	1038 100 %

$$\chi^2=93.766 \cdot df=14 \cdot \text{Cramer's } V=0.213 \cdot \text{Fisher's } p=0.000$$

Tabela 17: associação entre sentido de *(eu) acho que* e tópico discursivo

Fonte: elaborada pela autora

Na tabela 17, observamos que a maior distribuição de *(eu) acho que* para indicar certeza ocorreu nos tópicos educação (33.7%), questões de gênero (18.2%) e atitudes linguísticas (16.4%), enquanto a maior distribuição de dessa construção para indicar dúvida ocorreu em moradia (31.2%) e educação (21.9%). Para indicar incerteza, *(eu) acho que* foi majoritariamente utilizada em educação (26%) e atitudes linguísticas (26%). Os resultados sugerem associação

entre sentido e tópico ($\chi^2 = 93.766 \cdot df = 14, p = 0.000$) com força fraca ($V^2 = 0.213$). É possível que a maior distribuição de *(eu) acho que* para indicar dúvida tenha ocorrido em moradia porque nesse tópico os informantes responderam perguntas de caráter pessoal. Em alguns momentos eles tiveram que relembrar informações para responder o questionamento do documentador, como no excerto (24):

(24) DAN2MF: É isso *eu acho que* não tenho muito lugar que eu vá pelo menos não lembro agora.

No excerto acima, DAN2MF falou sobre o que gosta de fazer quando está no município de Itabaiana, informação baseada em sua experiência. Nos tópicos em que a informação pode partir tanto da experiência pessoal quanto individual, observamos que o sentido de *(eu) acho que* flutuou: essa construção apresentou altos percentuais de uso para indicar certeza (33.7%), dúvida (21.9%) e incerteza (26%) no tópico educação. Esses valores demonstram a relevância daquilo sobre o que se fala para a construção das pistas de evidencialidade que tanto Galvão (1999) quanto Freitag (2003) e Gonçalves (2003) apontaram como característica do uso da construção modalizadora *(eu) acho que*.

5.1.5 *Experiência do falante*

Freitag (2003), ao analisar o uso de *acho que* como indicativo de certeza e dúvida em entrevistas sociolinguísticas, afirmou que “na função *marcador de dúvida*, a forma *acho* tende a ocorrer em contextos semântico discursivos de maior envolvimento do falante com o assunto discorrido”. (FREITAG, 2003, p. 102, grifos da autora). O envolvimento foi definido por ela como o tipo de experiência em relação ao que foi dito, o que pode ser inferido por pistas presentes na fala do informante. Controlamos essa variável na intenção de observamos a relevância do tipo de experiência dos falantes na nossa amostra. Questionamos se o tipo de envolvimento (direto, indireto) estava associado aos sentidos de certeza, dúvida e incerteza indicados por *(eu) acho que*. As nossas hipóteses para o teste estatístico foram H_0 : não há associação entre sentido e experiência do falante, H_1 : há associação entre essas variáveis.

<i>Sentido</i>	<i>Experiência</i>		<i>Total</i>
	Direta	Indireta	
Certeza	490 60.1 %	325 39.9 %	815 100 %
Dúvida	94 97.9 %	2 2.1 %	96 100 %
Incerteza	5 3.9 %	122 96.1 %	127 100 %
<i>Total</i>	589 56.7 %	449 43.3 %	1038 100 %

$$\chi^2=214.376 \cdot df=2 \cdot \text{Cramer's } V=0.454 \cdot p=0.000$$

Tabela 18: associação entre sentido de (*eu acho que*) e a experiência do falante

Fonte: elaborada pela autora

Assim como nos dados apresentados por Freitag (2003), na amostra Deslocamentos 2020 (*eu acho que*) foi mais utilizada para indicar dúvida sobre fatos vivenciados diretamente pelos participantes da pesquisa em 97.9% das ocorrências (94/96). Os percentuais de ocorrência de (*eu acho que*) para indicar certeza e incerteza foram 60.1% (490/325) e 3.9% (5/127), respectivamente, distribuição que sugere que o que diferenciou os sentidos de dúvida e incerteza foi o tipo de experiência do falante. O teste de qui-quadrado sugeriu a associação entre essas duas variáveis ($\chi^2 = 214.376 \cdot df = 2, p = 0.000$) com força mediana ($V^2 = 0.454$).

(25) JAD3MI: *acho que* a principal venda da cidade é a agricultura

(26) DOCFR: mas pra você qual o principal problema dessas gestões?
JAD3MI: realmente eu não sei *acho que* é ser (os) eh os administradores né que estão fazendo alguma coisa errado né.

Em (25) *acho que* antecede a sentença “a principal venda da cidade é a agricultura” e indica dúvida a partir da experiência do falante: ele vivia em Canindé de São Francisco (SE) e se mudou para São Cristóvão (SE) para estudar. A partir dos seus conhecimentos de ex-morador, JAD3MI supôs qual a principal fonte de renda de sua antiga cidade de residência. É diferente do que ocorreu em (26): nesse excerto, depois da oração “eu não sei” o falante fez uso de *acho que*, de pausa preenchida (eh) e, em seguida, disse “os administradores né que estão fazendo alguma coisa errado né”, afirmação que não foi baseada em nenhum conhecimento prévio mencionado na entrevista. Além de corroborar a relevância do tipo de experiência do informante para a inferência de certeza, dúvida e incerteza, esses resultados favorecem a

proposta de Silva (2008, p. 39) para a definição de dúvida, incerteza e certeza, pois “demonstram uma gradação de acordo com o nível de conhecimento sobre um assunto”.

5.1.6 Polaridade

Ao discorrerem sobre aspectos pessoais e sociais nas entrevistas sociolinguísticas, os informantes demonstram os sentidos de certeza, dúvida e incerteza em orações com ou sem a presença de negação, ou seja, marcadas ou não marcadas, características que de acordo com Freitag (2007b) influenciam o processamento de sentenças e possivelmente sua interpretação. Por isso, questionamos se o sentido de *(eu) acho que* estava associado à polaridade do seu escopo. As nossas hipóteses para o teste estatístico foram: H_0 não há associação entre sentido e polaridade, H_1 há associação entre essas duas variáveis.

<i>Sentido</i>	<i>Polaridade</i>		<i>Total</i>
	Negativa	Positiva	
Certeza	98 12 %	717 88 %	815 100 %
Duvida	16 16.7 %	80 83.3 %	96 100 %
Incerteza	22 17.3 %	105 82.7 %	127 100 %
<i>Total</i>	136 13.1 %	902 86.9 %	1038 100 %

$\chi^2=3.890 \cdot df=2 \cdot Cramer's V=0.061 \cdot p=0.143$

Tabela 19: associação entre o sentido de *(eu) acho que* e a polaridade da

Fonte: elaborada pela autora

Na tabela 19, observamos que na amostra Deslocamentos 2020 houve, majoritariamente, ocorrência de sentenças de polaridade positiva. Em 88% (717/815) das ocorrências elas indicaram certeza, em 83.3% (80/96) dúvida e em 82.7% (105/127), incerteza. O teste de qui-quadrado não indicou associação entre polaridade e sentido de *(eu) acho que* ($\chi^2 = 3.890 \cdot df = 2, \cdot p = 0.143$), e a força entre elas foi fraca ($V^2 = 0.061$), resultados que apontam que, nesta amostra, a estrutura da sentença para indicar a força assertiva e, por conseguinte, o sentido da construção *(eu) acho que* não foi relevante.

As análises inferenciais com estatística qui-quadrado apontaram associação entre variáveis estruturais (ocorrência, escopo e presença do modalizador) e subjetivas (tópico discursivo, tipo de experiência) e os sentidos de certeza, dúvida e incerteza indicados por *eu acho que* e *acho que*, resultados alinhados ao que os estudos descritivos sobre essa construção pautados no paradigma da gramaticalização sugerem (GALVÃO, 1999; GONÇALVES, 2003; FREITAG, 2003). A fim de contribuirmos com esse panorama e possibilitarmos um aumento do escopo de análise, procedemos a uma análise de parâmetros acústicos para investigarmos sua relevância na inferência de certeza, dúvida e incerteza. Na próxima seção, apresentamos os resultados da análise da frequência fundamental, intensidade, duração, pausas silenciosas e preenchidas quanto aos sentidos indicados por (*eu*) *acho que*.

5.2 VARIÁVEIS ACÚSTICAS

Para análise das variáveis acústicas, utilizamos o pacote library (rstatix) para extrairmos os valores das médias e desvio padrão das ocorrências de (*eu*) *acho que*, da frequência fundamental, intensidade e duração dessa construção, mais os valores de duração das pausas silenciosas e preenchidas para testarmos, por meio de análises de variância multiníveis (ANOVAs), se as diferenças entre os parâmetros acústicos foram relevantes para a inferência de certeza, dúvida e incerteza

5.2.1 Frequência fundamental de (*eu*) *acho que*

De acordo com Silva (2008); Antunes; Aubergé e Sasa (2014); Antunes e Aubergé (2015); Fernandes e Antunes (2017), orações que indicam dúvida e incerteza apresentam média de frequência fundamental maior do que as que indicam certeza. A análise descritiva da amostra Deslocamentos 2020 apresentou resultado semelhante, embora a diferença tenha sido pequena:

Sentido	Variável prosódica	Nº	Média	Desvio padrão	Min	Max
Certeza	f0_media_modalizador	805	10.3	5.70	-3.70	26.8
Dúvida	f0_media_modalizador	93	10.9	6.16	-2.38	21.6
Incerteza	f0_media_modalizador	125	11.1	6.25	-1.45	25.5

Tabela 20: valores descritivos da média de f0 em relação ao sentido de (*eu*) *acho que*

Fonte: elaborada pela autora

A partir das médias expressas na tabela 20, calculamos, por meio de uma análise de variância, se a frequência fundamental foi relevante para a inferência de certeza, dúvida e

incerteza. O resultado indicou que não ($f(2, 1020) = 1.44, p = 0.237$), o que pode ser consequência do recorte que norteou este trabalho: analisar somente as ocorrências de *(eu) acho que* e não sentenças inteiras, decisão tomada em função do foco nessa construção e da amostra analisada, composta por entrevistas sociolinguísticas nas quais, durante em média 50 minutos, os informantes discorreram sobre diversos temas e enunciaram orações com extensões assimétricas, o que tornaria inviável a comparação entre elas.

Considerando que certeza, dúvida e incerteza foram expressos por *eu acho que* e *acho que*, executamos outra ANOVA para calculamos se a frequência fundamental diferenciou sentido e o tipo de ocorrência dessa construção. Na amostra Deslocamentos 2020, nem o sentido ($f(2, 1019) = 1.44, p = 0.236$) nem o tipo de ocorrência ($f(1, 1019) = 1.33, p = 0.230$) foram diferenciados pela frequência fundamental, resultados contrários ao que apontam os estudos prosódicos que descrevem sentenças que indicam certeza, dúvida e incerteza no português brasileiro.

5.2.2 Intensidade de *(eu) acho que*

A intensidade indica a força do som (BARBOSA, 2019, p. 26) e essa percepção pode exercer efeito na inferência de sentidos. O estudo de Oliveira (2011) sugeriu que a média da intensidade é maior em sentenças que indicam certeza. Na análise dos sentidos indicados por *(eu) acho que* na amostra Deslocamentos 2020 obtivemos os seguintes valores:

Sentido	Variável prosódica	Nº	Média	Desvio padrão	Min	Max
Certeza	inten_media_modalizador	815	69.3	2.78	55.0	73.8
Dúvida	inten_media_modalizador	96	68.0	3.70	55.8	72.8
Incerteza	inten_media_modalizador	127	69.1	2.65	54.7	73.0

Tabela 21: valores descritivos da média da intensidade em relação ao sentido de *(eu) acho que*

Fonte: elaborada pela autora

A partir da distribuição de dados registrados na tabela 21, executamos uma análise de variância para calculamos se a intensidade da construção modalizadora diferenciou certeza, dúvida e incerteza. Os resultados indicaram que sim ($f(2, 1035) = 8.88, p < .001$).

Mais uma vez, considerando que certeza, dúvida e incerteza foram expressos por *eu acho que* e *acho que*, em outra ANOVA, testamos se a intensidade foi relevante para diferenciar sentido e tipo de ocorrência. Tanto o sentido ($f(2, 1034) = 10.10, p < .001$) quanto o tipo de

ocorrência ($f(1, 1034) = 115.95, p < .001$) foram variáveis diferenciadas pela intensidade, ou seja: nas entrevistas sociolinguísticas, para indicar certeza os falantes emitiram um som mais forte que dúvida e incerteza, resultados semelhantes aos obtidos por Oliveira (2011).

5.2.3 Duração de *(eu) acho que*

A duração diz respeito ao tempo, em milissegundos, de emissão de uma oração, item lexical ou construção. De acordo com os estudos de Oliveira (2011); Antunes; Aubergé e Sasa (2014), quando incertos sobre algo os falantes tendem a alongar as sílabas das palavras, por isso a duração de orações que indicam dúvida e incerteza é maior, ao contrário daquelas que indicam certeza. Na amostra que analisamos, não encontramos resultados semelhantes: a média de *(eu) acho que* ao indicar incerteza foi maior (62.0). Diferente dos estudos sobre orações no português, na amostra que analisamos a média de certeza foi maior (53.4) do que a de dúvida (51.7).

Sentido	Variável prosódica	Nº	Média	Desvio padrão	Min	Max
Certeza	duracao_modalizador	815	53.4	24.7	14.2	187
Dúvida	duracao_modalizador	96	51.7	28.8	19.3	188
Incerteza	duracao_modalizador	127	62.0	33.3	20.8	197

Tabela 22: valores descritivos da média da duração do modalizador em relação ao sentido de *(eu) acho que*

Fonte: elaborada pela autora

A partir da distribuição dos dados registrados na tabela 22 calculamos por meio de uma análise de variância se a duração de *(eu) acho que* foi relevante para diferenciar certeza, dúvida e incerteza. Os resultados indicaram que sim: ($f(2, 1035) = 6.38, p = 0.002$).

Novamente, os referidos sentidos foram indicados por *eu acho que* e *acho que*, por isso executamos outra ANOVA para calcularmos se a duração diferenciou sentido e o tipo de ocorrência da construção analisada. Tanto o sentido ($f(2, 1034) = 6.99, p < 0.001$) quanto a ocorrência ($f(1, 1034) = 99.99, p < .001$) foram diferenciados pela duração, resultados semelhantes aos de Oliveira (2011); Antunes; Aubergé; Sasa (2014) para sentenças no português brasileiro. No caso de *(eu) acho que*, os resultados por nós obtidos sugerem que os falantes alongaram ou não a realização das sílabas a depender do sentido a ser indicado.

5.2.4 Duração das pausas silenciosas

Pausas silenciosas, ou seja, períodos de silêncio durante a fala também possibilitam a inferência de sentidos. Sua ocorrência, tanto em dados de laboratório quanto nos questionários do AliB (SILVA, 2008; OLIVEIRA, 2011; ANTUNES; AUBERGÉ; SASA, 2014; FERNANDES; ANTUNES, 2017) ocorreu apenas quando houve o sentido de incerteza, embora os resultados nem sempre tenham sido estatisticamente significativos. Na amostra Deslocamentos 2020 computamos pausas preenchidas com mais de 150 milissegundos de duração:

Sentido	Variável prosódica	Nº	Média	Desvio padrão	Min	Max
Certeza	duracao_pausa_silenciosa	11	158	32.0	107	221
Dúvida	duracao_pausa_silenciosa	5	158	34.2	118	210
Incerteza	duracao_pausa_silenciosa	12	199	52.0	118	300

Tabela 23: valores descritivos da duração das pausas silenciosas em relação ao sentido de *(eu) acho que*

Fonte: elaborada pela autora

A partir dos dados registrados na tabela 23, utilizamos uma análise de variância para calcularmos se a duração da pausa silenciosa diferenciou certeza, dúvida e incerteza. Os resultados indicaram que não ($f(2, 25) = 3.22, p = 0.057$).

Posteriormente, executamos outra ANOVA para calcularmos se a duração da pausa silenciosa diferenciou os sentido e ocorrência de *eu acho que* e *acho que* pelos falantes da amostra Deslocamentos 2020. Observamos que nem o sentido ($f(2, 24) = 3.23, p = 0.057$) nem o tipo de ocorrência ($f(1, 24) = 1.06, p = 0.313$) foram diferenciados pela duração desse tipo de pausa, o que pode ter relação com o seu pequeno número de ocorrências.

5.2.5 Duração das pausas preenchidas

Assim como os falantes silenciam por um breve período enquanto discorrem sobre algo, eles também fazem pausas em que enunciam sons vocálicos, que chamamos de pausas preenchidas. Na amostra Deslocamentos 2020 elas foram marcadas com “ah”, “eh”, “ih”, de acordo com as regras de transcrição adotadas para o banco de dados Falares Sergipanos. A seguir apresentamos a descrição dos resultados dessa variável:

Sentido	Variável prosódica	Nº	Média	Desvio padrão	Min	Max
Certeza	duracao_pausa_preenchida	52	61.6	26.3	25.8	159
Dúvida	duracao_pausa_preenchida	8	53.3	13.2	36.6	76.2
Incerteza	duracao_pausa_preenchida	8	80.6	33.1	41.2	123

Tabela 24: valores descritivos da duração das pausas preenchidas em relação ao sentido de *(eu) acho que*

Fonte: elaborada pela autora

A partir dos dados na tabela 24, calculamos se a duração das pausas preenchidas diferenciou certeza, dúvida e incerteza por meio de uma análise de variância. Os resultados indicaram que não ($f(2, 65) = 2.45, p = 0.094$). Em seguida, considerando as ocorrências *eu acho que* e *acho que*, executamos outra ANOVA para calculamos se a duração das pausas preenchidas diferenciou o sentido e o tipo de ocorrência da construção modalizadora. Nem o sentido ($f(2, 64) = 2.47, p = 0.093$), nem a ocorrência ($f(1, 64) = 1.43, p = 0.237$) foram diferenciados pela duração das pausas preenchidas. De acordo com Silva (2008); Antunes e Aubergé (2015, p. 216), sentenças que indicam certeza apresentam menor média de frequência fundamental, intensidade e duração, ao contrário daquelas que indicam dúvida e incerteza. Na amostra Deslocamentos 2020, as análises de variância apontaram que os sentidos de certeza, dúvida e incerteza indicados por *(eu) acho que* foram inferidos pelos parâmetros intensidade e duração. Certeza foi caracterizado como um sentido de maior intensidade e menor duração, ou seja, maior força e menor alongamento da construção modalizadora. Já dúvida e incerteza foram caracterizados como sentidos com menor intensidade e maior duração: quando não tinham certeza do que diziam, os falantes alongavam a realização de *eu acho que*, o que contribuiu para a diferenciação desses sentidos.

5.3 VARIÁVEIS PARALINGÜÍSTICAS: EXPRESSÕES FACIAIS

Apesar da relevância das variáveis linguísticas e acústicas para a inferência de certeza, dúvida e incerteza, nenhum procedimento de análise conseguiu explicar totalmente como esse processo ocorre. Estudos de prosódia gestual sinalizam que elementos paralingüísticos, ou seja, as expressões faciais são relevantes. Nesta subseção, apresentamos uma análise qualitativa dos padrões de expressões faciais de dois informantes: DAN2MF e FRA3FF, dos gêneros masculino e feminino, respectivamente, enquanto enunciaram *acho que* e *eu acho que*. As ocorrências foram selecionadas a partir da relevância dos tópicos (moradia e questões de gênero), tipo de ocorrência e experiência do informante (direta e indireta), fatores que de acordo

com o que expusemos nas seções anteriores exerceram influência no sentido indicado por (*eu acho que*). Os padrões de expressões faciais foram gerados a partir do output do OpenFace v. 2.2.0, que estimou a intensidade (*intensity*) dos movimentos dos músculos do rosto, reconheceu e classificou as expressões de raiva, nojo, medo, felicidade, deboche, neutra, tristeza e surpresa, de acordo com a proposta de Tejada et al (2021).

Segundo Ekman e Friesen (1975), toda a espécie humana demonstra reações a estímulos externos por meio de expressões faciais, que também atuam na manutenção das relações interpessoais (EKMAN, 2000) e contribuem para a inferência de sentidos. É o que sugerem os estudos de prosódia gestual sobre o holandês, português europeu e brasileiro (SWERTS ET AL, 2003; ANTUNES; AUBERGÉ; SASA, 2014; CRUZ, SWERTS, FROTA, 2017). Por isso, tomamos como hipótese que as expressões faciais, junto às variáveis linguísticas e parâmetros acústicos contribuíram para a inferência dos sentidos de certeza, dúvida e incerteza indicados por (*eu acho que*) nas entrevistas sociolinguísticas que analisamos. A partir dos resultados de Swerts et al (2003) esperamos observar nos resultados da amostra Deslocamentos 2020 uma expressão semelhante à *funny face* (figura 4, em 3.2), caracterizada pelas AUs (*action units*, movimentos dos músculos do rosto) 4, 15 e 17, presentes na expressão de tristeza. Elas foram ilustradas na imagem que expusemos em 3.2 e que reproduzimos a seguir:

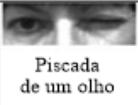
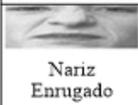
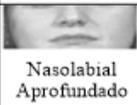
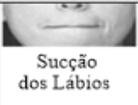
Unidades de Ação da Face Superior					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
 Sobrancelha Interna Levantada	 Sobrancelha Externa Levantada	 Sobrancelhas Baixas	 Pálpebra Superior Levantada	 Bochechas Levantadas	 Pálpebras Apertadas
 *AU 41 Abaixamento das Pálpebras	 *AU 42 Contração Retinal	 *AU 43 Olhos Fechados	 AU 44 Olhos Semicerrados	 AU 45 Piscada dos olhos	 AU 46 Piscada de um olho
Unidades de Ação da Face Inferior					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
 AU 9 Nariz Enrugado	 AU 10 Lábio Superior Levantado	 AU 11 Nasolabial Aprofundado	 AU 12 Cantos dos Lábios Puxados	 AU 13 Bochecha Inchada	 AU 14 Fazendo Covinhas
 AU 15 Cantos dos Lábios Pressionados	 AU 16 Lábio Inferiore Pressionado	 AU 17 Queixo Levantado	 AU 18 Lábios Enrugados	 AU 20 Lábios Esticados	 AU 22 Lábios Afunilados
 AU 23 Lábios Endurecidos	 AU 24 Lábios Pressionados	 *AU 25 Lábios Separados	 *AU 26 Mandíbula Caída	 *AU 27 Apertando a Boca	 AU 28 Sucção dos Lábios

Figura 14: facial action units (movimentos dos músculos do rosto)

Fonte: adaptado de <https://infograph.venngage.com/p/8769/microexpressions>

Quanto à expressão de certeza, a partir do que apontam os estudos de prosódia gestual que descrevemos em 3.2, tomamos como hipótese que ao indicar esse sentido o falante não demonstraria predomínio de nenhuma das AUs, pois os músculos do rosto estariam próximos a um estado de relaxamento. Não foi o que ocorreu no exemplo que destacamos a seguir:

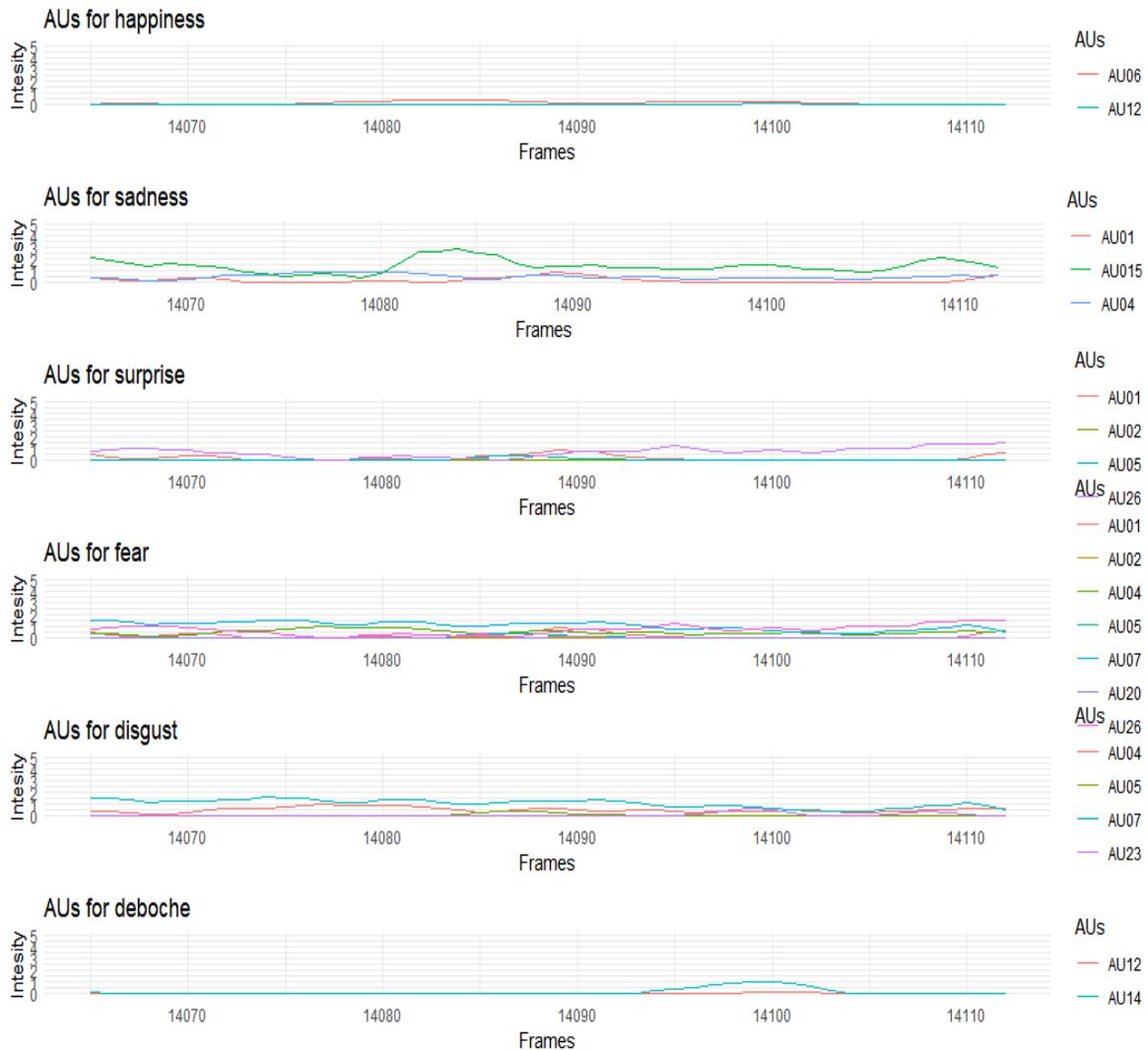


Figura 15: action units de DAN2MF (certeza)

Fonte: elaborada pela autora

Na figura 15, demonstramos os padrões das expressões faciais de DAN2MF ao usar *acho que* para indicar certeza ao proferir o excerto (27):

(27) DAN2MF: eu gosto bastante da cidade eh *acho que* por causa de tudo que eu já vivi lá.

As AUs que caracterizam felicidade, surpresa, medo, nojo e deboche (happiness, surprise, fear e disgust, respectivamente) apresentaram um comportamento bastante próximo, sem nenhum pico de proeminência. Houve, dentre os movimentos do rosto que caracterizam tristeza (sadness), predomínio da AU 15 no frame 14070, depois do frame 14080 ao 14110, que

demonstra curvatura dos lábios. Esses movimentos poderiam estar associados tanto a um sentimento do informante – o que não esperávamos, dada a função indicada por *acho que* e o tipo de assunto sobre o qual se falou – quanto à articulação da boca para enunciar as palavras.

Na figura 16, o falante demonstrou dúvida quanto às cidades que já viajou: talvez houvesse outras, mas ele não tinha certeza:

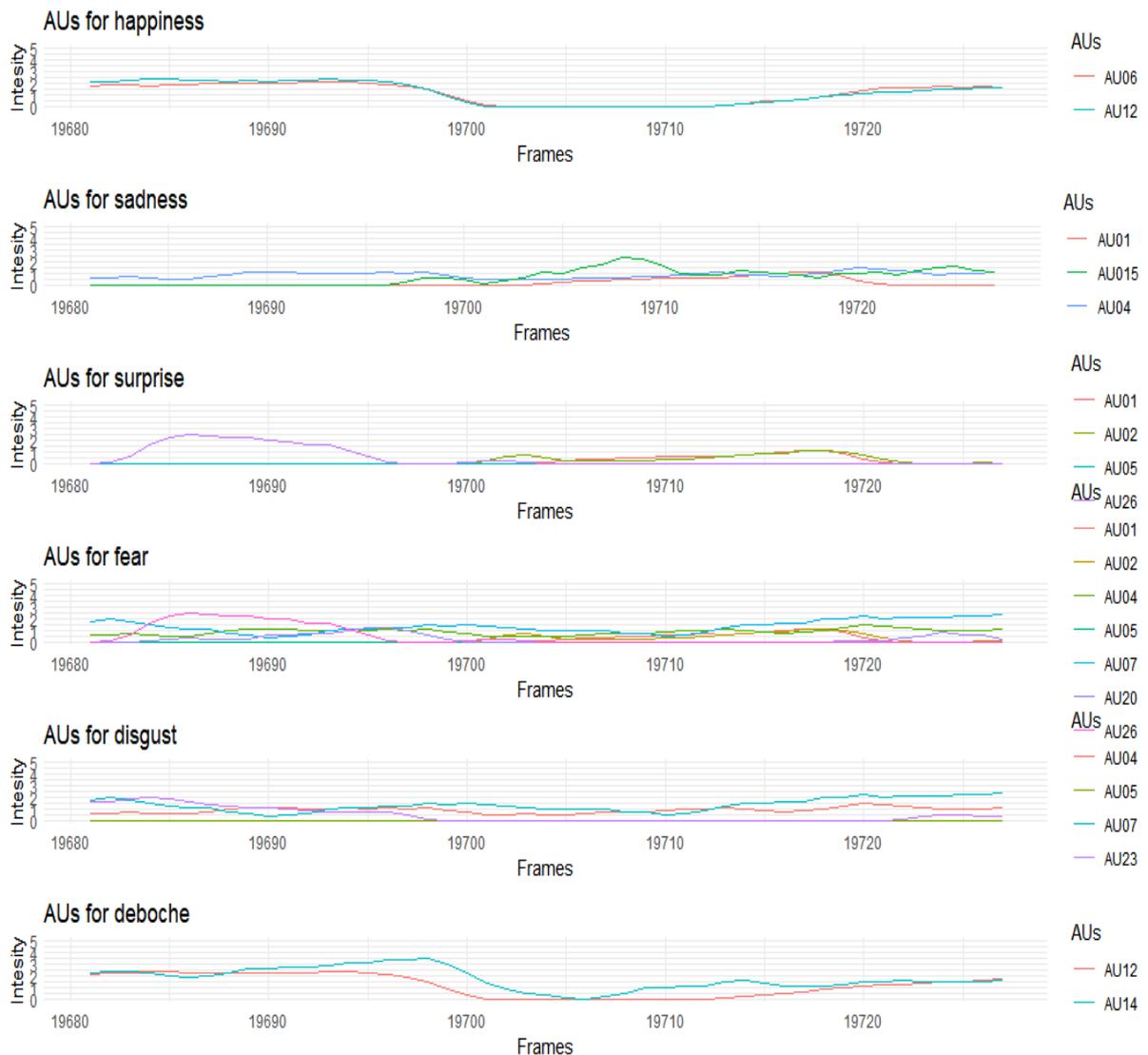


Figura 16: action units de DAN2MF (dúvida)

Fonte: elaborada pela autora

(28) DOCMS: quais lugares aqui em Sergipe você já visitou?

DAN2MF: cidades eu já fui Ribeirópolis, Aracaju, Socorro eh Campo do Brito Areia Branca hum não não sei *acho que* é isso

Ao indicar um aspecto experienciado diretamente por ele observa-se, nas AUs que caracterizam a expressão de tristeza (*sadness*), predominância da 4, do frame 19690 a 19700, e 15, do frame 19701 a 19710, que correspondem à contração da linha da sobrancelha,

movimento que para Swerts et al (2003); Antunes, Aubergé, Sasa (2014); Cruz, Swerts e Frota (2017) caracterizam ausência de certeza. Neste trabalho, assumimos que dúvida e incerteza, não são sentidos sinônimos: o primeiro não permite gradação, pois diz respeito a experiências pessoais, ao contrário do segundo. Por isso, esperávamos observar entre as duas um comportamento dos músculos da face com diferenças sutis:

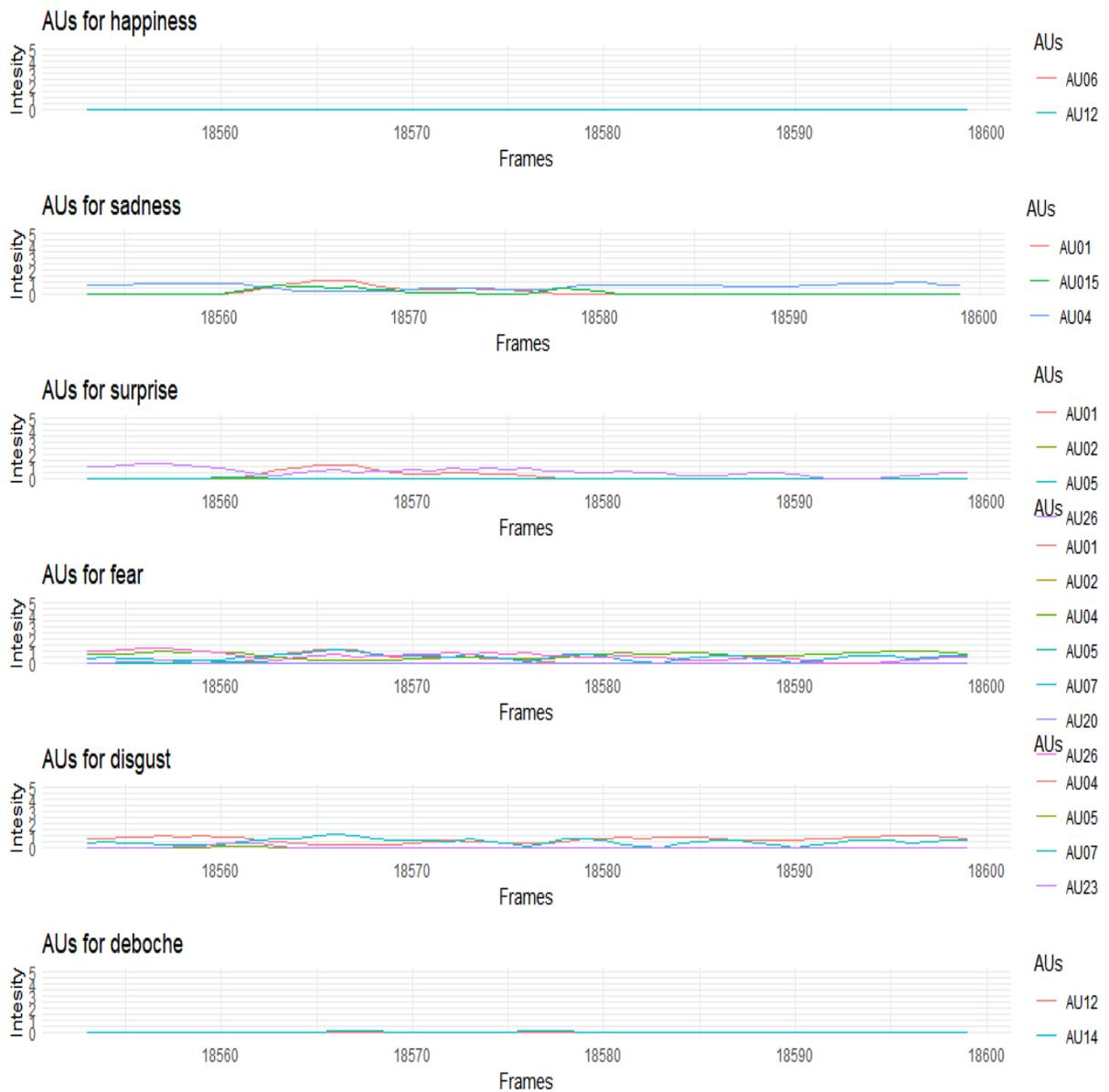


Figura 17: action units de DAN2MF (incerteza)

Fonte: elaborada pela autora

Dentre as AUs que caracterizam a expressão de tristeza (sadness) na figura 17, no frame 18560, depois dos frames 18580 a 18600 houve predomínio AU 15, que indica contração dos lábios, proeminência não captada na descrição das demais emoções. Nesse contexto, o falante

emitiu uma suposição sobre o que os moradores de outra cidade pensam sobre os problemas do município de Itabaiana:

(29) DOCMS: e em contraparte o que é menos atrativo para as pessoas lá qual o ponto de certa forma chato lá em Itabaiana?
 DAN2MF: (...) basicamente se aproximando mas é uma coisa que acho que os principais problemas pra quem mora lá.

Os padrões demonstrados nas figuras 16 e 17 para tristeza em comparação com a expressão de AUs nas demais emoções podem ser interpretados como gradação da *funny face*, do mesmo modo que há uma gradação entre dúvida e incerteza.

Ao contrário do que ocorreu na expressão de certeza por DAN2MF, na análise dos dados da informante FRA3FF ao utilizar *eu acho que* para demonstrar certeza sobre o tópico moradia observamos movimentos dos músculos do rosto esperados:

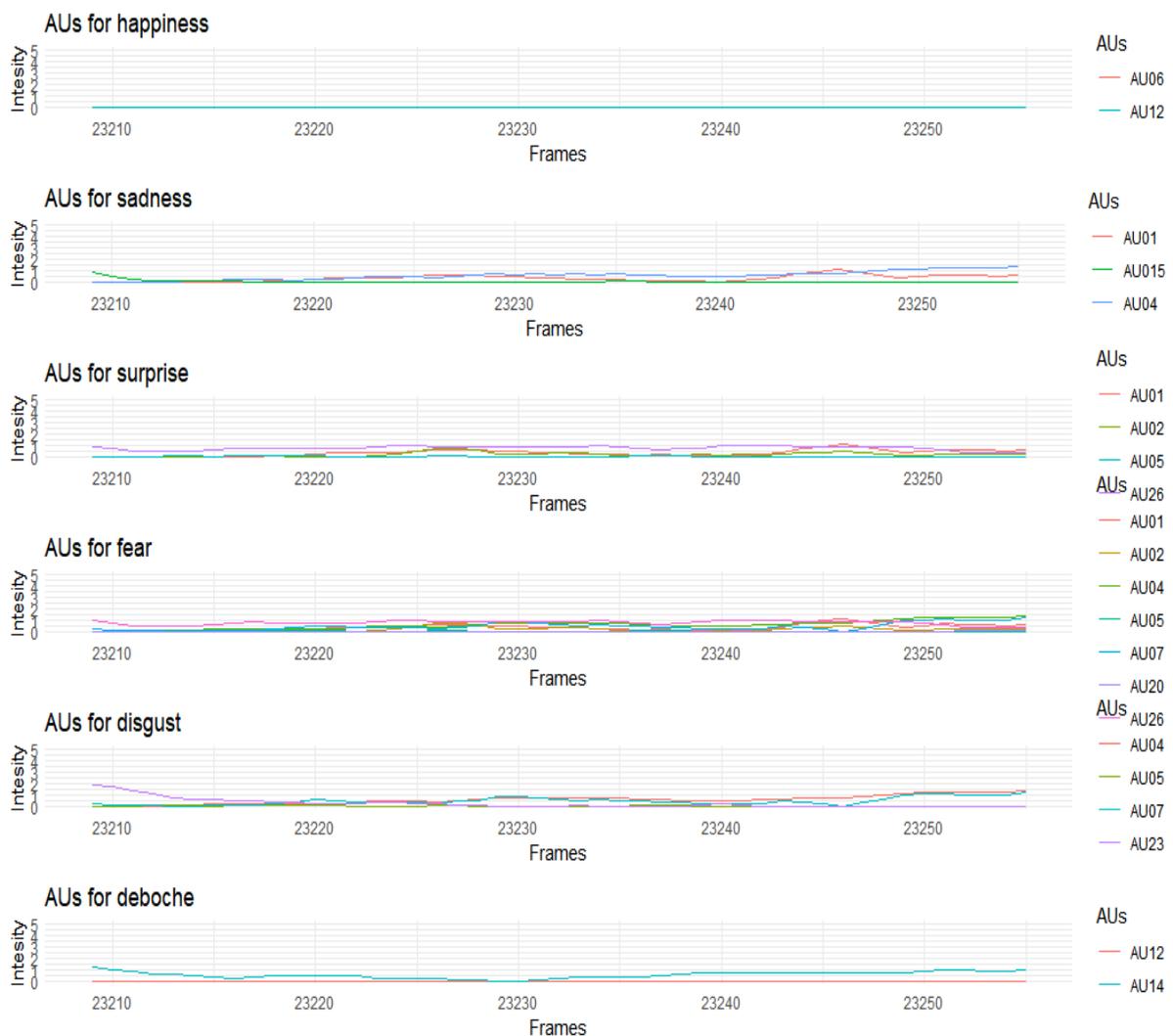


Figura 18: action units de FRA3FF (certeza)
Fonte: elaborada pela autora

As AUs demonstradas na expressão de certeza ilustrada na figura 18 foram registradas enquanto a informante fez um comentário sobre o que seria desagradável para os moradores da cidade em que ela reside: a falta de emprego para recém-formados:

(30) FRA3FF: vai aparecendo aos poucos mas não é uma coisa imediata
aí *eu acho que* seria isso.

Esses resultados aproximam-se do que esperávamos: nenhum movimento dos músculos do rosto se sobressaiu de modo a indicar reação muscular ao que havia sido perguntado. Para essa informante, também encontramos um comportamento esperado na expressão de dúvida:

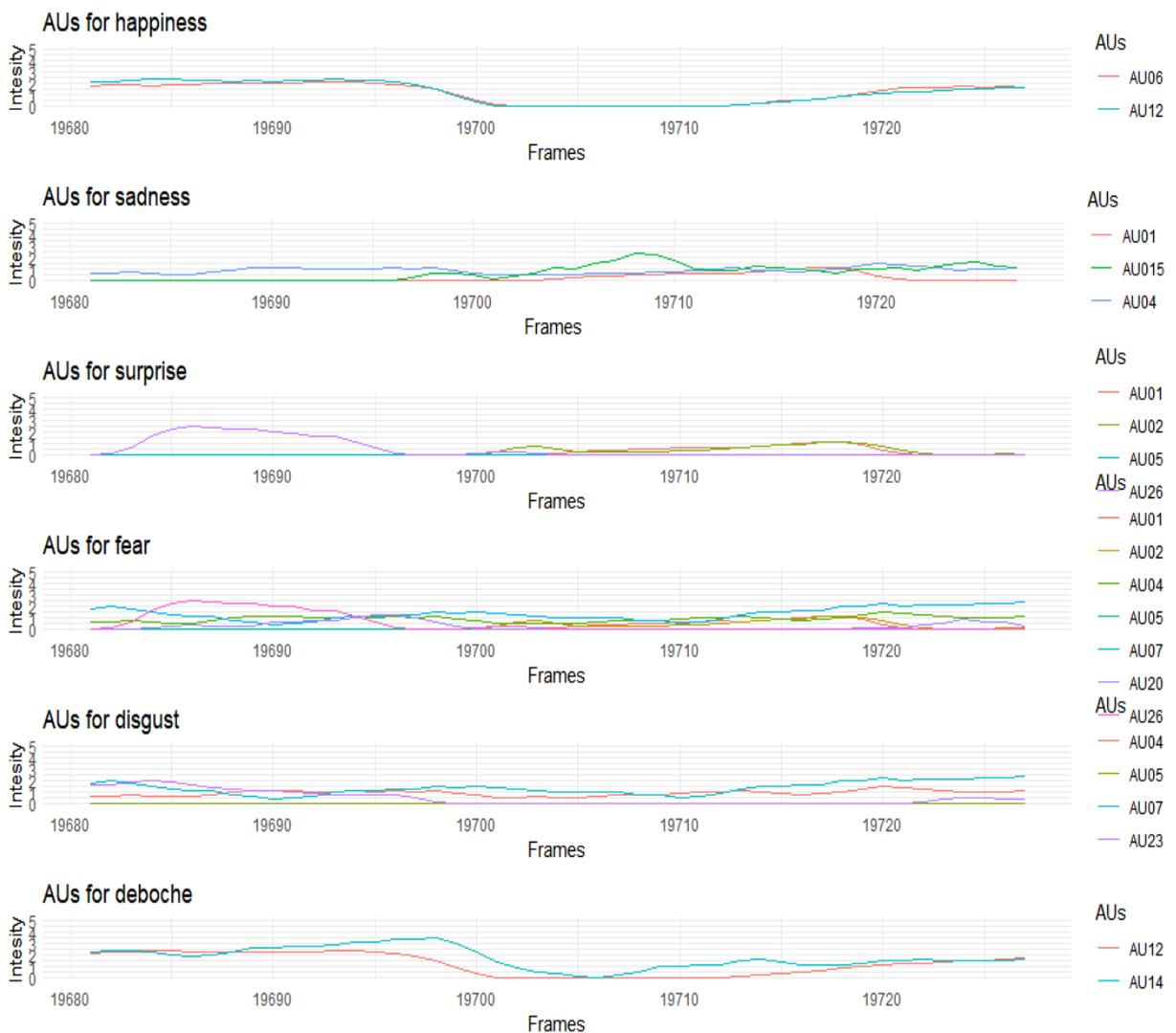


Figura 19: action units de FRA3FF (dúvida)

Fonte: elaborada pela autora

No excerto acima, FRA3FF falava sobre o crescimento da cidade em que viveu e a dificultosa oferta de vagas de emprego:

(31) FRA3FF: *eu acho que* não seria muito

Apesar de haver na expressão de tristeza, entre os frames 19580 e 19690, um pico da AU 4 que indica contração da linha das sobrancelhas e no frame 19710 um pico da AU 15 que corresponde à contração do lábio inferior, ao compararmos os resultados dessa expressão com as demais podemos observar que também houve picos das AUs que caracterizam as expressões de surpresa e deboche, o que não esperávamos dado o tópico discursivo sobre o qual a informante falou.

Para FRA3FF, a expressão de incerteza foi a que esteve mais próxima do que hipotetizamos que aconteceria:

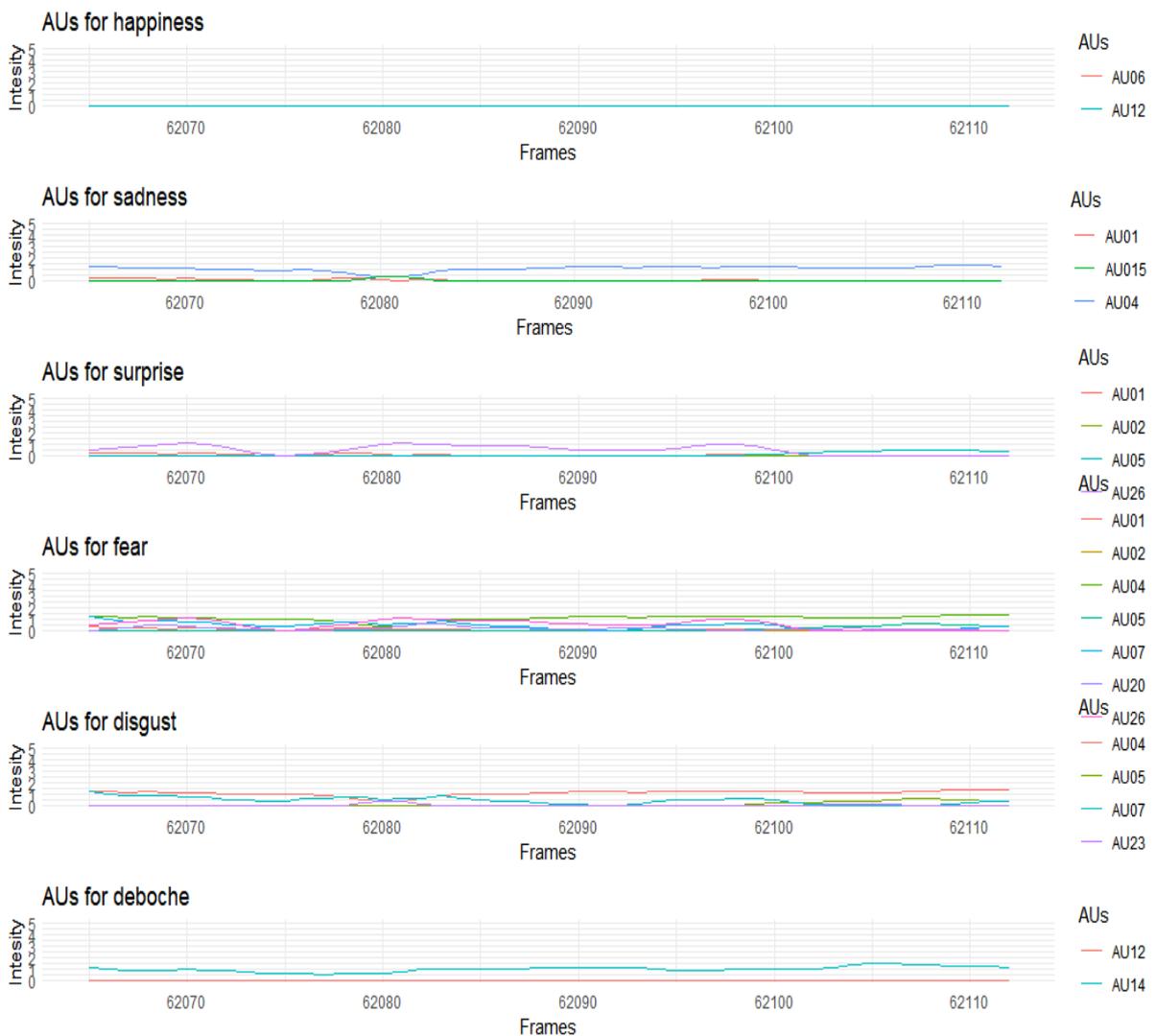


Figura 20: action units de FRA3FF (incerteza)
Fonte: elaborada pela autora

A figura 20 representa o momento em que a informante falou sobre questões de gênero, tópico discursivo que também exerceu forte influência na categorização das funções indicadas por *eu acho que*:

(32) DOCMS: a gente vê cada vez mais na mídia eh pessoas da comunidade (LGBTQIA+) tendo mais espaço tipo Pablo Vittar Lara Goove eh tipo Thammy Gretchen etc o que você acha que tem levado essas pessoas a terem mais espaço na mídia brasileira?
FRA3FF: *eu acho que* vai muito de liberdade interior né?

Dentre as AUs que caracterizam a expressão de tristeza, do frame 62080 ao 62110 houve predomínio da 4, que indica contração das sobrancelhas, movimento esperado para a expressão de incerteza. Contudo, nota-se que houve proeminência de AUs que indicam deboche, resultado inesperado.

Dentre os dados de expressões faciais de DAN2MF e FRA3FF encontramos ocorrência das AUs que caracterizam o que Swerts et al (2003) definiu como *funny face* e que dá indícios de dúvida e incerteza. Todavia, houve também picos de movimentos dos músculos do rosto em expressões que não esperávamos, como deboche. Esses resultados sugerem a influência de variáveis que na presente análise não puderam ser controladas e evidenciam a necessidade de expandirmos a exploração dos dados de expressões faciais de todas as entrevistas gravadas em vídeo, a fim de observarmos se, nas entrevistas sociolinguísticas, os movimentos dos músculos do rosto atuaram de modo significativo na inferência dos sentidos de certeza, dúvida e incerteza indicados por *(eu) acho que*.

6 CONSIDERAÇÕES FINAIS

(*Eu acho que*) é uma construção modalizadora parentética epistêmica que, em decorrência dos processos de mudança semântica e gramaticalização, se tornou polissêmica e passou a indicar certeza, dúvida e incerteza, sentidos diferenciados pela experiência direta e indireta do falante em relação ao que é dito. Nos estudos linguísticos, o sentido (*eu acho que*) é descrito a partir de variáveis estruturais e subjetivas: tipo de ocorrência, escopo, tópico discursivo e experiência GALVÃO, 1999; GONÇALVES, 2003; FREITAG, 2003). Nesta pesquisa, tivemos como objetivo ampliar o escopo de análise da referida construção por meio da inclusão de parâmetros acústicos e paralinguísticos (expressões faciais).

Do ponto de vista acústico, sentenças que indicam dúvida e incerteza apresentam maior tempo de latência, ou seja, os falantes falam mais devagar, fazem pausas e produzem maiores valores de frequência fundamental e duração, ao contrário das que indicam certeza (SILVA, 2008; OLIVEIRA, 2011; FERNANDES; ANTUNES, 2017). Certeza, dúvida e incerteza também apresentam correlatos gestuais: enquanto o primeiro sentido é caracterizado por uma expressão facial de neutralidade, os dois últimos são caracterizados pela contração da linha das sobrancelhas e boca (SWERTS ET AL, 2003; ANTUNES; AUBERGÉ; SASA, 2014; CRUZ, SWERTS, FROTA, 2017). Neste trabalho, enfatizamos um item lexical específico: a construção modalizadora parentética epistêmica (*eu acho que*) em entrevistas sociolinguísticas da amostra Deslocamentos 2020, realizadas com estudantes de graduação da Universidade Federal de Sergipe, *campus* Professor José Aloísio de Campos.

A amostra que analisamos foi constituída por 30 entrevistas da amostra Deslocamentos 2020 gravadas em áudio e vídeo, a partir das quais computamos 1038 ocorrências de (*eu acho que*). Testes com estatística qui-quadrado indicaram que o sentido dessa construção estava associado ao tipo de ocorrência (*eu acho que*, *acho que*) ($\chi^2 = 10.011 \cdot df = 2 \cdot \text{Cramer's } V = 0.098 \cdot p = 0.007$), ao escopo interno ou externo ($\chi^2 = 11.826 \cdot df = 2 \cdot \text{Cramer's } V = 0.107 \cdot p = 0.003$), à presença de outro modalizador ($\chi^2 = 14.329 \cdot df = 2 \cdot \text{Cramer's } V = 0.117 \cdot \text{Fisher's } p = 0.002$), ao tópico discursivo ($\chi^2 = 93.766 \cdot df = 14 \cdot \text{Cramer's } V = 0.213 \cdot \text{Fisher's } p = 0.000$) e à experiência do falante ($\chi^2 = 214.376 \cdot df = 2 \cdot \text{Cramer's } V = 0.454 \cdot p = 0.000$), resultados que corroboram o que estudos anteriores sugeriram (GALVÃO, 1999; FREITAG, 2003; GONÇALVES, 2003). Ou seja: para indicar maior grau de certeza, os falantes demarcaram seu posicionamento por meio do uso do sintagma nominal “eu”. Isso ocorreu principalmente quando eles falaram das próprias experiências e de assuntos pessoais, demonstrando que o

sentido da construção modalizadora que investigamos carrega traços de evidencialidade, parte da relação do informante com o que foi dito, conforme Freitag (2003) pontuou. Ainda assim, entendemos que para compreendermos como os diferentes sentidos de certeza, dúvida e incerteza indicados por *(eu) acho que* são inferidos, seria necessário considerar fatores tradicionalmente não incluídos na análise dessa construção: acústicos e de expressões faciais.

Estudos de prosódia acústica que descreveram a realização de sentenças que indicam certeza, dúvida e incerteza no português brasileiro controlaram como variáveis medidas de frequência fundamental, intensidade, duração, mais a ocorrência de pausas silenciosas e preenchidas, parâmetros que diferenciam cada um dos sentidos que analisamos neste trabalho (SILVA, 2008; OLIVEIRA, 2011; FERNANDES; ANTUNES, 2017). Esses mesmos sentidos também são indicados pela construção modalizadora parentética epistêmica *(eu) acho que*. Por isso, tomamos como hipótese que variáveis acústicas junto às linguísticas, atuariam de modo integrado na inferência dos sentidos indicados por *(eu) acho que*. Para analisarmos os sentidos indicados por *(eu) acho que* quanto à frequência fundamental, intensidade e duração dessa construção modalizadora mais duração de pausas silenciosas e preenchidas, realizamos análises de variância multiníveis. Os resultados indicaram que somente intensidade ($f(2, 1035) = 8.88$, $p < .001$) e duração ($f(2, 1035) = 6.38$, $p = 0.002$) de *(eu) acho que* diferenciaram seus sentidos. Ao indicar certeza, a intensidade, ou seja, a força dessa construção foi maior e sua duração foi menor, ao contrário de quando ela indicou dúvida e incerteza: nestes casos, os falantes falaram com menos força e alongaram as sílabas de *(eu) acho que*.

Do ponto de vista gestual, a análise das expressões faciais de DAN2MF e FRA3FF demonstrou que há correlatos entre os sentidos de dúvida e incerteza: na expressão deles houve ocorrência da contração das sobrancelhas (AU4) e do lábio inferior (AU15), resultados semelhantes aos encontrados por Swerts et al (2003), Antunes; Aubergé; Sasa (2014), Cruz; Swerts; Frota (2017). Contudo, a análise qualitativa que executamos demonstraram a necessidade de aumentarmos o poder explanatório dos dados, considerando que também houve ocorrência de AUs que não prevíamos, como a de deboche, possivelmente em decorrência de variáveis que não controlamos. Ainda assim, as análises linguística, acústica e gestual sugeriram um funcionamento integrado e sistematizado entre essas variáveis, o que proporcionou uma ampliação do escopo de análise de *(eu) acho que* e delineou a potencialidade de estudos descritivos que incluem expressões faciais como fator relevante para a inferência de sentidos.

REFERÊNCIAS

- ANTUNES, L. B.; AUBERGÉ, V.; SASA, Y. Certainty and uncertainty in Brazilian Portuguese: methodology of spontaneous corpus collection and data analysis. *Proceedings of the 7th Speech Prosody. Dublin, Trinity College*, p. 110-114, 2014.
- ANTUNES L. B.; AUBERGÉ, V. Análise prosódica da certeza e da incerteza em fala espontânea e atuada. *Revista Diadorim*, v. 17, n. 2, p. 212-237, 2015.
- BARBOSA, P. A. *Prosódia*. Editora Parábola, 2019.
- BARBOSA, P. A.; MADUREIRA, S. *Manual de fonética acústica experimental: aplicações a dados do português*. Cortez editora, 2015.
- BOERSMA, P.; WEENINK, D. *Praat: Doing phonetics by computer* (Versão 6.0.33). 2017. Disponível em: <<http://www.praat.org/>>. Acesso em: 5 de jan. de 2020.
- BYBEE, J.; PERKINS, M. R.; PAGLIUCA, W. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press, 1994.
- BYBEE, J. *Língua, uso e cognição*. Trad. Maria angélica Furtado da Cunha. São Paulo: Cortez, 2016.
- COUPER-KUHLEN, E. *An introduction to English Prosody*. Tübingen: Niemeyer, 1986.
- CRUTTENDEN, A. *Intonation*. Cambridge: Cambridge University Press, 1986.
- CRUZ, M.; SWERTS, M.; FROTA, S. The role of intonation and visual cues in the perception of sentence types: Evidence from European Portuguese varieties. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, v. 8, n. 1, 2017.
- CRYSTAL, D. *Prosodic systems and intonation in english*. Cambridge: Cambridge University Press, 1969.
- CUNHA, M. A. F.; OLIVEIRA, M. R.; VOTRE, S. A interação sincronia/diacronia no estudo da sintaxe. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, v. 15, n. 1, p. 00-00, 1999.
- EKMAN, P.; FRIESEN, W.V. *Unmaking the face: a guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ, Prentice Hall, 1975.
- EKMAN, P. Basic emotions. *Handbook of cognition and emotion*, v. 98, n. 45-60, p. 16, 2000.
- EKMAN, P.; FRIESEN, W. V. *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk, 2003.
- EKMAN, P.; CORDARO, D. What is meant by calling emotions basic. *Emotion review*, v. 3, n. 4, p. 364-370, 2011.
- Expressão de dúvida. Disponível em: <https://br.freepik.com/fotos-vetores-gratis/duvida>. Acesso em: 20 de jan, de 2021.

- FERNANADES, L. D. P.; ANTUNES, L. B. O papel da prosódia na expressão da certeza e da incerteza em português brasileiro: respostas ao questionário ALiB em cidades do norte, nordeste e sul. In: *Anais do Congresso Brasileiro de Prosódia*, 2017.
- FREITAG, R. M. K. Gramaticalização e variação de acho (que) e parece (que) na fala de Florianópolis. Dissertação (Mestrado em Linguística), 2003.
- FREITAG, R. M. K. Mudar para variar, variar para mudar: tratando da variação e mudança de acho (que) e parece (que) parentéticos epistêmicos na fala de Florianópolis. *Fórum Linguístico*, v. 4, n. 1, p. 81-113, 2007a.
- FREITAG, R. M. K. A expressão do passado imperfeito no português: variação/gramaticalização e mudança. Tese (Doutorado em Linguística), 2007b.
- FREITAG, R. M. K. Banco de dados falares sergipanos. *Working Papers em Linguística*, v. 14, n. 2, p. 156-164, 2013.
- FREITAG, R.M.K. ET AL. Julgamentos de traços linguísticos e expressões faciais: uma abordagem do processamento da variação. *Cadernos de Linguística*, 2020.
- FREITAG, R. M. K. *Variáveis categóricas*. Disponível em: <https://rkofreitag.github.io/Categorica.html/>. Acesso em: 2020-10-07.
- GALVÃO, V. C. C. O achar no português no Brasil: um caso de gramaticalização. Dissertação (Mestrado em Linguística). Instituto de Estudos da Linguagem da Universidade Estadual de Campinas. Campinas, 1999.
- GONÇALVES, S. C. L. ET AL. Gramaticalização, modalidade epistêmica e evidencialidade: um estudo de caso no português do Brasil. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem da Universidade Estadual de Campinas. Campinas, 2003.
- GRICE, H. P. Logic and conversation. In: *Speech acts*. Brill, 1975. p. 41-58.
- HALLIDAY, M. A. K. *A course in spoken English: intonation*. London: Oxford University Press, 1970.
- HEINE, B.; CLAUDI, U.; HÜNNEMEYER, F. *Grammaticalization: A conceptual framework*. University of Chicago Press, 1991.
- HIST, D.; DI CRISTO, A. A survey of intonation systems. In: HIST, D; DI CRISTO, A. *Intonation Systems: A Survey of Twenty Languages*. New York: Cambridge University Press, 1998. cap. 1, p. 1-44.
- LABOV, W. *Padrões sociolinguísticos*. São Paulo: Parábola Ed., 2008.
- KANADE, T.; COHN, J.; TIAN, Y. *Comprehensive database for facial expression analysis*. In: Proceedings of the Fourth IEEE, v. 1, 2000, France. International Conference on Automatic Face and Gesture Recognition. France: Grenoble, p. 46-53.
- KOWAL, S.; WIESE, R.; O'CONNELL, D. C. The use of time in storytelling. *Language and Speech*, v. 26, n. 4, p. 377-392, 1983.

- LAVER, J. *The phonetic description of voice quality*. Cambridge University Press, 1980.
- LEHMANN, C. *Thoughts on grammaticalization: A programmatic sketch*. Universalienprojekt, Institut für Sprachwissenschaft Universität, 1982.
- MACHADO, A. Descartes e a psicologia da dúvida. Colóquio Descartes da Academia Brasileira de filosofia. Faculdade da Cidade. Rio de Janeiro: 1996.
- MARTINS, F. Bases linguísticas em sistemas de reconhecimento de fala. In: RANCHO, E. M. (Org). *Tratamento das línguas por computador: uma introdução à linguística computacional e suas aplicações*. Lisboa: Editorial Caminho, 2001, p.295-229.
- MCDUGALL, W. *An Introduction to Social Psychology*. Boston: Luce, 1926.
- MIGUEL, F. K. Psicologia das emoções: uma proposta integrativa para compreender a expressão emocional. *Psico-usf*, v. 20, n. 1, p. 153-162, 2015.
- MORAES, J. A. From a prosodic point of view: remarks on attitudinal meaning. *From a Prosodic Point of View: Remarks on Attitudinal Meaning*, p. 19-37, 2011.
- NASCENTES, A. *Dicionário Etimológico da Língua Portuguesa*, vol. 1. Rio de Janeiro, 1955.
- OLIVEIRA, B. F. V. de. A prosódia na expressão das atitudes de dúvida, incerteza e incredulidade no português brasileiro. Dissertação (Mestrado em Linguística), 2011.
- OLIVEIRA Jr., M. Aspectos técnicos na coleta de dados linguísticos orais. In: FREITAG, R. M. K. (Organizadora). *Metodologia de Coleta e Manipulação de Dados em Sociolinguística*, São Paulo: Editora Edgard Blücher, 2014.
- OUSHIRO, L. Transcrição de entrevistas sociolinguísticas com o ELAN. In: FREITAG, R. M. K. *Metodologia de Coleta e Manipulação de Dados em Sociolinguística*. São Paulo: Blucher, p. 46-50, 2014.
- PIOT, O.; LYAGHAT, M. Expression et reconnaissance de onze attitudes assertives et interrogatives en Persan standard. In: *Speech Prosody 2002, International Conference*. 2002.
- PLUTCHIK, R. A general psychoevolutionary theory of emotion. In: *Theories of emotion*. New York: Academic Press, 1980. p. 3-33.
- ROACH, P; ET AL. Transcription of Prosodic and Paralinguistic Features of Emotional Speech. *Journal of the International Phonetic Association*, 1998.
- ROSA, M. M. *Marcadores de atenuação*. Contexto, 1992.
- SCHÖTZ, S. Linguistic & paralinguistic phonetic variation in speaker recognition & text-to-speech synthesis. In: In Speech Technology, GSLT, 2002. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.104.1401>> Acesso em: 25 de mar. de 2020>
- SEARLE, J. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press, 1969.
- SEARLE, Jonh R. *Intencionalidade*. São Paulo: Martins Fontes, 1995.

- SILVA, J. P. G. *Análise dos aspectos prosódicos na expressão da certeza e da dúvida no português brasileiro*. Dissertação (Mestrado em Linguística), Universidade Federal de Minas Gerais, Belo Horizonte, 2008.
- SWERTS, M. ET AL. Audiovisual cues to uncertainty. In: *ISCA Tutorial and Research Workshop on Error Handling in Spoken Dialogue Systems*. 2003.
- TEJADA ET AL. Building and validation of a set of facial expression images to detect emotions: a transcultural study, 2021 (no prelo).
- TOMKINS, S. S. Affect theory. *Approaches to emotion*, v. 163, n. 163-195, 1984.
- TEASDALE, J. D. Multi-level theories of cognition – emotion relations. *Handbook of cognition and emotion*, v. 98, p. 665-682, 1999.
- TRAUNMÜLLER, H. Evidence for demodulation in speech perception. In: *Sixth International Conference on Spoken Language Processing*, 2000. Disponível em: <https://www.isca-speech.org/archive/icslp_2000/i00_3790.html> Acesso em: 31 de mar. de 2020.
- TRAUGOTT, E. C.; DASHER, R. B. *Regularity in semantic change*. Cambridge University Press, 2004.
- VIANA, M. Síntese da fala. In: RANCHHOD, E. M. (Org). *Tratamento das línguas por computador: uma introdução à linguística computacional e suas aplicações*. Lisboa: Editorial Caminho, 2001, p. 133-193.
- VIOLA, P.; JONES, M. *Rapid object detection using boosted cascade of simple features*. IEEE Conference on Computer Vision and Pattern Recognition, 2001.

6.1 ANEXO 1

ROTEIRO PARA ENTREVISTA SOCIOLINGUÍSTICA**Perguntas de checagem**

1. Qual o seu nome?
2. Qual a sua idade?
3. Em que cidade/UF você nasceu?
4. Onde você mora atualmente?
5. Quais os nomes dos seus pais?
6. Em que cidade/UF eles nasceram? Onde eles vivem?
7. Qual o nível de escolaridade deles?
8. Quais as profissões deles?
9. Com quem você mora?
10. Qual o curso você faz? Está em qual período?
11. Por que você veio estudar na Universidade Federal de Sergipe?
12. Qual meio de transporte você mais utiliza para vir para a universidade? Por quê?
13. Caso seja um informante que more no interior e que venha de ônibus: o ônibus que você pega é público ou locado pela associação de estudantes da sua cidade?
14. Você trabalha ou já trabalhou? Como foi essa experiência?
15. Você acha que trabalhar e estudar pode dar certo?

Moradia

16. Há quanto tempo mora no bairro/cidade X?
17. Sempre morou nesse lugar?
18. Como foi sua infância onde você mora (ou morou)? Mudou muito em relação aos dias de hoje?
19. Qual a memória mais marcante de sua infância nesse lugar?
20. Você gosta de morar lá?
21. Se tivesse oportunidade, moraria em outro lugar? Qual? Por quê?
22. O que você mais gosta de fazer no local onde mora?
23. O que é atrativo para os moradores da comunidade?
24. O que é chato para os moradores da comunidade? (pedir para o informante falar sobre problemas da localidade)
25. Você já viajou para fora do seu estado?
26. Quais os lugares que você já visitou? Como foi? Por que você foi?
27. Gostaria de voltar?
28. Tem alguma viagem que você sonha em fazer? Qual? Por quê?

Lazer

29. O que você gosta de fazer no final de semana?
30. Quais lugares você costuma frequentar por lazer?
31. Você costuma frequentar cinema, museu, teatro?
32. Gosta de ler livros? Qual seu favorito? Por quê? Pode resumir?

33. Ler algo pela internet, quais sites?
34. Sua família tem hábito de ler?
35. Você acha o hábito de leitura importante? Por quê?

Educação

36. Você acha que a educação pública deve ser assegurada pelo governo? Por quê?
37. A educação pública melhorou ou piorou nos últimos anos? Por quê?
38. Você considera importante que as bolsas (como PIBIC, PIDID, bolsas auxílios) ofertadas na universidade tenham como critério de escolha a vulnerabilidade socioeconômica para pessoas com baixa renda?
39. Você considera importante às políticas para garantir o acesso e permanência dos alunos na educação básica: transporte escolar, plano nacional do livro didático, merenda escolar? Por quê?
40. Quais pontos negativos e os positivos os quais você observa na educação pública?
41. O Brasil investe o suficiente ou não em educação?
42. O que você pensa sobre o governo federal financiar o acesso de estudantes ao ensino superior em instituições privadas?
43. Muitos jovens ingressam na universidade por meio dos programas PROUNI e o FIES, você considera que esses programas diminuí a desigualdade social entre os jovens?
44. Você concorda ou não que a política de cotas para ingresso de negros/as na universidade diminuí a desigualdade entre brancos e negros?
45. Como foi a experiência de começar um curso superior?
46. Você sentiu ou sente alguma dificuldade em seu curso?
47. Por que você escolheu esse curso?
48. A infraestrutura da UFS atende às necessidades do seu curso?
49. O que você pensa em fazer depois que terminar a graduação? Pretende ingressar no mercado de trabalho de imediato ou fazer uma especialização ou mestrado? Por quê?

Segurança pública

50. O bairro em que você mora é violento?
51. Você concorda ou não com o ditado “Bandido bom é bandido morto”?
52. Qual o papel da segurança pública?
53. A segurança pública [de Aracaju/São Cristóvão] tem conseguido proteger os cidadãos?
54. A polícia é ágil? Consegue atender as demandas da população?

Políticas de igualdade de gênero

55. Você concorda com o ditado “Se as mulheres não usassem roupas tão curtas haveria menos violência contra elas”?
56. Você acha que as mulheres atualmente “tem conquistado muito espaço na sociedade”, por quê?
57. A violência atinge homens e mulheres da mesma forma?
58. É recorrente de uns tempos pra cá ouvirmos a frase: “a comunidade de gays e lésbicas tem ganhado muitos privilégios”. Você concorda com essa frase ou não? Por quê?
59. Recentemente, o Supremo Tribunal Federal decretou que a homofobia é crime. Você concordou com “essa conquista” de toda comunidade lgbtqi+ ?
60. A questão de gênero tem ganhado muita evidência na mídia ultimamente. Dragqueens cantoras, como Pablio Vittar, mulheres e homens transexuais, como o Tamy Miranda, a

filha que agora é filho da Gretchen, gays já tem aparecido em novelas. Quais motivos você considera que levou essas pessoas terem visibilidade nos meios de comunicação?

61. Você considera a discussão de gênero ou até mesmo o empoderamento feminino necessários ou mimimi?

Saúde pública

62. Já teve algum problema de saúde? Qual?
63. A saúde pública é um direito que deve ser assegurado pelo governo?
64. A saúde pública no Brasil funciona?
65. Como você avalia o Sistema Único de Saúde (SUS)?
66. Quais os pontos negativos e os positivos do SUS?
67. Há problemas no funcionamento do SUS? Quais?
68. Como o SUS poderia ser melhorado?
69. A quantidade de médicos (e especialistas) no SUS é suficiente para atender a demanda da população?
70. No bairro em que você mora há unidade de saúde da família?
71. Você já precisou usar algum serviço do SUS? Como você descreve a experiência?

Teste de percepção linguística

72. Aqui no estado de Sergipe, você já percebeu alguma diferença no modo como as pessoas falam? Você seria capaz de imitar (caracterizar) essa fala?
73. Você acredita que, estando em outro lugar ou em contato com pessoas de outro lugar outras pessoas podem identificar de onde você é (a sua origem) apenas pela maneira como você fala? Por quê?
74. Você acha que as pessoas que moram na capital falam diferente das que moram no interior? Como?
75. O que mais chama sua atenção quando as pessoas estão falando?
76. Você acha que seu modo de falar mudou depois que você entrou na universidade?
77. Falar **otfo, mutfo, biscofto, otfenta** é comum onde você mora?
78. Falar essas mesmas palavras é comum aqui na universidade?
79. O que você acha desse jeito de falar?
80. Falar **pranta; broco; praneta e franela** é comum onde você mora?
81. Você acha que esse jeito de falar é típico de alguma localidade?
82. Você acha que esse modo de falar tem alguma relação com a escolarização da pessoa?
83. Falar **tjia,dzia, dentje,médzico** é comum onde você mora?
84. Falar essas mesmas palavras é comum aqui na universidade?
85. O que você acha desse jeito de falar?
86. Você acha que esse jeito de falar é típico de alguma localidade?
87. Você acha que esse modo de falar tem alguma relação com a escolarização?
88. Você acha que esse modo de falar é considerado mais bonito?
89. Você acha que **dadzo e doido** tem a mesmo significado?
90. Você usa essas palavras nos mesmos contextos?
91. Quando você está falando sobre você junto a mais alguém você fala mais **nós** ou fala mais **a gente**? Por quê?
92. Você acha que é melhor usar mais **nós** ou **a gente**?
93. O que você acha de construções como “**nós comemu**”, “**a gente vamu**”, “**nós vai**”?
94. Você acha que essas formas determinam de onde o falante é? (região do país, Estado)
95. Você acha que a escolarização tem relação com a escolha de uma dessas formas?

6.2 ANEXO 2

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Caro participante,

Estamos convidando-o a participar como voluntário(a) de uma pesquisa de campo a ser realizada por meio da gravação de entrevistas sobre temas relacionados às experiências de vida e práticas acadêmicas universitárias.

A coleta será realizada com o objetivo de desenvolvermos o Projeto Falares Sergipanos virtual: variedade, diversidade, contato e os direitos linguísticos (Edital 02/2015 SENACON/MJ; Edital CAPES/FAPITEC/SE 10/2016 PROMOB). A partir da identificação de valores e atitudes da comunidade universitária, o projeto tem como finalidade coletar práticas das vivências dos estudantes universitários.

A entrevista coletada ficará disponível no banco de dados *Falares Sergipanos*, do Grupo de Estudos em Linguagem, Interação e Sociedade – GELINS, aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Sergipe (CAAE: 0386.0.107.000-11), para ser utilizada em pesquisas futuras. Serão resguardadas todas as informações de identificação de forma que se mantenha o anonimato.

Para participar desta pesquisa, o responsável por você deverá autorizar e assinar um Termo de Consentimento. Esta pesquisa apresenta risco mínimo, isto é, o mesmo risco existente em atividades rotineiras como conversar, tomar banho, ler e etc. Não será cobrado nada, não haverá gastos e não estão previstos ressarcimentos ou indenizações.

Você poderá solicitar esclarecimento sobre a pesquisa em qualquer etapa do estudo.

PESQUISADORA RESPONSÁVEL: Raquel Meister Ko. Freitag (Universidade Federal de Sergipe, Departamento de Letras Vernáculas), telefone (79) 2105-6602.

Você será esclarecido (a) em qualquer aspecto que desejar e estará livre para participar ou recusar-se. O responsável por você poderá retirar o consentimento ou interromper a sua participação a qualquer momento. A sua participação é voluntária e a recusa em participar não irá acarretar qualquer penalidade ou perda de benefícios.

Desde já, agradecemos sua atenção e participação e colocamo-nos à disposição para maiores informações.

Assentimento para participação

Eu, _____, portador (a) do documento de Identidade _____, fui informado (a) dos objetivos da presente pesquisa, de maneira clara e detalhada e esclareci minhas dúvidas. Sei que a qualquer momento poderei solicitar novas informações, e me retirar do estudo a qualquer momento sem qualquer prejuízo, e o meu responsável poderá modificar a decisão de participar se assim o desejar. Tendo o consentimento do meu responsável já assinado, declaro que concordo em participar dessa

pesquisa. Recebi uma cópia deste termo de assentimento e me foi dada a oportunidade de ler e esclarecer as minhas.

_____, ____ de _____ de _____

Assinatura do (a) participante:

Assinatura do (a) pesquisador (a):

Assinatura do (a) documentador (a)

6.3 ANEXO 3

FICHA SOCIAL (PERGUNTAS DE CHECAGEM)

Nome: _____

Qual o curso? _____

Campus/UFS: _____ Período: _____ Idade: _____

Qual a sua ocupação? _____ Gênero: _____

Profissão dos seus pais: _____ / _____

Cidade/UF onde nasceu: _____

Cidade/UF onde mora atualmente: _____

Bairro _____

Mora

Na casa dos pais República Residência Estudantil Casa própria

Outro, onde? _____

Onde almoça quando está aqui na UFS?

No Resun Nos restaurantes Vai almoçar em casa Traz de casa

Outro, onde? _____

Como você vem para UFS?

A pé Transpo coletivo Setrans Transporte escolar Carro Próprio/familiar Táxi/lotação/Uber/moto taxi

Moto Bicicleta

É bolsista?

Sim Não

Recebe algum auxílio estudantil?

Sim Não

Número do Celular: _____