

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE CEILÂNDIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA REABILITAÇÃO

LÍVIA MARIA SANTOS DE SOUZA

**REORGANIZAÇÃO DOS PARÂMETROS TEMPORAIS DA FALA APÓS
TREINAMENTO DE DESEMPENHO COMUNICATIVO ORAL**

Brasília

2021

LÍVIA MARIA SANTOS DE SOUZA

**REORGANIZAÇÃO DOS PARÂMETROS TEMPORAIS DA FALA APÓS
TREINAMENTO DE DESEMPENHO COMUNICATIVO ORAL**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Reabilitação da Universidade de Brasília como requisito necessário para a obtenção do título de mestre.

Linha de pesquisa: Aspectos Biomecânicos e Funcionais Associados à Prevenção, Desempenho e Reabilitação.

Tema de pesquisa: Comunicação humana e seus distúrbios.

Orientadora: Letícia Corrêa Celeste

Coorientador: Philippe Martin

Brasília

2021

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

SS729r Souza, Livia Maria Santos de
Reorganização dos parâmetros temporais da fala após
treinamento de desempenho comunicativo oral / Livia Maria
Santos de Souza; orientador Letícia Corrêa Celeste; co
orientador Philippe Martin. -- Brasília, 2021.
102 p.

Dissertação (Mestrado - Mestrado em Ciências da
Reabilitação) -- Universidade de Brasília, 2021.

1. Comunicação. 2. Fonoaudiologia. 3. Fala. 4. Acústica
da Fala. 5. Fonética. I. Celeste, Letícia Corrêa, orient.
II. Martin, Philippe, co-orient. III. Título.

LÍVIA MARIA SANTOS DE SOUZA

**REORGANIZAÇÃO DOS PARÂMETROS TEMPORAIS DA FALA APÓS
TREINAMENTO DE DESEMPENHO COMUNICATIVO ORAL**

Dissertação de mestrado apresentada ao programa de Pós-Graduação em Ciências da Reabilitação (PPGCR) da Universidade de Brasília (UnB).

BANCA EXAMINADORA:

Profa. Dra. Letícia Corrêa Celeste
Presidente da banca e membro do PPGCR

Profa. Dra. Luciana Mendonça Alves
Membro externo credenciado no PPGCR

Profa. Dra. Aveliny Mantovan Lima
Membro interno credenciado no PPGCR

Prof. Dr. Eduardo Magalhães da Silva
Membro suplente interno credenciado no PPGCR

Brasília, 26 de março de 2021.

*Dedico este trabalho
À minha família,
sem a qual eu não seria.
A Deus,
por tudo.*

Agradecimentos

Agradeço a Deus, por me dar a vida, me sustentar ao longo dela e montar o quebra-cabeça complexo e impressionante que me trouxe até aqui.

Agradeço à minha família, por prover e me cercar dos elementos necessários ao meu desenvolvimento físico, psicológico e emocional.

Agradeço à minha mãe, por tanto lutar e fazer sacrifícios em prol de minha educação, quando os tempos eram mais difíceis.

Agradeço à minha orientadora, Letícia Corrêa Celeste, por me orientar pelos caminhos da academia desde a graduação. Sou imensamente grata por todas as oportunidades oferecidas e pela confiança depositada em minhas habilidades e potencial.

Agradeço ao meu coorientador, Philippe Martin, pela surpreendente receptividade e alegria, por todas as informações, paciência e tempo.

Agradeço à colega de profissão, e amiga, Dianete do Valle, por me receber como parceira de pesquisa e por dedicar seu tempo e paciência de forma a me auxiliar no aprimoramento deste trabalho.

Agradeço aos meus amigos, por tornarem minha vida mais alegre, diversa e jovial.

Agradeço aos membros da banca por suas sugestões, as quais tanto contribuíram para o enriquecimento desta dissertação.

Agradeço, ainda, a todos os meus pacientes, por me ensinarem de tudo um pouco e me inspirarem a ser uma profissional cada vez melhor.

Agradeço a todos que fizeram parte da minha caminhada até aqui, aos que ainda seguem ao meu lado e aos que não mais.

“...mas quem tem vontade de aprender e quer fazer alguma coisa, prefere a lição que melhora ao ruído que lisonjeia.”

Machado de Assis (1872)

SUMÁRIO

Apresentação	14
Resumo.....	15
Abstract.....	16
1. Introdução.....	17
2. Revisão da Literatura.....	20
2.1 <i>Produção da fala</i>	20
2.2 <i>Organização temporal da fala</i>	22
2.2.1 <i>Fluência e disfluência de fala</i>	23
2.2.2 <i>Pausa</i>	25
2.2.3 <i>Velocidade de fala</i>	27
2.2.4 <i>Grupo acentual e contornos melódicos</i>	29
2.3 <i>Percepção da fala</i>	31
2.3.1 <i>Influência dos eventos temporais na percepção de fala</i>	33
2.3.2 <i>Sensação de certeza e incerteza</i>	35
3. Justificativa.....	37
4. Objetivos	38
5. Métodos	39
5.1 <i>Tipo de estudo</i>	39
5.2 <i>Aspectos éticos</i>	39
5.3 <i>Descrição da amostra</i>	39
5.3.1 <i>Critérios de inclusão</i>	39
5.3.2 <i>Critérios de exclusão</i>	40
5.4 <i>Procedimentos</i>	40
5.4.1 <i>Coleta das amostras de fala</i>	40
5.4.2 <i>Treinamento de desempenho comunicativo oral</i>	41

5.4.3 <i>Análise perceptivo-auditiva</i>	42
5.4.4 <i>Análise das amostras de fala</i>	43
5.4.5 <i>Análise da fluência de fala</i>	45
5.4.6 <i>Grupo acentual e contornos melódicos</i>	46
5.4.7 <i>Pausas</i>	50
5.5 <i>Análise estatística</i>	51
6. Resultados	52
6.1 <i>Análise perceptivo-auditiva</i>	52
6.2 <i>Análise da fluência de fala</i>	55
6.3 <i>Grupo acentual e contornos melódicos</i>	57
6.4 <i>Pausas</i>	62
6.5 <i>Correlações</i>	64
6.6 <i>Maiores notas vs. menores notas</i>	65
6.7 <i>Outros achados</i>	70
7. Discussão.....	72
8. Conclusão	80
Referências	81
Anexos	99
Anexo 1 – Parecer Consubstanciado do CEP	100
Anexo 2 – Textos argumentativos apresentados aos participantes	103

Relação de tabelas, gráficos e figuras

Tabela 1. Vogais salientes do enunciado e sua ordem de anotação.....	48
Tabela 2. Tipos de contorno entoacional e suas respectivas formas de notação.....	49
Tabela 3. Teste t de amostras independentes entre os grupos GEf e GCf em relação às notas na APA.....	54
Tabela 4. Teste t de amostras independentes entre os grupos GEi e GCi em relação às notas na APA.....	54
Tabela 5. Resultado da análise de variância unidirecional para as notas na APA.....	55
Tabela 6. Resultados da análise de comparações múltiplas.....	57
Tabela 7. Resultados das correlações de Pearson.....	64
Tabela 8. Resultados das correlações de Spearman.....	64
Tabela 9. Média dos valores das variáveis dos indivíduos que obtiveram as maiores notas e menores notas na APA.....	68
Tabela 10. Teste t de amostras independentes entre os grupos com as maiores e menores notas na análise perceptivo-auditiva.....	69
Tabela 11. Teste Mann-Whitney entre os grupos com as maiores e menores notas na análise perceptivo-auditiva.....	70
Gráfico 1. Boxplot simples apresentando os dados descritivos das notas obtidas na análise perceptivo-auditiva pelos grupos experimental e controle nas condições inicial e final (pré e pós treinamento).....	52
Gráfico 2. Média das notas individuais recebidas na análise perceptivo-auditiva pelos sujeitos do grupo experimental nas condições inicial e final.....	53
Gráfico 3. Média das notas individuais recebidas na análise perceptivo-auditiva pelos sujeitos do grupo controle nas condições inicial e final.....	53

Gráfico 4. Boxplot simples apresentando os dados descritivos em relação aos valores da diferença em Hz apresentada pelos sujeitos na palavra 'vou', por grupo.....	58
Gráfico 5. Tipos de contornos melódicos apresentados pelo grupo experimental nas condições inicial (a) e final (b).....	58 e 59
Gráfico 6. Tipos de contornos melódicos apresentados pelo grupo controle nas condições inicial (a) e final (b).....	59
Gráfico 7. Número total de pausas produzidas por grupo.....	63
Gráfico 8. Quantitativo de pausas utilizadas e sua localização em relação às palavras do enunciado, independente de grupo.....	63
Gráfico 9. Quantitativo de pausas utilizadas e sua localização em relação às palavras do enunciado, considerando os grupos experimental (GE) e controle (GC) nas condições inicial (i) e final (f).....	64
Gráfico 10. Quantitativo de pausas utilizadas nos enunciados dos quinze indivíduos que obtiveram as maiores e menores notas na APA.....	65
Gráfico 11. Duração, sem segundos, da primeira pausa dos enunciados dos quinze indivíduos que obtiveram as maiores e menores notas na APA.....	66
Gráfico 12. Duração, sem segundos, da segunda pausa dos enunciados dos quinze indivíduos que obtiveram as maiores e menores notas na APA.....	66
Gráfico 13. Tipos de contornos melódicos apresentados pelos grupos que obtiveram as maiores (a) e menores (b) notas na análise perceptivo-auditiva.....	67
Figura 1. Fluxogramas de decisão utilizados no treinamento de aprimoramento do desempenho comunicativo oral.....	42
Figura 2. Exemplo do padrão de anotação adotado.....	44
Figura 3. Grupos acentuais presentes na emissão esperada da frase padronizada para a amostra de fala.....	47
Figura 4. Exemplo de marcação das sílabas salientes do enunciado.....	48
Figura 5. Exemplo de demarcação das pausas nos enunciados.....	50

Figura 6. Tipos de organização prosódica apresentada pelos sujeitos com melhora significativa no GE nas condições pré e pós treinamento.....60 e 61

Figura 7. Comparação do tipo de estrutura prosódica dos enunciados que obtiveram a maior (esquerda) e menor (direita) média de notas da APA.....61

Relação de siglas e abreviaturas

OTF	Organização temporal da fala
DTG	Disfluência típica da gagueira
DC	Disfluência comum
CEP	Comitê de Ética em Pesquisa
TCLE	Termo de Consentimento Livre e Esclarecido
APA	Análise perceptivo-auditiva
SPM	Sílabas por minuto
SPS	Sílabas por segundo
PPM	Palavras por minuto
GA	Grupo acentual
GE	Grupo experimental
GC	Grupo controle
TE	Taxa de elocução
TA	Taxa de articulação
F0	Frequência fundamental
ms	Milissegundos
Hz	Hertz
s	Segundo(s)

Apresentação

Me chamo Livia Maria Santos de Souza, sou fonoaudióloga graduada pela Universidade de Brasília – UnB. Tenho o privilégio de dizer que também possuo outra graduação nesta universidade, a qual me proporcionou inestimável crescimento pessoal e profissional.

Durante a graduação em fonoaudiologia, participei de projetos acadêmicos voltados para as áreas de extensão e de pesquisa. Participei como bolsista e extensionista do Programa Institucional de Bolsas de Extensão (PIBEX) e do Programa de Iniciação Científica (ProlC). O projeto de iniciação científica do qual participei na graduação em fonoaudiologia abordava o tema da fluência de fala, mais especificamente, tratava de um método de intervenção na gagueira em adultos. Elaborei esse projeto sob a orientação da professora doutora Letícia Corrêa Celeste, pela qual também tive o privilégio de ser orientada, juntamente com a professora doutora Vanessa de Oliveira Martins-Reis, durante a elaboração do meu trabalho de conclusão de curso. Assim como na iniciação científica, meu trabalho de conclusão de curso foi elaborado na área de fluência de fala, campo de conhecimento que é foco de meu interesse desde o início da graduação em fonoaudiologia e, não por acaso, área de estudo na qual desenvolvo, agora, minha dissertação de mestrado.

Após a conclusão da graduação, tive a oportunidade de participar de forma ativa da coleta de dados, da revisão da literatura e da escrita da dissertação de mestrado da colega de profissão, e amiga, Dianete do Valle. Os dados coletados nesse período foram utilizados para o desenvolvimento da presente pesquisa.

Durante o período do mestrado, tive o prazer de elaborar o Manual básico do usuário em português brasileiro do programa WinPitch. O manual teve como coautor o criador do programa WinPitch, e coorientador da presente dissertação, Philippe Martin. A orientadora deste trabalho, Letícia Corrêa Celeste, também foi coautora do manual. A revisão por pares do manual foi realizada por meio da banca examinadora desta dissertação.

Ressalto, por fim, que o objetivo principal desta dissertação é contribuir para o corpo de conhecimento por ela abordado, devolvendo para a sociedade todo o investimento realizado.

Resumo

Introdução: Componentes da organização temporal da fala (OTF) podem agir como pistas acústicas que informam e guiam a compreensão. Dentre estes efeitos, pode-se citar a sinalização de problemas de produção da fala, a informação a respeito do nível de conhecimento do falante sobre um tema e o fornecimento de pistas facilitadoras da compreensão da mensagem transmitida. Os componentes da OTF podem ser modificados pelo treinamento fonoaudiológico, trazendo benefícios para o desempenho comunicativo oral. **Objetivo:** O objetivo geral do estudo é verificar a reorganização dos parâmetros temporais da fala de estudantes de nível superior após um treinamento de aprimoramento do desempenho comunicativo oral. **Método:** O grupo experimental (GE) foi composto por 20 adultos do sexo masculino, entre 18 e 35 anos de idade, que passaram por treinamento de desempenho comunicativo oral. O grupo controle (GC) foi composto por 21 adultos do sexo masculino entre 18 e 40 anos de idade. Foram realizadas gravações em áudio dos participantes nas condições inicial e final. Os enunciados foram submetidos à análise acústica e perceptivo-auditiva. A análise acústica consistiu na avaliação das pausas, velocidade de fala, fluência de fala, grupo acentual e de contornos melódicos dos enunciados emitidos nas condições inicial e final pelos grupos GE e GC. 23 juízes leigos realizaram a análise perceptivo-auditiva, na qual deveriam julgar o grau de certeza de 82 enunciados (41 na condição inicial e 41 na condição final). Os juízes desconheciam as condições dos enunciados e realizaram o julgamento de forma cega. 1.886 respostas foram computadas. **Resultados:** O GE na condição final se diferenciou dos demais grupos (GCi, GCf e GEi) e obteve as maiores notas na análise perceptivo-auditiva. Variáveis como a taxa de articulação de sílabas, o número de palavras emitidas por minuto e a variação de F0 foram distintas no GE na condição final. A análise dos enunciados que obtiveram as maiores e as menores notas nas avaliações perceptivo-auditivas indicou tendências de relações entre as variáveis de organização temporal da fala e o nível de certeza expresso pelos sujeitos. **Conclusão:** O treinamento de desempenho comunicativo oral aumentou o nível de certeza da fala, conforme resultados da análise perceptivo-auditiva realizada por juízes leigos. Os achados auxiliam o direcionamento de intervenções clínicas que abordam a OTF visando o aumento do nível de certeza da fala.

Palavras-chave: Comunicação, Fonoaudiologia, Fala, Acústica da Fala, Fonética

Abstract

Introduction: Temporal structure of speech features can act as acoustic cues that inform and guide listener's comprehension. Such effects include signaling speech production problems and cues to the listeners such as speakers' level of certainty and message's comprehension enhancement. Temporal structure of speech features can be modified by speech training, benefiting oral communication performance. **Purpose:** The general purpose of this study is to verify the reorganization of temporal structure of speech of college students after a communicative oral skills training. **Methods:** The experimental group (EG) was formed by 20 males aged between 18 and 35 years old who went through a communicative oral skills training. The control group (CG) consisted of 21 males aged between 18 and 40 years old. Audio recordings of the participants were made on the initial and final conditions. The utterances were submitted to acoustic analysis and auditory-perceptual analysis. The acoustic analysis verified the pauses, speech rate, speech fluency, accent phrases and melodic contours in the utterances produced by EG and CG both in the initial and final condition. 23 lay judges judged the 82 utterances, being 41 on the initial condition and 41 on the final condition. The judges were not aware of the utterances' conditions and judged them blindly. 1.886 answers were computed. **Results:** The EG on the final condition set itself apart from the other groups (CGi, CGf and EGf) and had the best ratings on the auditory-perceptual analysis. Variables such as syllable articulation rate, words per minute and F0 variation were distinct on the EG on the final condition. The analysis of the utterances that had the higher and lower ratings on the auditory-perceptual analysis showed tendencies regarding the relation between the temporal structure of speech and the judgment of certainty. **Conclusion:** The communicative oral skills training improved the rating of certainty on the auditory-perceptual analysis carried out by lay judges. Findings help clinicians set goals for the use of temporal structure of speech features on speech interventions aiming at the improvement of certainty.

Keywords: Communication, Speech, Language and Hearing Sciences, Speech, Speech Acoustics, Phonetics

1. Introdução

A comunicação humana dispõe de componentes verbais e de componentes não verbais, os quais transmitem uma mensagem do falante ao interlocutor. Tais componentes podem ser utilizados de forma intencional ou de forma não intencional. A comunicação não verbal abrange múltiplos fatores, dentre os quais podem-se citar os gestos, as expressões faciais, a postura corporal, o vestuário e o contato visual. Atos comunicativos verbais, por sua vez, incluem a linguagem escrita e a linguagem oral.

Por comunicação verbal entende-se a codificação de um sistema simbólico representativo da língua, codificação esta que pode ocorrer de forma escrita (por meio de grafemas) e de forma oral (por meio de fonemas), os quais integram os componentes linguísticos(1) da comunicação. A comunicação não verbal, por sua vez, é composta pela interação entre elementos paralinguísticos e elementos extralinguísticos(2). Por aspectos paralinguísticos entende-se as características fonéticas e prosódicas (como a velocidade de fala, a entonação e as pausas)(3,4), as quais, dentre outras funções, complementam os elementos linguísticos, fornecem pistas temporais (como para a troca de turno entre o falante e o interlocutor) e concedem características de expressividade à fala(5). Os componentes extralinguísticos, por sua vez, podem incluir elementos contextuais e visuais diversos, como o vestuário e o corte de cabelo(2). A comunicação humana, portanto, é composta por aspectos linguísticos, paralinguísticos e extralinguísticos(2), os quais se relacionam para compor a mensagem final a ser expressa.

A fala e a voz são componentes linguísticos da comunicação oral que, associados aos seus aspectos estruturais e constituintes (como a fluência, a velocidade de fala, a pausa, a articulação e a intensidade), codificam e transmitem a atitude do falante que, por sua vez, imprime em suas características expressivas sua individualidade e sua intenção(6). A partir do momento no qual o falante se expressa, ele se torna sujeito a avaliações e julgamentos, sendo o seu desempenho território livre para as interpretações de seus ouvintes.

Por desempenho compreende-se a competência da execução de determinado ato (no presente trabalho, um ato comunicativo oral) e o modo pessoal de

representação de suas habilidades frente a situações comunicativas. O desempenho, por ser reflexo da habilidade, é afetado por diferentes variáveis, dentre elas a familiaridade com a situação em questão e a prática na execução do ato, podendo o desempenho ser influenciado, então, pelo treinamento intencional de tais aspectos.

Considerando o panorama exposto até o presente momento, o treinamento das habilidades comunicativas orais deve considerar, portanto, os aspectos linguísticos, paralinguísticos e extralinguísticos da comunicação(3,6,7), visando o aperfeiçoamento do desempenho do falante na tarefa de fala envolvida. Dentre os aspectos linguísticos e paralinguísticos da comunicação, encontra-se a organização temporal da fala, caracterizada pelos atributos que dizem respeito à duração e à organização do enunciado e de seus componentes.

Os componentes da organização temporal da fala sofrem ajustes e modificações constantes a depender do falante e da tarefa de fala envolvida. Sabe-se, por exemplo, que a velocidade de fala é suscetível a ajustes a depender do interlocutor(8). Por sua vez, as pausas, a fluência e a entonação também compõem a organização temporal da fala e tais parâmetros, conseqüentemente, sofrem alterações em sua organização a depender da situação comunicativa e estão suscetíveis a padrões e características próprias do falante.

Além dos fatores que dizem respeito à produção dos aspectos temporais, há que se considerar o efeito que essa organização provoca nos ouvintes, ou seja, o papel da organização temporal na percepção de fala e na imagem do falante. Considerando tal influência, a busca pelo aprimoramento pessoal em relação à competência comunicativa é recorrente em indivíduos que reconhecem a sua fala enquanto forma de ação em seu meio e objeto de avaliação externa, principalmente quando ingressam em ambientes que demandam e evidenciam tais papéis exercidos pela fala, como a universidade e o mercado de trabalho.

Os componentes da organização temporal da fala podem ser modificados pelo treinamento fonoaudiológico, o qual pode trazer benefícios para o desempenho comunicativo oral, conforme resultados previamente descritos na literatura(9–12). Em levantamento de estudos publicados a respeito do treinamento comunicativo oral, observa-se que tais estudos avaliaram os componentes da organização temporal da fala ora associados aos demais aspectos da comunicação ora de forma isolada, não

sendo encontrados estudos que tratam apenas da influência da organização temporal e de seus componentes na atitude comunicativa de certeza, verificando sua reorganização após um treinamento de desempenho comunicativo.

Em estudo recente(11) publicado na temática do treinamento comunicativo oral, foram avaliados os valores de frequência em Hertz (Hz) do enunciado como um todo, a velocidade de fala em palavras por minuto, as pausas foram analisadas quanto ao número de pausas e pausas emitidas por minuto e houve ausência de análise perceptivo-auditiva dos dados e de grupo controle. Tais escolhas metodológicas serão discutidas no presente trabalho.

Em busca de novas evidências, o presente estudo buscou analisar os aspectos da organização temporal tanto do ponto de vista acústico quanto do ponto de vista perceptivo, visando identificar o comportamento das variáveis temporais da fala (tais como os aspectos entoacionais e de fluência) e sua possível reorganização após um treinamento de desempenho comunicativo oral em estudantes de nível superior.

2. Revisão da Literatura

Para haver comunicação oral é preciso que haja, a princípio, a intenção de se comunicar. A intenção comunicativa surge quando há alguma razão ou desejo de expressar um conteúdo. Este conteúdo, por sua vez, deve ser expresso por um meio de comunicação. No caso da fala, o meio físico de comunicação é o ar, no qual as ondas sonoras se propagam em direção ao ouvinte(13). O conteúdo emitido pelo falante, portanto, é direcionado ao receptor da mensagem por meio de ondas sonoras, as quais devem ser captadas e interpretadas. Para que ocorra a comunicação, ainda, é necessário que haja uma condição favorável para tal, um contexto.

Para interpretar a mensagem, o receptor faz uso, além das pistas físicas acústicas da fala, de informações perceptivas visuais que complementam a mensagem, como as expressões faciais do falante, os gestos que este pode vir a executar, sua postura corporal e seu vestuário. Estas informações podem, inclusive, exercer tal influência perceptiva que o ouvinte pode vir a considerá-las como mais relevantes para o julgamento do falante do que a mensagem acústica em si.

Para fins teóricos e de análise no presente trabalho, trataremos unicamente das pistas acústicas envolvidas na organização temporal e no processo de percepção de atitudes comunicativas. Como expressão comunicativa oral, ou seja, como emissão, entender-se-á, aqui, a mensagem elaborada pelo falante e transmitida ao seu parceiro comunicativo por meio de uma sequência de sinais acústicos que codificam uma língua. A esta codificação física de um sistema linguístico, por meio de atos motores que alteram as propriedades de ondas sonoras, denomina-se fala.

2.1 *Produção da fala*

A produção do sinal de fala se inicia com mecanismos neurofisiológicos que programam e desencadeiam eventos neuromotores, os quais culminam na produção dos sons da fala(13). A produção da fala pode ser considerada uma atividade rítmica na qual grupos de fonemas são produzidos a uma determinada velocidade(14) como produto de sequências articulatórias que codificam uma língua. O processo

articulatório envolve mecanismos respiratórios prévios e movimentos coordenados ao longo do trato vocal de forma a codificar fisicamente o conteúdo a ser emitido.

Para que haja a produção vocal, é necessário que um fluxo de ar deixe os pulmões em direção à laringe, na qual ocorrerá a sonorização/vozeamento, por meio das pregas vocais(13). As ondas sonoras produzidas pela vibração das pregas vocais seguem, então, em direção às estruturas articulatórias e às caixas de ressonância em seu caminho pelo trato vocal(13). As propriedades das ondas sonoras são, então, modificadas pela constrição e modificação das estruturas do trato vocal e pela ressonância de suas cavidades(13). Assim descreve-se o processo de articulação da fala. A articulação e as suas formas de uso influenciam o fluxo da fala, podendo alterar o seu ritmo e as suas características tonais(15).

Os eventos e os movimentos musculares necessários para a produção da fala se organizam de forma linear do tempo e são percebidos como um fluxo contínuo e integrado de gestos articulatórios. Os gestos articulatórios, porém, são produzidos em combinações entre si para formarem unidades progressivamente maiores e, portanto, os efeitos do contexto fonético não podem ser ignorados na descrição da produção da fala(16). A produção de sons em contexto geralmente requer ajustes temporais, inclusive nos parâmetros de duração(17). Na fala encadeada, os gestos articulatórios que codificam os segmentos não são produzidos de forma individual e isolada e podem afetar os demais segmentos em sua vizinhança; esse fenômeno se denomina coarticulação(16).

A coarticulação (ou coprodução) é um fenômeno constante que pode ocorrer em menor ou maior grau na fala. A coarticulação, ou seja, a sobreposição temporal dos movimentos articulatórios(18) pode ter efeito antecipatório (no qual o traço de uma unidade de fala é antecipado durante a produção de uma unidade anterior) ou preservatório (no qual o traço é retido e persiste durante a produção de uma unidade seguinte)(17,19). Os eventos coarticulatórios podem ser descritos por meio dos conceitos de espraiamento (quando ocorre expansão articulatória e um traço transpõe suas fronteiras temporais), sobreposição (na qual há produção simultânea e invasão de um traço pertencente a outro fonema) e fusão (na qual os movimentos se fundem e o fonema é alterado de forma mais intensa por seu contexto)(17).

Cabe ressaltar, por fim, que a fala não se apoia no movimento de articuladores individuais, mas surge da dinâmica que ocorre entre eles na busca de um objetivo comum(20). A partir do conhecimento a respeito de seus mecanismos de produção e de seu contexto de ocorrência, a produção da fala pode ser entendida enquanto uma série de eventos físicos e acústicos que ocorrem em uma determinada sequência(13) e velocidade no tempo. A característica dos gestos motores acima descritos pressupõe uma ordenação no domínio temporal. A organização temporal, portanto, exerce uma função essencial dentre os aspectos relacionados à produção da fala.

2.2 Organização temporal da fala

Não há língua que organize seus componentes sem regras temporais(21,22). Todo ato de fala ocorre de forma linear no tempo e é manifestado por meio da organização sequenciada de fonemas, os quais compõem unidades progressivamente maiores (fonemas → sílabas → palavras → enunciados → discurso). A organização temporal da fala influencia diretamente na inteligibilidade da fala(23–25) e os aspectos temporais transmitem informações linguísticas(26), como a duração de fonemas, de sílabas e de palavras, unidades que interagem entre si para transmitirem a mensagem final ao ouvinte.

A fala é organizada temporalmente pelos parâmetros de velocidade, entonação e pausa, parâmetros estes que dizem respeito à duração do enunciado e de seus segmentos. A fala, porém, não pode ser produzida totalmente de forma contínua; interrupções são necessárias para as funções fisiológicas (respiratória) e linguísticas, como o planejamento motor da fala, a ênfase ou a reorganização após erros de produção. Podem ser citadas, ainda, as influências extralinguísticas na organização temporal da fala, como o contexto comunicativo e o estado de humor do falante(14). Pode-se citar, ainda, o fato de que as características temporais de segmentos e sílabas são alteradas na fala direcionada às crianças, havendo aumento do VOT (*Voice Onset Time*: tempo entre o relaxamento dos articuladores de fonemas plosivos e o início do vozeamento)(19,27) e maior variabilidade de duração na última unidade linguística do enunciado(28). Ainda, alterações na modulação temporal e espectral da fala podem alterar o seu reconhecimento(29).

Mais recentemente, estudos publicados nas áreas de engenharia e da física acústica vêm abordando a temática da produção e do reconhecimento automatizado da fala, trazendo novos conhecimentos a respeito da fisiologia, da acústica e do processo global envolvido na organização temporal da fala e em sua percepção(30–32).

Certas interrupções na fala, como as disfluências e as pausas, podem ser indicativas de alterações tanto no planejamento quanto na execução motora da fala. A organização temporal pode refletir, portanto, uma organização cognitiva, na qual componentes como a hesitação, por exemplo, podem ser utilizados de forma a exercerem funções relacionadas ao planejamento do discurso(33).

Dentre os componentes da organização temporal, podemos citar a fluência de fala, a pausa, a velocidade de fala e a organização tonal enquanto marcadores importantes do discurso e indicadores de características de produção, além de influenciadores na percepção de fala.

2.2.1 Fluência e disfluência de fala

Por fluência de fala entende-se o padrão no qual a fala é emitida e organizada pelo falante. Esta emissão, quando fluente, deve ser contínua, suave e com velocidade apropriada. A fala fluente não exige esforço motor, emocional, cognitivo ou linguístico por parte do falante, podendo sofrer alterações a depender da situação comunicativa(34).

A disfluência de fala consiste na presença de características que divergem das supracitadas. A disfluência é, então, compreendida enquanto uma interrupção no fluxo da fala, podendo ser acompanhada por tensão e esforço por parte do falante. As disfluências tendem a ocorrer quando o falante se depara com várias possibilidades semânticas e sintáticas(35), assim como em casos nos quais outros tipos de atitudes de incerteza ocorrem(36).

As disfluências podem ser classificadas em disfluências comuns e disfluências típicas da gagueira, sendo as disfluências comuns mais associadas à fala fluente, enquanto as disfluências típicas da gagueira tendem a caracterizar a fala de falantes com gagueira. As disfluências comuns e as disfluências típicas da gagueira, porém,

encontram-se presentes no discurso de todos os falantes(37), sendo o aspecto qualitativo e quantitativo de sua presença o fator diferencial para a determinação de quadros de alterações da fluência, como a gagueira(38) e a taquifemia(39).

As disfluências típicas da gagueira, de forma geral, ocorrem dentro das palavras, alterando o seu aspecto temporal e a sua produção(40–43). As disfluências comuns, por sua vez, ocorrem fora das palavras, não as segmentando(37,44). Dentre as disfluências típicas da gagueira, podemos citar os bloqueios (permanência em uma posição articulatória para além de seu tempo habitual de produção), os prolongamentos dentro de palavra (aumento atípico da duração de um fonema) e as repetições de sons e de sílabas. As disfluências comuns consistem em hesitações (intrusão de um fonema ou de uma palavra, geralmente com emissão prolongada), interjeições (como: aí, daí e né), revisões (correções de um erro prévio de produção), repetições de palavras e de segmentos, pausas e palavras não terminadas.

As disfluências podem ser interpretadas como subprodutos da dificuldade na fala(45) e algumas disfluências aumentam de acordo com o aumento da demanda imposta ao sistema de planejamento de fala(46–48). Disfluências aumentam o tempo do sinal de fala, fato este associado a dificuldades nos aspectos de produção da fala, incluindo o planejamento sintático e o acesso lexical(45,47,49).

Disfluências tendem a preceder elementos menos previsíveis da fala(50–52), podem estar associadas a um vocabulário mais amplo(35) e ocorrem mais frequentemente quando o tópico de fala é desconhecido(46,53). As disfluências são mais prováveis de ocorrerem no início de um enunciado(54,55) e os enunciados maiores são associados a maiores taxas de ocorrência de disfluências quando comparados a enunciados menores(54,55).

Aproximadamente seis em cada cem palavras são afetadas por disfluências de fala, incluindo repetições, correções e hesitações, como os *fillers* “um” e “er”(46,56). Dentre as disfluências, os *fillers* (junto com os prolongamentos) marcam hesitações por parte dos falantes. Em estudo recente sobre a avaliação automática da fluência na tarefa de fala espontânea(57), as informações fornecidas por disfluências com a repetição de parte de palavra foram efetivas para a avaliação da fluência, principalmente quando associadas às pistas prosódicas de velocidade de fala e o tempo de silêncio. Por outro lado, as informações fornecidas por pausas preenchidas

não foram efetivas na avaliação da fluência de fala(57). Ainda, conforme achados de estudo a respeito de pausas e de *fillers*(58), “ums” e pausas silenciosas não são o mesmo fenômeno, como sugere o termo pausa preenchida. Considerando o exposto acima, o presente estudo considerará, portanto, os *fillers*/pausas preenchidas enquanto hesitações(37,40,42–44,46,59).

Resultados de um estudo recente a respeito de hesitações(60) realçam a variação de ajustes na fala no que diz respeito à sua organização temporal e evidenciam o fato de que a produção de disfluências é variável de acordo com a tarefa de fala, sua complexidade e os mecanismos de adaptação individuais, assim como a variação interpessoal.

Em estudo recente sobre a temática, foi encontrada correlação entre alterações cognitivas leves e a fluência de fala(61), com destaque para os parâmetros de taxa de articulação e a média do tempo de emissão. Já em estudo recente abordando a fluência de fala em uma segunda língua(62), os autores sugerem a divisão entre uma visão mais holística da fluência e uma visão direcionada ao processamento léxico-gramatical automático, sendo esta última visão uma reflexão mais direta de processos cognitivos envolvidos na produção da fala.

Conforme exposto acima, as pausas podem ser classificadas enquanto disfluências de fala, a depender de características como a presença de tensão e sua localização no enunciado. As pausas, porém, podem carregar mais significado do que a mera interrupção involuntária da fala, podendo ser utilizadas de maneira intencional para atribuir características de expressividade à fala.

2.2.2 Pausa

A pausa pode ser definida enquanto a interrupção silenciosa de uma emissão. Sob a perspectiva fonética, as pausas são caracterizadas pela ausência do sinal acústico. Em estudo recente, porém, foi encontrado que a maioria das pausas consideradas silenciosas são marcadas pela presença de ruídos respiratórios e, portanto, há uma minoria de pausas que podem ser consideradas foneticamente silenciosas(63).

A produção de pausas pode ser associada ao processo de seleção lexical(64), ao tempo de oclusão do trato vocal na produção fisiológica da fala(14), ao auxílio na interpretação perceptiva do enunciado(14), aos processos de coordenação interpessoal, como a solicitação de respostas(46), às funções relacionadas ao *feedback* e ao gerenciamento de turno(14), à sinalização de processos de planejamento do discurso(65) e à demanda cognitiva na definição de objetos abstratos e concretos(66).

Do ponto de vista da análise, as pausas podem ser estudadas em relação à sua função, à sua duração e à sua localização no enunciado. No que diz respeito à função, as pausas podem ser utilizadas para exercerem a função fisiológica da respiração (ou seja, para a inspiração do ar durante um conteúdo que está sendo emitido pelo falante, interrompendo-o) e para exercerem funções prosódicas (como o uso de pausa após uma emissão, constituindo a ênfase naquele conteúdo), além de serem marcadores dos limites entoacionais do discurso(67). A ocorrência de pausas se relaciona, ainda, à naturalidade da fala(14) e à fluência em uma língua(68).

O parâmetro de duração das pausas pode depender de múltiplos fatores (como, por exemplo, o estilo de fala)(69). A duração das pausas interage com os parâmetros da vogal precedente, e a influência desses valores parecem aumentar conforme a duração das pausa aumenta(70). A duração das pausas também sofre influência da tarefa de fala envolvida, sendo as pausas mais longas na fala espontânea(71). As pausas silenciosas tendem a ter duração maior em narrativas do que em conversas(69) e a duração das pausas pode variar, ainda, de acordo com o tipo de informação contida no enunciado e o conhecimento do falante a respeito dos fatos(72).

Sobre o uso das pausas na fala, há grande variação no que diz respeito à localização da pausa em relação ao enunciado. Em estudos realizados sobre a duração, as pausas em fronteiras foram mais longas do que pausas dentro das frases(69) e as pausas longas tenderam a ocorrer em enunciados curtos(33). Pode-se citar, ainda, a relação entre a localização da pausa e a percepção de fluência(73).

A duração parece ser o parâmetro essencial para a percepção das pausas na fala, influenciando em sua taxa de identificação(70,74,75). Em estudo de fluência em segunda língua(76), os autores concluem que o valor de corte tradicional de 250 ms

é uma boa escolha e que utilizar um limite maior do que 300 ms possui a desvantagem de, dentre outras, confundir o número de pausas com a duração de pausas. Pausas curtas (0.13 – 0.25 seg), porém, são psicologicamente funcionais e sua exclusão da análise articulatória é injustificada(77). A categorização em pausa breve (< 200 ms), média (200 – 1000 ms) e longa (> 1000 ms), sendo as pausas longas presentes apenas na fala espontânea, foi sugerida em estudo prévio(78). Neste mesmo estudo, os autores defendem, baseados em seus achados, a importância da consideração de pausas com duração 'extrema', ou seja, acima ou abaixo dos limites comumente sugeridos pela literatura(78).

Considerando o exposto acima, o presente trabalho considerou como pausa qualquer tempo de silêncio durante a emissão, sem o estabelecimento de valores de corte para limites mínimos ou máximos de duração das pausas. Tal decisão foi tomada considerando que uma definição de pausa baseada apenas em sua duração pode não abranger os aspectos linguísticos e os aspectos psicológicos pertinentes à comunicação.

Assim como as pausas, a velocidade é um componente temporal importante da fala que pode indicar características específicas de sua produção e de sua percepção.

2.2.3 Velocidade de fala

A velocidade de fala pode ser definida como o fluxo de sílabas ou de palavras emitidas em um determinado espaço de tempo, podendo apresentar variações de acordo com fatores como a faixa etária, a situação comunicativa, o estado do falante e as variações interpessoais.

Valores objetivos da velocidade de fala podem ser obtidos a partir do cálculo do número de sílabas e de palavras emitidas por minuto ou por segundo em uma amostra de fala, considerando ou não o tempo de não emissão do enunciado expresso. A velocidade de fala, portanto, é um termo geral que pode abranger tanto a percepção subjetiva da velocidade da fala pelo ouvinte quanto às diversas aferições objetivas que podem ser realizadas para determinar tal velocidade.

Estudos brasileiros que realizam a análise da velocidade de fala utilizam mais frequentemente as medidas de sílabas por minuto e palavras por minuto(39,79–82).

Em contrapartida, estudos internacionais utilizam mais frequentemente a unidade silábica como forma de medida de velocidade de fala(83–85), inclusive contabilizando o número de sílabas emitidas por segundo(86–89). De forma geral, o uso da unidade silábica como medida de velocidade de fala fornece um cenário mais preciso da emissão(90,91), pois as unidades silábicas apresentam menores variações em sua duração quando comparadas à alta variabilidade de duração apresentada pelas palavras(90). As sílabas indicam, então, a regularidade rítmica da fala de forma mais fidedigna, considerando que são produzidas de maneira relativamente consistente dentre aos falantes, tarefas de fala e línguas(20). Ainda, a velocidade de fala expressa em sílabas por segundo possui mais relevância perceptiva, pois iguala a função psicométrica entre ouvintes do sexo masculino e do sexo feminino(89).

Considerando os achados supracitados, em relação às medidas objetivas, o presente trabalho adotará como medidas de velocidade de fala as palavras emitidas por minuto, as sílabas emitidas por minuto e as sílabas emitidas por segundo.

Todos os valores anteriormente citados dizem respeito à taxa de elocução. A taxa de elocução abrange o tempo total do enunciado emitido, levando em consideração os tempos de pausa/não emissão e fornecendo o valor de duração médio do enunciado como um todo.

Serão adotadas no presente trabalho, ainda, as taxas de articulação, que, por sua vez, compreendem o tempo no qual o falante emite somente os componentes articulados da fala, fornecendo, então, o valor médio de duração dos movimentos articulatórios emitidos pelo falante em um determinado espaço de tempo. Para encontrar seu valor, portanto, descarta-se o tempo de não emissão da amostra de fala.

Os cálculos das taxas de elocução e taxas de articulação fornecem, conforme supracitado, valores médios de duração do enunciado. Valores elevados de taxa de elocução e articulação reduzem a precisão articulatória, podendo provocar a distorção ou o apagamento de fonemas, de sílabas e de segmentos. As medidas de velocidade de fala podem demonstrar, portanto, a interação entre os elementos prosódicos e o segmento.

A velocidade de fala afeta a inteligibilidade da fala e influencia o esforço do ouvinte(92–95), podendo provocar, inclusive, respostas fisiológicas, como a dilatação

das pupilas(93). Tais achados evidenciam a relevância desse parâmetro temporal para a comunicação humana. A regularidade das características prosódicas da fala, como a velocidade de fala, pode determinar a espontaneidade da fala, sendo a prosódia mais variável na fala espontânea e em diálogo do que na tarefa de leitura(96). Em estudo realizado com adolescentes(97), a tarefa de fala narrativa teve velocidade de fala muito mais lenta quando comparada à tarefa de fala espontânea. Na tarefa de narrativa, a redução das taxas de articulação e de elocução foi caracterizada por um aumento significativo na duração das hesitações e das pausas, quando comparada à tarefa de fala espontânea(97).

Conforme exposto previamente, a unidade silábica e sua duração influenciam nas medidas de velocidade de fala. Há uma tendência para a duração igual das sílabas do enunciado, exceto quando as sílabas imediatamente anteriores a uma pausa são significativamente maiores(98). Os núcleos da sílaba (as vogais) possuem variação em sua duração, a depender de sua nasalidade e das características da consoante seguinte(98). Uma vogal típica varia sua duração entre 100 e 150 ms. As consoantes, por sua vez, quando em posição inicial na palavra, podem ter duração média de 101-176 ms, com as consoantes mais longas consistindo nas fricativas, as de ponto posterior e as desvozeadas(99).

A duração da sílaba pode variar para indicar acento (sílabas salientes), ênfase e prolongamento. A duração de uma sílaba típica varia entre 200 e 300 ms, enquanto a duração mínima da sílaba é de 100 ms e a duração máxima é de 250 ms, sendo sílabas com duração maior que 250 ms acentuadas ou prolongadas(99). As sílabas, portanto, além de influenciarem na velocidade de fala, são unidades complexas que carregam características e informações que não se limitam a este aspecto, sendo as sílabas elementos que interagem entre si e o enunciado fornecendo características de tonicidade e de expressividade à fala.

2.2.4 Grupo acentual e contornos melódicos

Os padrões entoacionais da fala são compostos por unidades denominadas grupos acentuais ou grupos tonais(100). Entende-se por grupo acentual a unidade prosódica delimitada por sílabas salientes no enunciado(19), por sílaba saliente

entende-se a sílaba na qual ocorre a maior variação melódica, seja esta variação ascendente ou descendente(100).

Os grupos acentuais demonstram a variação da duração ao longo do enunciado e imprimem características rítmicas à produção da fala. Normalmente, no português brasileiro, o número de grupos acentuais/sílabas salientes esperado corresponde ao número de palavras de conteúdo (ou seja, substantivos, verbos, adjetivos e advérbios) presentes no enunciado. As sílabas salientes apontam para a presença de eventos fonológicos, como a organização hierárquica do enunciado e a ênfase(101), e movimentos articulatórios mais amplos tendem a ocorrer em segmentos de sílabas acentuadas(17).

No presente trabalho, será seguida a proposta de Martin(99,101–103) para a definição dos contornos melódicos que constituem as características de entonação do enunciado. Os contornos melódicos podem ser classificados de acordo com o seu efeito perceptivo em decorrência de sua relação com o limiar do glissando (transição entre frequências). Quando abaixo deste limiar, um glissando de F0 é percebido enquanto uma frequência plana/estática e, quando acima do limiar, ocorre uma percepção distinta pelo ouvido humano entre os padrões ascendentes e descendentes de F0(104). Os contornos melódicos C0, Cc, C1, C2 e Cn codificam classes de eventos prosódicos localizados nas sílabas salientes (mais especificamente, em seu núcleo vocálico) dos grupos acentuais(103).

Os contornos C1 e C2 ocorrem acima do limiar do glissando e são contornos que marcam a continuidade do enunciado, enquanto o contorno Cn ocorre abaixo do limiar do glissando e é um contorno que marca a neutralidade(99). O contorno C0 marca conclusão, um contorno terminal descendente conclusivo declarativo(99). O contorno Cc, por sua vez, possui características complexas, podendo ser plano, ascendente ou descendente, a depender da sílaba na qual ocorre. No português brasileiro, assim como nas demais línguas românicas (exceto o francês), a classificação hierárquica dos contornos prosódicos ocorre como se segue: $Cn < C1 < C2 < Cc < C0$ (103).

O sistema de hierarquia dos contornos indica que a existência de um contorno C1, por exemplo, pressupõe a ocorrência de um contorno C0 vindouro. Palavras que contém contornos melódicos em determinada posição na hierarquia ($Cn < C1 < C2 <$

Cc < C0), serão subordinadas, na organização melódica do enunciado, a contornos de posição mais alta na hierarquia. Os contornos, então, formam grupos prosódicos com contornos de hierarquia mais alta que a sua(102). Os contornos melódicos acima elencados são, portanto, tipos de eventos prosódicos que se diferenciam por meio de uma série de características acústicas e perceptivas que indicam uma relação de dependência em relação a um evento futuro(102).

No que diz respeito à influência dos grupos acentuais na percepção da fala, a localização das pausas em relação aos grupos acentuais é um indicador de fluência. Pausas, de forma típica, não costumam ocorrer dentro de grupos acentuais, mas sim em seus limites. A presença de pausas dentro de grupos acentuais, portanto, indica alterações na fluência de fala ou demais alterações, como, por exemplo, a presença de fatores externos (extralinguísticos) que podem vir a interromper a fala.

Ainda em relação à percepção, o contorno melódico da expressão de dúvida, em oposição à expressão de certeza, se inicia de forma interrogativa, mas termina de forma declarativa moderadamente descendente, combinando, portanto, indicações contraditórias: a solicitação de uma informação e a recusa da resposta que o falante pode vir a obter(102).

Considerando os efeitos que os componentes da organização temporal podem exercer na percepção da fala, tanto isoladamente quanto na interação entre si, é pertinente abordar a temática da percepção da fala, seus mecanismos e suas formas de avaliação.

2.3 Percepção da fala

A percepção da fala consiste na detecção, na discriminação e na interpretação de pistas acústicas em representações linguísticas(105). Dentre as pistas acústicas do sinal de fala, podem ser citadas as suas características temporais. Por parte do ouvinte, habilidades como a memória e a capacidade de reconhecimento de sinais de fala que compõem o sistema linguístico são necessárias para a percepção da fala(106,107).

Além das habilidades do ouvinte, a percepção da fala é influenciada, também, pelas habilidades do falante, sejam elas intencionais ou não(108). A correspondência

entre as pausas e as estruturas prosódicas e fronteiras sintáticas do enunciado, por exemplo, são ferramentas utilizadas pelo falante para facilitar a compreensão do ouvinte(109). Sabe-se, ainda, que as próprias disfluências do falante podem contribuir para a percepção do ouvinte e trazer mais efetividade para a interação entre os falantes(110–112), não estando as disfluências de fala relacionadas apenas à falta de competência comunicativa.

Constata-se, a partir da exposição anterior, que os fatores que influenciam a percepção da fala são diversos e estão interligados. Os estudos que abordam a percepção da fala, portanto, muito dificilmente serão realizados de forma a isolar todas as variáveis que interferem neste processo de percepção. É necessário considerar, ainda, que os mecanismos exatos de processamento neural do sinal acústico da fala, e como é feita a interface com as representações linguísticas por ele evocadas, ainda não foram completamente descritos(113).

No que diz respeito às medidas de avaliação da percepção da fala, as análises perceptivo-auditivas vêm sendo utilizadas amplamente como instrumentos confiáveis para avaliar a percepção dos ouvintes em relação a um estímulo de fala, sendo um método de avaliação consolidado(114,115), considerado padrão ouro(114–117) e utilizado em estudos mais recentes(93,118–121).

Resultados baseados na avaliação do ouvinte podem ser obtidos por meio de análises nas quais os juízes, especialistas ou leigos, a depender do objetivo do estudo, analisam e julgam enunciados apresentados de forma auditiva. Este julgamento pode ser expresso por meio de escalas numéricas ou escalas qualitativas de avaliação ou, ainda, por meio de aferições neurofisiológicas do ouvinte em resposta a um estímulo auditivo(118,119). Quanto à confiabilidade do julgamento dos ouvintes, sabe-se que juízes leigos, por exemplo, são capazes de realizar a distinção adequada da expressão de atitudes(122).

Quanto ao número de juízes utilizados em estudos que conduziram análises perceptivo-auditivas, o número médio encontrado em estudos recentes que utilizaram este método(118,123–125) consistiu em 8 juízes, com o número máximo de 18 juízes(118) e o número mínimo de 3 juízes(125).

Partindo para a neurofisiologia da percepção da fala, em estudo sobre a percepção de aspectos temporais(126), foram identificadas regiões bilaterais do sulco

temporal superior cujas respostas variaram de acordo com a duração do segmento, sendo que as respostas dos ouvintes aumentaram com segmentos de até aproximadamente 500 ms, esse efeito, porém, não ocorreu com outros sons naturais ou com sons sintetizados acusticamente.

O processamento musical e o processamento de fala compartilham substratos neurais, mas parece haver diferença no processamento da estrutura temporal dinâmica(21). O processamento de ruídos com estrutura temporal complexa não é claramente dominado por nenhum dos hemisférios cerebrais(127).

Quanto ao processamento dos parâmetros da organização temporal da fala, o hemisfério direito parece predominar na percepção de transições acústicas lentas, enquanto a percepção do estímulo de fala com transições acústicas rápidas é dominada pelo hemisfério esquerdo(127). Parece haver, ainda, um processamento lateralizado de elementos acústicos de sílabas, sendo o processamento de sílabas vozeadas e desvozeadas distinto no hemisfério esquerdo, mas não no hemisfério direito.

Em estudo recente(113), resultados revelaram que o sulco temporal superior processa propriedades acústicas da organização temporal da fala, enquanto o giro frontal inferior esquerdo mapeia tal informação acústica para transformá-la em representações linguísticas. Tal mapeamento realiza a modulação do processamento de propriedades acústicas da fala em áreas auditivas anteriores no córtex(113).

2.3.1 Influência dos eventos temporais na percepção de fala

Dentre os efeitos que a organização temporal pode provocar na percepção de fala, sabe-se que as disfluências afetam os ouvintes tanto em curto quanto em longo prazo(112,128–130), podendo afetar sua compreensão(56,130). Consequências a longo prazo da disfluência incluem o falante ser classificado como menos provável de saber respostas para perguntas de conhecimento geral quando suas respostas são precedidas por hesitações(36).

As disfluências não são um sinal intencional no diálogo(48). Entretanto, há autores que sugerem que os falantes produzem disfluências específicas de forma

intencional para informar o ouvinte, por exemplo, da duração de uma interrupção antecipada da fala(110,111).

Cabe ressaltar que as disfluências não são, necessariamente, obstáculos para a compreensão(56), podendo inclusive facilitar o entendimento do ouvinte(131). Os ouvintes são sensíveis às disfluências que encontram e, do ponto de vista do ouvinte, nem todas as disfluências são iguais(132). Por exemplo, as pausas possuem efeitos similares no processamento da linguagem e na memória dos ouvintes aos observados nas hesitações, mas efeitos amplamente diferentes dos efeitos observados nas repetições disfluentes(133).

Da perspectiva do ouvinte, uma característica crítica da disfluência é o atraso temporal que ela adiciona ao sinal de fala(133). Pausas disfluentes podem aumentar a expectativa dos ouvintes para a menção de um item lexical que é mais difícil para o falante produzir(133). Hesitações têm um efeito de longo prazo: palavras após o “er” tendem a serem mais reconhecidas no teste de memória(130,134). As hesitações aumentam a atenção imediata do ouvinte para o conteúdo seguinte(112) e os ouvintes são mais rápidos em uma tarefa de monitoramento de palavras quando as palavras são precedidas por uma hesitação(112).

Quanto à duração silábica, o padrão rítmico das sílabas fornece uma característica temporal crítica para a compreensão da fala(24,135). O silêncio, por sua vez, tem consequências para o processamento do ouvinte e para a retenção do material linguístico que ele interrompe(133). A detecção de pausas é uma tarefa psicoacústica e as pausas e as hesitações têm uma influência comparável no julgamento dos ouvintes, porém os ouvintes consideram as pausas como marcadores mais importantes e influentes da fala(136).

Ainda em relação aos efeitos dos aspectos temporais na percepção da fala, estudos demonstraram que adultos com audição saudável compreendem 50% do conteúdo de enunciados emitidos com a velocidade de 12,5 sílabas por segundo(89) e, em estudo de avaliação de emissões aceleradas, sujeitos relataram cansaço após a exposição auditiva a 30 minutos de estímulos com velocidade de 17,5 sílabas por segundo, apesar de não haver prejuízos para a compreensão do conteúdo expresso(137).

Os aspectos perceptivos da organização temporal podem, ainda, afetar a própria produção da fala, podendo ser citado, por exemplo, o efeito que o atraso da retroalimentação auditiva provoca na fluência de fala (afetando tanto seus aspectos quantitativos quanto os aspectos qualitativos)(138–140), na velocidade de fala (reduzindo-a)(138,141,142) e na intensidade vocal (aumentando-a)(143).

2.3.2 Sensação de certeza e incerteza

Por certeza entende-se a qualidade do que é certo, o conhecimento exato, a convicção e a afirmação categórica decorrente de conhecimento(144). A sensação de certeza na fala pode ser transmitida por meio de uma série de características acústicas, dentre as quais os aspectos de organização temporal também exercem seu papel.

A influência da organização temporal na percepção de certeza na fala pode ser atestada por meio de estudos nos quais enunciados com baixa sensação de certeza tenderam a conter hesitações(36,129,145) e, ainda, em estudo de investigação do papel de indicadores prosódicos para a percepção da incerteza, as pistas prosódicas tiveram efeito aditivo (a combinação de hesitação, pausa e entonação resultaram em piores julgamentos do nível de certeza), sendo que a pausa não foi um indicador claro na percepção da incerteza(146).

Pode-se citar, ainda, o fato de que o ouvinte é mais exigido na tarefa de compreensão da fala na ausência de componentes prosódicos, como a entonação e sua correspondência com as fronteiras sintáticas do enunciado. Tal fato é relevante pois os componentes supracitados influenciam no esforço do ouvinte(131,147) e um discurso que requer menos esforço cognitivo por parte do ouvinte é mais agradável e mais bem recebido por ele.

Da mesma forma que elementos paralinguísticos (como as hesitações e a entonação)(148) e neurocognitivos causam impressões no ouvinte, afetando seu julgamento do nível de certeza(108,118), elementos visuais, como as expressões faciais, também podem afetar a sensação de incerteza do ouvinte(129). O ouvinte, ainda, é capaz de detectar características dos falantes, como a personalidade e a confiança, a partir das informações acústicas fornecidas por uma única palavra(149,150).

No que diz respeito à própria percepção do falante em relação ao seu nível de certeza, valores de velocidade de fala muito elevados ou muito reduzidos foram associados à sensação de certeza por parte do falante(151). Em relação ao tempo de silêncio do enunciado, valores baixos de porcentagem de silêncio no enunciado correspondem à sensação de certeza do falante, enquanto altos valores de porcentagem de silêncio estão associados à sensação de incerteza do falante(151). Ainda, em trabalho anterior dos mesmos autores(152) também houve correlação entre a porcentagem de silêncio e a percepção do nível de certeza.

Em estudo recente(118), dezoito participantes julgaram o nível de confiança de quarenta e quatro enunciados emitidos nas condições de certeza e dúvida utilizando uma escala de cinco pontos (na qual 1 representava nenhuma confiança e cinco representava muita confiança). Neste estudo, foi evidenciada a relevância dos aspectos sociais na expressão vocal da confiança e da dúvida no que diz respeito às inferências realizadas quanto ao falante, pois houve distinção entre o julgamento de falantes do mesmo grupo sociolinguístico (obtiveram melhores avaliações) quando comparados a grupos externos (obtiveram julgamentos menores do nível de certeza em relação aos demais grupos)(118).

Em estudo a respeito da expressão de certeza e incerteza(108), os autores discutem que um fator que contribui para a dúvida em relação ao papel da mudança de F0 na percepção de tais atitudes é o fato de que os dados presentes na literatura são originados de medidas realizadas no enunciado como um todo. O presente trabalho se propõe a analisar as alterações melódicas presentes em cada uma das palavras salientes do enunciado (de forma individual, e não apenas como média de valores de F0 para o enunciado inteiro), removendo este viés metodológico por meio da análise local de parâmetros acústicos.

Não se pode perder de vista, e nem aqui se desconsidera, o fato de que o contexto completo de produção do enunciado é relevante para a percepção do nível de certeza(152). De forma a seguir a proposta do presente trabalho, porém, as características visuais/extralinguísticas que influenciam a sensação de certeza não foram analisadas, por tratar-se de *corpus* auditivo.

3. Justificativa

O treinamento fonoaudiológico de desempenho comunicativo oral pode trazer benefícios por meio de sua influência nos parâmetros prosódicos da fala, dentre eles os parâmetros temporais(9–12). A verificação detalhada da reorganização temporal provocada por um treinamento do desempenho comunicativo oral permitirá a avaliação do comportamento dos parâmetros de organização temporal da fala como um todo antes e após a aplicação do treinamento. Considerando o fato de que a organização temporal é relevante para a comunicação, tal investigação faz-se necessária e pertinente.

Desta forma, o presente estudo pretende contribuir para a análise e discussão já existentes na literatura a respeito da efetividade de treinamentos comunicativos orais, norteados as intervenções fonoaudiológicas para parâmetros específicos, indicando quais irão se destacar no julgamento dos juízes como preditores de certeza.

Além dos avanços acadêmicos, os resultados aqui descritos têm implicações clínicas tanto para a avaliação quanto para o treinamento fonoaudiológico voltado para o aprimoramento do desempenho comunicativo oral.

4. Objetivos

O objetivo geral da presente pesquisa consiste em verificar a reorganização dos parâmetros temporais da fala de estudantes de nível superior após um treinamento de aprimoramento do desempenho comunicativo oral.

De forma a responder ao objetivo geral, os seguintes objetivos específicos foram estabelecidos, a saber:

1. Analisar os parâmetros de organização temporal da fala de estudantes de nível superior nas condições inicial e final (no caso do grupo experimental, nas condições pré e pós treinamento e, no caso do grupo controle, condição inicial e final, sem treinamento).
2. Realizar a análise da fluência de fala dos participantes na condição inicial e na condição final.
3. Identificar os tipos de contornos melódicos produzidos pelos sujeitos da pesquisa e sua organização no enunciado proposto para análise.
4. Verificar a possível correlação entre a percepção dos efeitos do treinamento e as medidas de organização temporal da fala e de fluência de fala.

5. Métodos

5.1 Tipo de estudo

O presente trabalho consiste em um estudo longitudinal retrospectivo realizado com o banco de dados de um ensaio clínico controlado não aleatório com juízes e avaliadores cegos(153).

5.2 Aspectos éticos

O Comitê de Ética em Pesquisa da Faculdade de Ceilândia da Universidade de Brasília (UnB), por meio do parecer nº 3.420.515, CAAE: 64822717.6.0000.8093 (**Anexo 1**), aprovou a pesquisa que gerou o banco de dados do qual são oriundas as amostras de fala aqui analisadas. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido (TCLE).

5.3 Descrição da amostra

As amostras de fala desta pesquisa são provenientes de 41 estudantes de nível superior do sexo masculino, com idade entre 18 e 40 anos, sem alterações auditivas e/ou neurológicas conhecidas, alocados por conveniência nos grupos experimental e controle. A alocação em grupos não foi violada por quaisquer fatores conhecidos, independente da etapa da pesquisa.

O grupo experimental foi composto por 20 indivíduos do sexo masculino com faixa etária entre 18 a 35 anos (média = 21,65, mediana = 20,5 e desvio padrão = 3,89) e o grupo controle foi composto por 21 indivíduos do sexo masculino com faixa etária entre 18 e 40 anos (média = 25,85, mediana = 25 e desvio padrão = 6,23).

5.3.1 Critérios de inclusão

- i. Grupo controle (GC): estar no primeiro semestre da graduação e não ter cursado ou não estar cursando a disciplina “Voz e dicção”.
- ii. Grupo experimental (GE): estar no primeiro semestre da graduação e estar matriculado na disciplina “Voz e dicção”.

5.3.2 Critérios de exclusão

- iii. Ter participado de quaisquer treinamentos visando o aprimoramento do desempenho comunicativo oral.
- iv. Apresentar alterações auditivas e/ou neurológicas conhecidas.
- v. Grupo experimental (GE): possuir menos de 75% de presença nas aulas da disciplina “Voz e Dicção”.

5.4 Procedimentos

Os procedimentos do presente estudo foram realizados nas seguintes etapas: coleta das amostras de fala na condição inicial (GE e GC), aplicação do treinamento de desempenho comunicativo oral (GE) ou não intervenção (GC), coleta das amostras de fala na condição final (GE e GC), codificação das amostras de fala, aplicação de análise perceptivo-auditiva com juízes leigos e análise das amostras de fala (análise acústica e análise da fluência de fala). Tanto a coleta de dados quanto a análise das amostras de fala foram realizadas de forma cega.

5.4.1 Coleta das amostras de fala

A gravação das amostras de fala foi direcionada por dois textos argumentativos que apresentavam dois pontos de vista distintos sobre um mesmo tema (**Anexo 2**). Os textos foram apresentados aos participantes oralmente por um fonoaudiólogo que não estava envolvido no processo de intervenção. Os participantes foram orientados a argumentar de forma espontânea e livre sobre o tópico apresentado por um período de um minuto, expondo seus argumentos pessoais a respeito do tema.

Visando a maior padronização do enunciado a ser analisado (e a manutenção das características segmentais da fala), os participantes foram orientados a iniciar suas argumentações utilizando a seguinte frase: *Vou explicar agora porque concordo plenamente com os seguintes argumentos.*

O fonoaudiólogo responsável pela leitura dos textos permaneceu na sala com os participantes e foi responsável por cronometrar o tempo de argumentação dos sujeitos. Tanto a apresentação dos textos aos participantes quanto as gravações das amostras foram realizadas de forma individual, sendo o participante e o fonoaudiólogo

responsável pela coleta os únicos indivíduos presentes no ambiente da coleta de dados. As gravações foram realizadas em ambiente silencioso e reservado.

O procedimento descrito acima foi utilizado tanto para a coleta das amostras de fala na condição inicial quanto para a coleta das amostras de fala na condição final, nos grupos experimental e controle.

O gravador utilizado foi o Gravador Zoom H1 Handy Recorder - gravador digital estéreo com ajuste em formato de áudio .wav 48k/16 bit, apoiado sobre tripé de mesa com distância de 30 centímetros entre o gravador e o falante.

Após a finalização da coleta, as frases iniciais foram editadas dos áudios dos participantes e codificadas, de forma que informações como nome do participante, grupo e condição inicial e/ou final não fossem identificáveis.

5.4.2 Treinamento de desempenho comunicativo oral

O treinamento de desempenho comunicativo oral ao qual o grupo experimental foi submetido possuía o objetivo de aumentar o nível de certeza da fala. Não foi informado aos participantes, entretanto, o objetivo do treinamento. O treinamento foi realizado ao longo de 14 semanas, com periodicidade semanal, por meio de encontros presenciais com duração de 50 minutos, totalizando, assim, 11 horas e 40 minutos de intervenção.

O treinamento foi aplicado por uma fonoaudióloga da equipe que possui dez anos de experiência no treinamento de desempenho comunicativo oral de alunos. Os fluxogramas de decisão utilizados no treinamento aplicado podem ser visualizados na **Figura 1**. Os fluxogramas foram elaborados com base em uma revisão de literatura prévia(153).

O treinamento abordou a velocidade de fala, as pausas, as disfluências de fala e a variação melódica. Em relação à velocidade de fala, no caso de velocidade lenta ou moderada por parte do sujeito, foi treinado o aumento do número de palavras emitidas por minuto. Com relação às pausas, foi treinado o uso de pausas curtas. Foi treinada, ainda, a redução das disfluências de fala por parte dos sujeitos e, por fim, uma curva melódica com variação moderada.

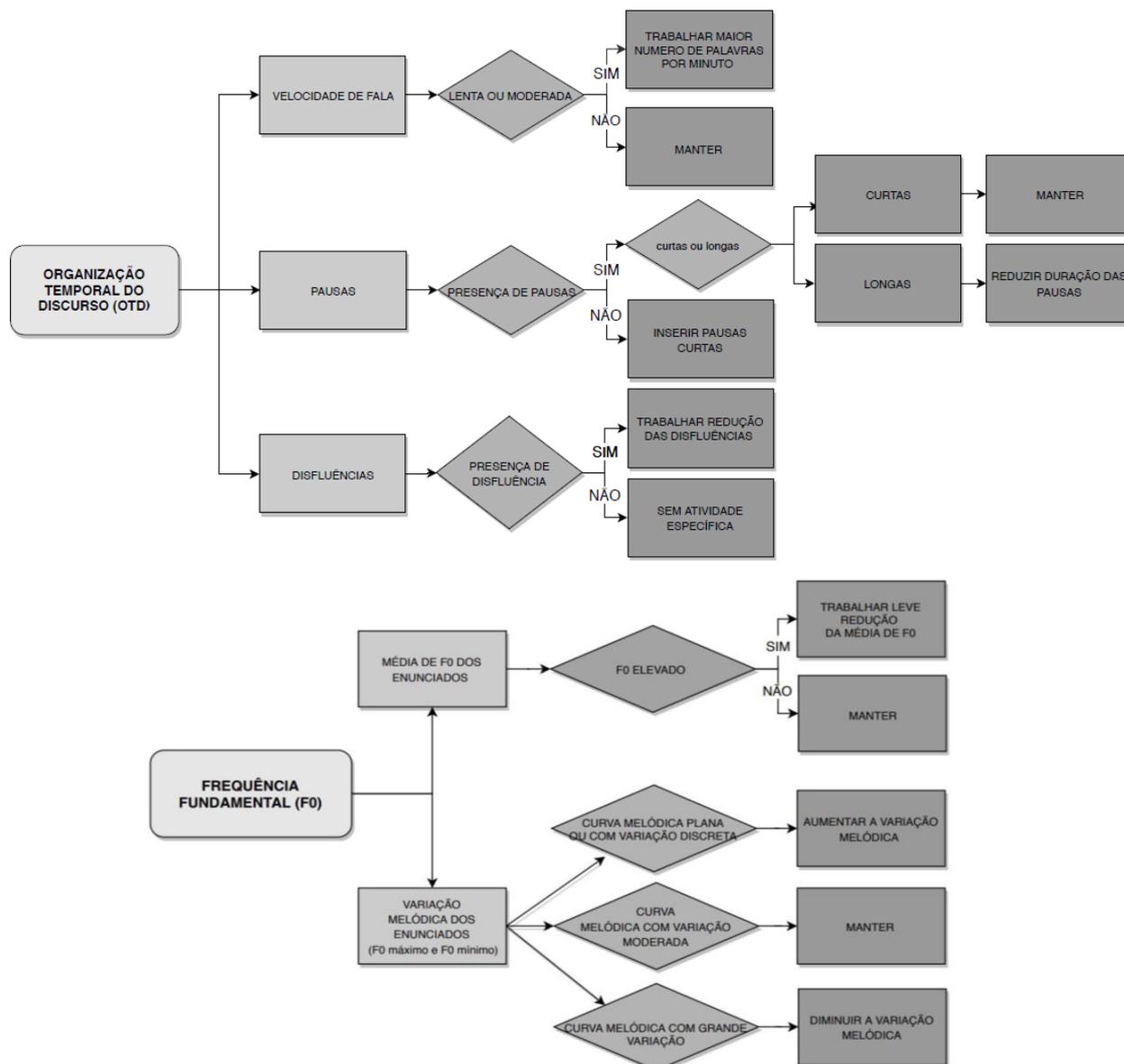


Figura 1. Fluxogramas de decisão utilizados no treinamento de aprimoramento do desempenho comunicativo oral.

5.4.3 Análise perceptivo-auditiva

A análise perceptivo-auditiva foi realizada por vinte e três juízes leigos que participaram da pesquisa de forma voluntária. Os juízes possuíam entre 19 e 61 anos de idade (média = 41,3, mediana = 44 e desvio padrão = 12,3), sendo 15 do sexo feminino e 8 do sexo masculino, com grau de escolaridade variado: especialização (n = 10), ensino superior completo (n = 7), ensino superior incompleto (n = 3), ensino médio completo (n = 1), mestrado (n = 1) e doutorado (n = 1).

A análise perceptivo-auditiva foi realizada de forma *online*, por meio da plataforma Google Forms. Os juízes foram orientados a utilizar fones de ouvido, de forma a favorecer a captação de mais detalhes dos áudios. Os juízes não receberam quaisquer informações a respeito do desenho metodológico do estudo e não estavam cientes da presença de um treinamento de habilidades comunicativas em um dos grupos de participantes. Os juízes deveriam julgar 82 enunciados, compostos pela frase inicial padrão dos participantes dos grupos controle e experimental nas condições inicial (n = 41) e final (n = 41). Os enunciados foram apresentados de forma aleatória aos juízes.

Os juízes foram orientados a avaliar os enunciados quanto ao grau de certeza da fala utilizando a escala psicométrica Likert, com opções de 1 a 5, na qual 1 representa pouca certeza e 5 representa muita certeza. De forma a evitar possíveis consequências da fadiga auditiva nos julgamentos, os participantes foram orientados a realizar intervalos periódicos ao longo da análise perceptivo-auditiva. Não houve restrição quanto ao número de vezes que os juízes poderiam ouvir os enunciados. A duração da análise perceptivo-auditiva foi estimada em 45 minutos.

No total, 1.886 respostas foram computadas.

5.4.4 Análise das amostras de fala

Todas as análises acústicas e conversões no formato dos arquivos de áudios foram realizadas utilizando o programa WinPitch W10 versão 1.0(154). Inicialmente, os arquivos de áudio foram convertidos para mono canal, 16 bits, com frequência de 22050 Hz, em formato .wav (conversão realizada pelo próprio WinPitch).

Em seguida, o conteúdo expresso no enunciado foi transcrito. Os enunciados foram analisados quanto ao número de sílabas por minuto e por segundo(87,155–160); palavras por minuto; número, duração e localização das pausas; presença de disfluências; porcentagem de descontinuidade de fala; porcentagem de disfluências típicas da gagueira; tipo de contorno melódico nas sílabas salientes; diferença em Hertz (Hz) nas sílabas salientes; média das vogais salientes do enunciado (em Hz) e; tipo de organização prosódica.

Os dados foram tabulados utilizando o *software* Microsoft® Excel® para Microsoft 365 MSO (16.0.13328.20334) 32 bits.

Conforme pode ser visualizado na anotação realizada no programa WinPitch, no exemplo da **Figura 2**, as pausas foram sinalizadas pelo caractere especial “_”, na mesma camada de transcrição das palavras (L1). Na segunda camada de texto (L2), foram transcritas as informações adicionais do enunciado, como disfluências, sons de respiração e possíveis ruídos ambientais. A anotação das sílabas salientes do enunciado (tipo de contorno melódico) foi realizada após as demais análises. Ainda na **Figura 2**, pode ser visualizada a janela do WinPitch com o espectrograma ao fundo (ainda não alinhado à curva de F0) e os contornos melódicos demarcados na amostra em questão.

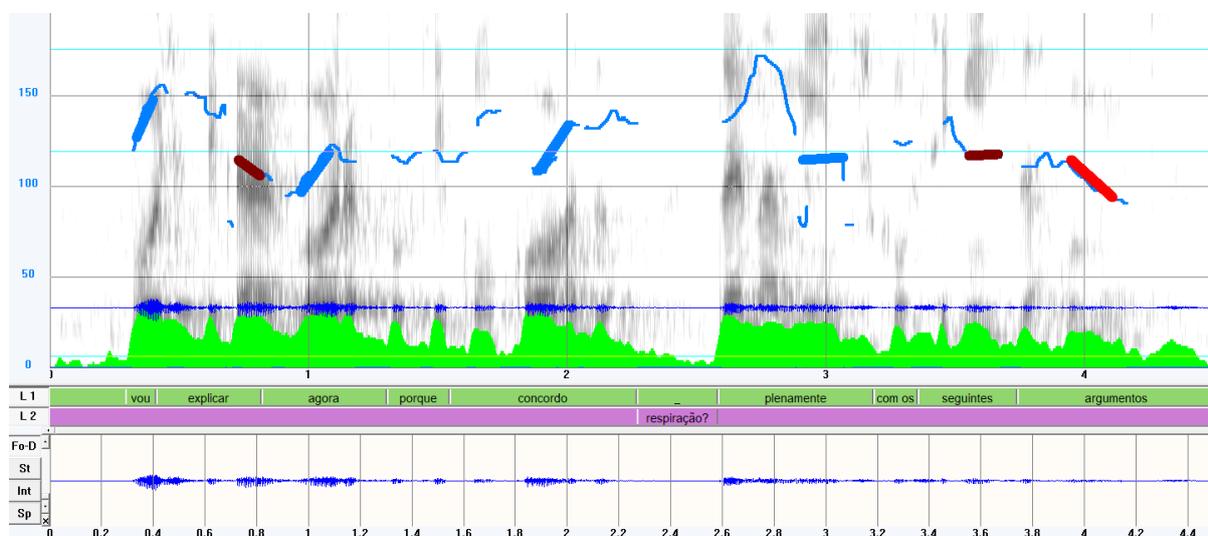


Figura 2. Exemplo do padrão de anotação adotado.

Os mesmos procedimentos de análise e tabulação foram aplicados em todos os 82 arquivos de áudio.

Reuniões de consenso entre os pesquisadores foram realizadas para sanar eventuais dúvidas quanto às aferições das medidas acústicas e/ou demais análises realizadas na amostra. Todas as anotações de contornos melódicos foram revisadas pelo coorientador do presente trabalho, assim como a classificação das estruturas prosódicas dos enunciados.

5.4.5 Análise da fluência de fala

Para a análise da fluência de fala dos enunciados coletados, foram utilizados os procedimentos descritos por Andrade(59). Os valores de referência para as medidas de fluência de fala de falantes do português brasileiro (rupturas e velocidade de fala) foram considerados de acordo com o estudo de Martins e Andrade(161).

Quanto à tipologia das disfluências, foram consideradas disfluências comuns as hesitações, interjeições, revisões, repetições de palavras e segmentos, pausas e palavras não terminadas. Já enquanto disfluências típicas da gagueira foram considerados os bloqueios, os prolongamentos dentro de palavras e as repetições de sons e sílabas.

No que diz respeito à frequência de rupturas, foi calculada a porcentagem de descontinuidade de fala a partir do número de sílabas do enunciado e do número total de disfluências comuns (DC) e disfluências típicas da gagueira (DTG) produzidas, seguindo a seguinte fórmula:

$$\frac{DC + DTG}{\text{número de sílabas emitidas}} \times 100$$

A porcentagem de disfluências típicas da gagueira (DTG) foi calculada pela contabilização do número de sílabas emitidas e o número total de disfluências típicas da gagueira emitidas na amostra, conforme fórmula abaixo:

$$\frac{DTG}{\text{número de sílabas emitidas}} \times 100$$

Ressalta-se que, conforme o conhecimento do português brasileiro, foram consideradas como sílabas as unidades com conformação CVC, CV, CCV, CVV, V, CVCC(162), VC, VCC, CCVC, VV, CCVVC e CCVCC. Unidades sem núcleo vocálico, como as africadas, não foram contabilizadas como sílabas para fins de análise na presente pesquisa.

Os cálculos para a mensuração da velocidade de fala foram realizados a partir do número de sílabas e de palavras emitidas em relação à duração total do enunciado

(em segundos), incluindo ou não o tempo de não emissão (em segundos), conforme as seguintes fórmulas:

Sílabas por segundo (SPS):

$$\frac{\text{número de sílabas da amostra}}{\text{tempo total da amostra}}$$

Sílabas por minuto (SPM):

$$\frac{\text{número de sílabas da amostra} \times 60 \text{ segundos}}{\text{tempo total da amostra}}$$

Palavras por minuto (PPM):

$$\frac{\text{número de palavras da amostra} \times 60 \text{ segundos}}{\text{tempo total da amostra}}$$

Taxa de articulação (TA) de sílabas:

$$\frac{\text{número de sílabas da amostra} \times 60 \text{ segundos}}{\text{tempo da amostra} - \text{tempos de pausa/não emissão}}$$

Taxa de articulação (TA) de palavras:

$$\frac{\text{número de palavras da amostra} \times 60 \text{ segundos}}{\text{tempo da amostra} - \text{tempos de pausa/não emissão}}$$

5.4.6 Grupo acentual e contornos melódicos

As sílabas salientes e os grupos acentuais presentes na emissão esperada da frase utilizada para as amostras de fala no presente trabalho podem ser visualizados na **Figura 3**. A transcrição fonética foi realizada conforme o dialeto predominante na região de coleta das amostras de fala (Brasília – DF). Foram considerados os parâmetros acústicos para a definição das sílabas enquanto salientes, de forma que as sílabas que apresentaram destaque em seu contorno duracional (pico de duração) e em sua frequência fundamental (picos e vales da curva de F0) foram demarcadas

enquanto constituintes centrais dos grupos acentuais (acentos frasais, na função perceptiva).

vou	explicar			agora			porque			concordo			plenamente		
vou	ɪsplɪkɐh			agɔrɐ			pɐhke			kõkõhdɔ			plɛnɛmẽtʃɪ		
vou	ex	pli	car	a	go	ra	por	que	con	cor	do	ple	na	men	te
GA	GA			GA			GA			GA					

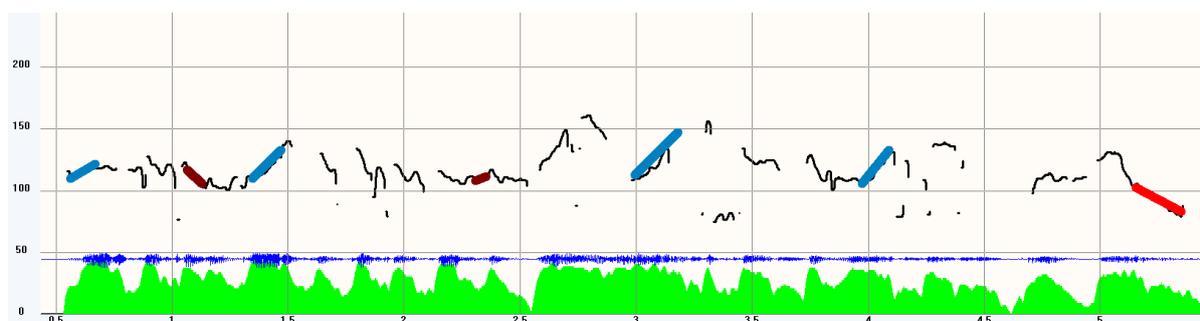
com	os	seguintes			argumentos			
kõ	us	segĩtʃɪs			ahgumẽtus			
com	os	se	guin	tes	ar	gu	men	tos
GA					GA			

Legenda:
 Sílabas salientes em negrito.
 GA = grupo acentual

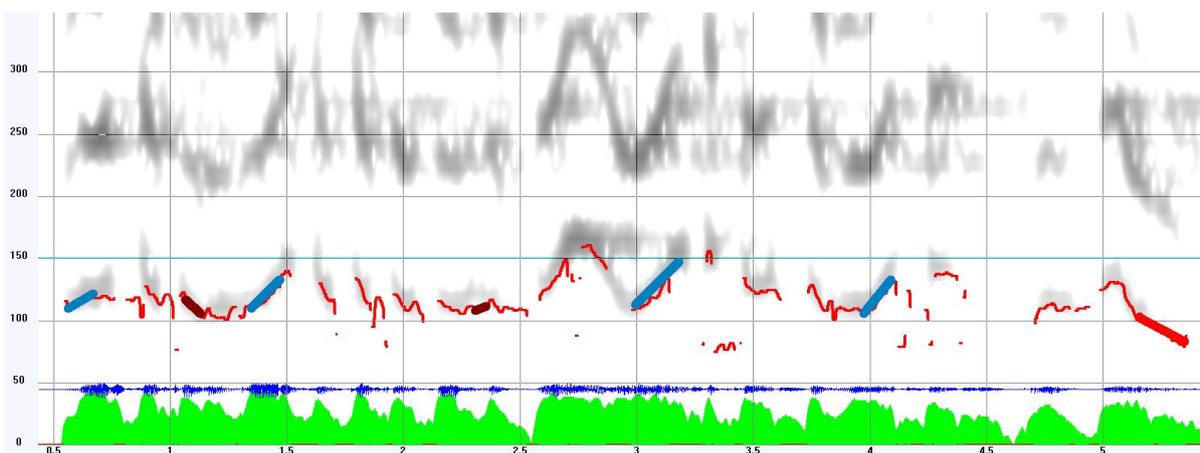
Figura 3. Grupos acentuais presentes na emissão esperada da frase padronizada para a amostra de fala.

A análise da curva de F0 e de grupos acentuais (GA) foram realizadas por meio do programa WinPitch(154). Foram utilizados os métodos gerais de rastreamento de F0 *Spectral comb* e *Autocorrelation*, assim como a análise da curva de F0 utilizando o espectrograma de banda estreita, com auxílio da função Align F0. O algoritmo utilizado pelo programa WinPitch por meio do rastreamento de F0 *Spectral comb* e *Autocorrelation* elimina a problematização da utilização do espectrograma de banda estreita para o estudo da dimensão tempo.

Inicialmente, realizou-se a marcação das sílabas salientes e vogais salientes dos enunciados (**Figura 4**). A vogal da última sílaba saliente do enunciado foi a primeira a ser demarcada, fornecendo, assim, parâmetros para que o programa WinPitch pudesse realizar a comparação tonal com as demais vogais salientes do enunciado (**Tabela 1**). Ressalta-se que a marcação das sílabas salientes foi realizada manualmente, sendo a definição do tipo de contorno melódico presente na sílaba em questão realizada de forma automática pelo programa WinPitch.



(a) Marcação das sílabas salientes sem a visualização do espectrograma. Podem ser observados, respectivamente e da esquerda para a direita, os contornos C1, Cn, C1, Cn, C1, C1 e C0.



(b) Marcação das sílabas salientes no mesmo enunciado exemplificado em (a), porém contendo a visualização do espectrograma (harmônicos) alinhado à curva de F0.

Figura 4. Exemplo de marcação das sílabas salientes do enunciado.

Tabela 1. Vogais salientes do enunciado e sua ordem de anotação.

Vogais salientes	v <u>o</u> u	explic <u>a</u> r	ag <u>o</u> ra	conco <u>o</u> rdo	plena <u>m</u> ente	segu <u>i</u> ntes	argum <u>e</u> ntos
Ordem de anotação	2º	3º	4º	5º	6º	7º	1º

O contorno da curva de F0 no local das sílabas salientes do enunciado foi anotado e codificado (**Tabela 2**) por meio de esquema de cores, conforme a seguinte notação(102,103):

Tabela 2. Tipos de contorno entoacional e suas respectivas formas de notação.

Sigla	Definição	Cor de notação
C0	Contorno descendente, de entonação final.	Vermelho
C1	Mudança de F0 percebida pelo ouvido humano como alteração melódica ascendente.	Azul
C2	Queda de frequência percebida como mudança descendente pelo ouvido humano.	Verde
Cn	Mudança de F0 muito discreta para ser percebida pelo ouvido humano; tom estático.	Marrom
Cc	Contorno complexo. Plano ou descendente na sílaba saliente do GA. Quando na sílaba final do GA, é ascendente.	Azul claro

As informações a respeito de cada contorno melódico saliente foram extraídas do programa WinPitch(154) para uma planilha do software Microsoft® Excel®. A planilha oriunda deste processo, única para cada enunciado, continha informações a respeito das variáveis acústicas de cada sílaba saliente da amostra, sendo elas: tipo de contorno, sua duração (ms), intensidade (dB), frequência (Hz), frequência máxima e mínima e diferença em Hz. O tipo de contorno e sua diferença em Hz de todos os contornos salientes de todos os enunciados foram extraídos e compilados.

A estrutura prosódica do enunciado foi gerada pelo programa WinPitch(154) por meio da função *Prosodic structure*, com base nas marcações realizadas nas sílabas salientes. A estrutura prosódica do enunciado foi determinada conforme a hierarquia existente entre os contornos melódicos ($C_n < C_1 < C_2 < C_c < C_0$). As estruturas prosódicas de todos os enunciados foram extraídas e classificadas quanto à sua organização. A classificação das estruturas prosódicas dos enunciados foi realizada pela pesquisadora principal e verificada pelo coorientador do trabalho.

5.4.7 Pausas

As pausas foram analisadas quanto à sua localização, quantidade e duração. A localização das pausas foi avaliada em relação às palavras e aos grupos acentuais, de forma a verificar se as pausas corresponderiam ou não às fronteiras sintáticas do enunciado.

A análise de duração das pausas foi realizada por meio do programa WinPitch W10 versão 1.0(154). As pausas foram transcritas na camada de texto (L1) utilizando o caractere especial “_” e foram demarcadas utilizando os seguintes critérios: foi considerado, como ponto de partida, o último ciclo periódico/regular das vogais em final de palavra e, como ponto final, o início da emissão do fonema seguinte(17,163) (**Figura 5**). Foi utilizada a presença de harmônicos no espectrograma, quando necessário, para determinar a presença ou a ausência do sinal de fala.

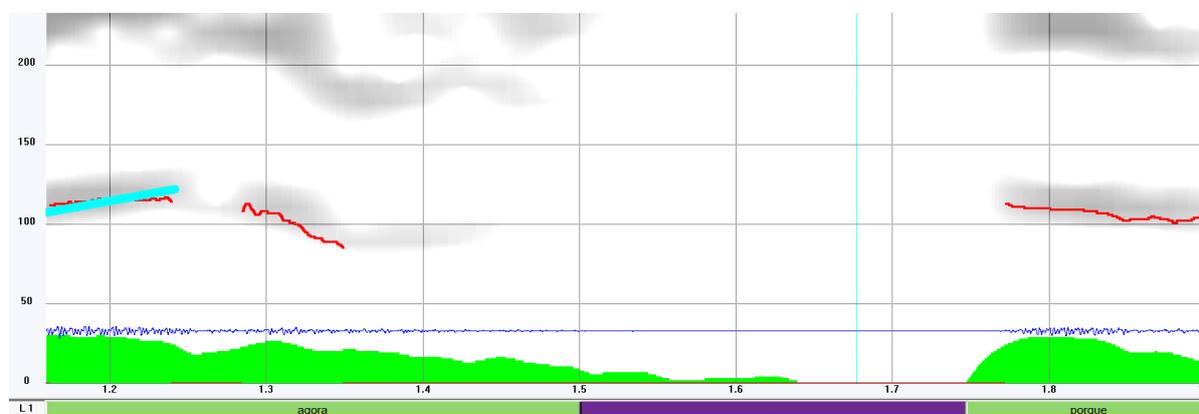


Figura 5. Exemplo de demarcação das pausas nos enunciados.

Considerando a capacidade de reconhecimento do ouvinte(102), a precisão de demarcação de 20 ms foi considerada como ótima para a demarcação dos limites da pausa. Foi considerado como pausa qualquer tempo de silêncio entre o início da emissão da primeira sílaba da primeira palavra do enunciado e o final da emissão da última sílaba da última palavra da frase proposta para análise, independente se menor que 0,250 s(164–166).

5.5 Análise estatística

O teste de normalidade de Kolmogorov-Smirnov foi utilizado para verificar a distribuição dos dados. A análise dos dados foi realizada tanto de forma inferencial quanto de forma descritiva. A análise estatística inferencial foi realizada tanto por meio de testes paramétricos quanto por meio de testes não paramétricos, a depender da distribuição e características da análise pretendida. A análise estatística descritiva consistiu no uso de gráficos com valores absolutos e médios e tabelas de frequência. Todos os valores médios foram calculados por meio de médias aritméticas.

A análise da variância (ANOVA) unidirecional foi realizada para verificar as medidas repetidas intra grupos e entre grupos. Para a análise estatística inferencial dos resultados da análise perceptivo-auditiva, foi utilizado o teste t de amostras independentes para a comparação entre as médias dos grupos. A correlação entre os parâmetros acústicos e as notas da análise perceptivo-auditiva foi realizada por meio do teste de correlação de Pearson ou Spearman, a depender da distribuição das variáveis a serem analisadas.

As análises estatísticas foram realizadas por meio do software IBM® SPSS® Statistics Versão 25. A significância adotada foi de 5% ($p < 0,05$).

6. Resultados

Foram analisados 82 enunciados emitidos por 41 participantes da pesquisa, sendo 20 indivíduos do GE e 21 indivíduos do GC. Os enunciados foram coletados e analisados considerando as condições inicial (para GE, pré treinamento) e final (para GE, pós treinamento), assim como as classificações relativas às notas obtidas por cada enunciado na análise perceptivo-auditiva (APA). Vinte e três juízes leigos originaram 1.886 respostas na APA, as quais foram computadas de forma individual e, posteriormente, em forma de média para cada um dos enunciados.

6.1 Análise perceptivo-auditiva

A análise descritiva das notas médias atribuídas ao nível de certeza dos sujeitos dos grupos experimental e controle pelos juízes na análise perceptivo-auditiva (**Gráfico 1**) apontou que o grupo experimental na condição pós treinamento apresentou a maior média de avaliação (média = 2,958, mediana = 3,260, desvio padrão = 0,725). O grupo controle na condição pós treinamento, por sua vez, apresentou média de 2,262 (mediana = 2,217, desvio padrão = 0,827).

Quanto aos resultados descritivos das notas da análise perceptivo-auditiva na condição pré treinamento, o grupo experimental apresentou média de 2,617 (mediana = 2,500, desvio padrão = 0,668) e o grupo controle apresentou média de 2,430 (mediana = 2,565, desvio padrão = 0,819).

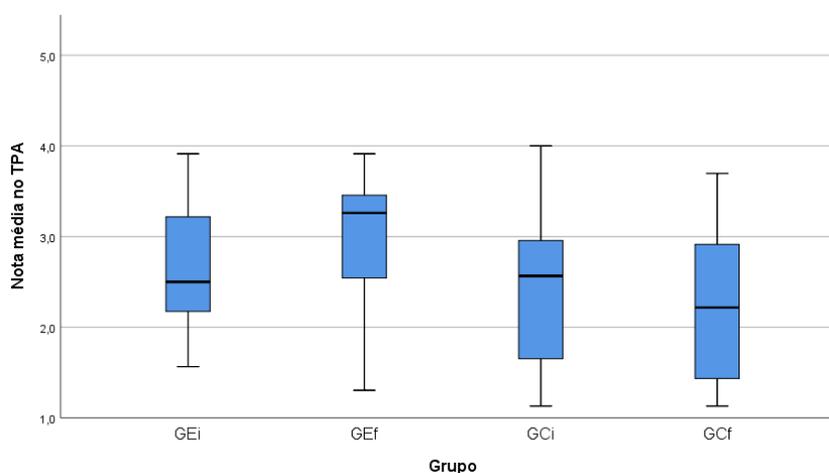


Gráfico 1. Boxplot simples apresentando os dados descritivos das notas obtidas na análise perceptivo-auditiva pelos grupos experimental e controle nas condições inicial e final (pré e pós treinamento).

As notas médias obtidas individualmente na análise perceptivo-auditiva nas condições inicial e final pelos sujeitos que compuseram o grupo experimental (**Gráfico 2**) e o grupo controle (**Gráfico 3**) podem ser visualizadas abaixo.

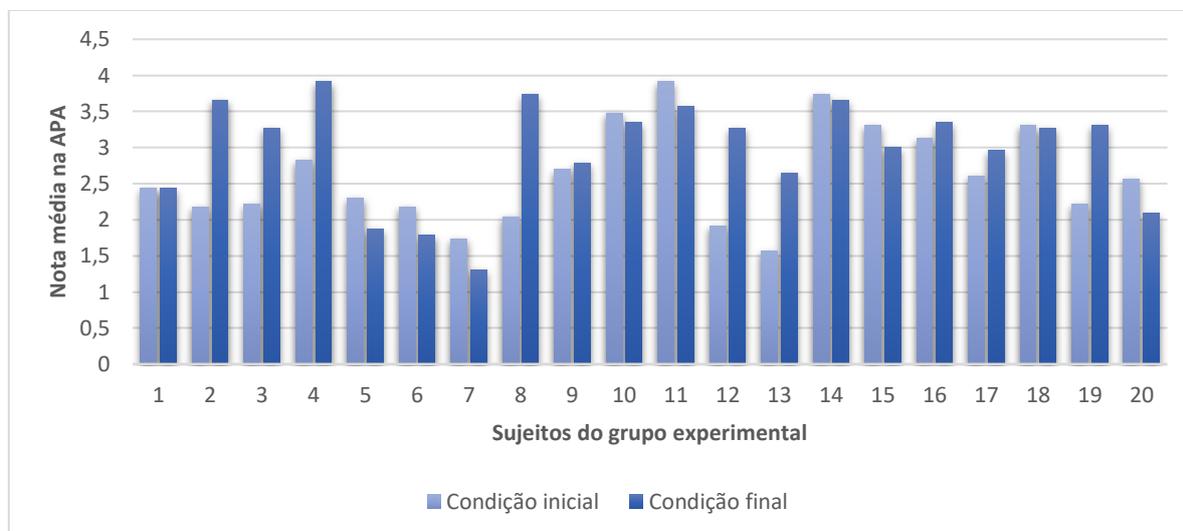


Gráfico 2. Média das notas individuais recebidas na análise perceptivo-auditiva pelos sujeitos do grupo experimental nas condições inicial e final.

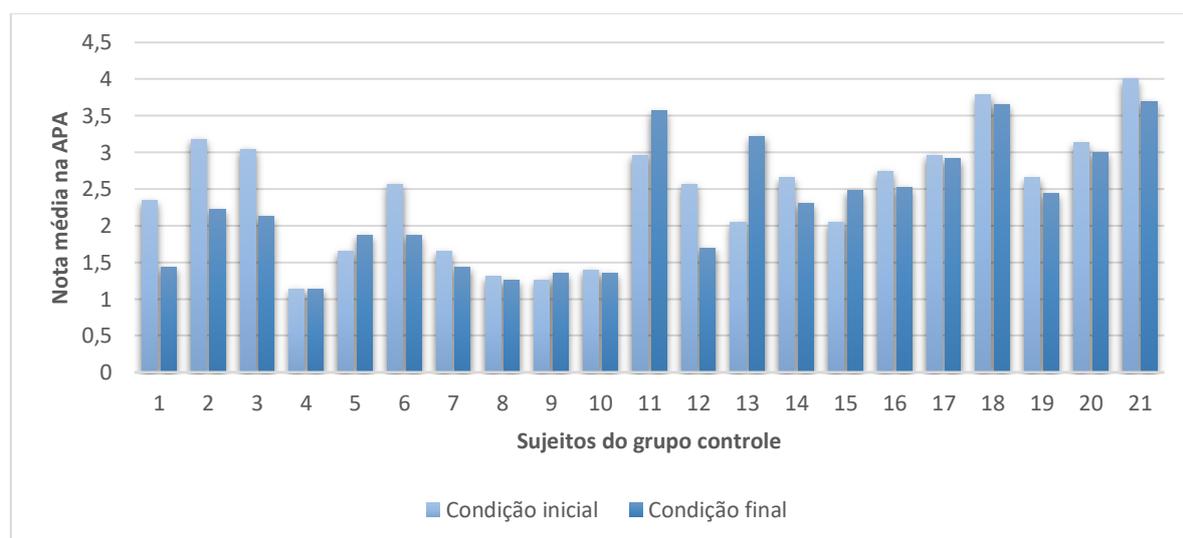


Gráfico 3. Média das notas individuais recebidas na análise perceptivo-auditiva pelos sujeitos do grupo controle nas condições inicial e final.

Na comparação entre os grupos experimental e controle na condição final, o teste de Levene indicou igualdade de variâncias ($p=0,456$). O teste t indicou diferença entre as médias dos grupos ($p=0,007$) (**Tabela 3**). O teste t independente mostrou que, em média, os participantes do grupo experimental na condição final apresentaram nota superior na análise perceptivo-auditiva quando comparados aos participantes do grupo controle na condição final ($t(39) = 2,857$; $p < 0,05$).

Tabela 3. Teste t de amostras independentes entre os grupos GEf e GCf em relação às notas na APA.

		Teste de amostras independentes								
		Teste de Levene para igualdade de variâncias		teste t para Igualdade de Médias						
		Z	Sig.	T	df	Sig. (2 extremidades)	Diferença média	Erro padrão de diferença	95% Intervalo de Confiança da Diferença	
									Inferior	Superior
Nota média na APA	Variâncias iguais assumidas	,567	,456	2,857	39	,007	,695756	,243548	,203133	1,188378

Por sua vez, o teste t realizado entre os grupos experimental e controle na condição inicial não demonstrou diferença entre as médias ($p=0,430$) (**Tabela 4**). Houve igualdade de variâncias ($p=0,328$). O teste t independente mostrou que, em média, os participantes do grupo experimental na condição inicial não apresentaram nota na análise perceptivo-auditiva superior à nota dos participantes do grupo controle na condição inicial ($t(39) = 0,798$; $p > 0,05$).

Tabela 4. Teste t de amostras independentes entre os grupos Gei e Gci em relação às notas na APA.

		Teste de amostras independentes								
		Teste de Levene para igualdade de variâncias		teste t para Igualdade de Médias						
		Z	Sig.	t	df	Sig. (2 extremidades)	Diferença média	Erro padrão de diferença	95% Intervalo de Confiança da Diferença	
									Inferior	Superior
Nota média na APA	Variâncias iguais assumidas	,980	,328	,798	39	,430	,186749	,234149	-,286862	,660361

A análise de variância (ANOVA) unidirecional (**Tabela 5**) também apontou o efeito do grupo sobre a nota média na análise perceptivo-auditiva.

Tabela 5. Resultado da análise de variância unidirecional para as notas na APA.

		Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Nota média na APA	Entre Grupos	5,449	3	1,816	3,107	,031
	Nos grupos	45,601	78	,585		
	Total	51,050	81			

6.2 Análise da fluência de fala

Dentre os 82 enunciados analisados, apenas dois apresentaram disfluências típicas da gagueira. Ambos enunciados foram produzidos por indivíduos do grupo controle, na condição inicial. Um dos enunciados que apresentou disfluências típicas da gagueira compôs o grupo dos quinze enunciados que obtiveram as menores notas na análise perceptivo-auditiva.

As disfluências típicas da gagueira emitidas foram o prolongamento dentro de palavra e o bloqueio, ambas na palavra 'seguintes'. O prolongamento dentro de palavra (s_eguintes) foi a única disfluência presente em seu enunciado. O bloqueio foi precedido por um prolongamento em final de palavra (com_ /seguintes). Os enunciados contendo disfluências típicas da gagueira (ambos produzidos por sujeitos do GCi) apresentaram % de DTG de 4,347% e 4,166%.

Quanto às disfluências comuns, as revisões estavam presentes em quatro enunciados. Todos os quatro enunciados foram emitidos na condição final, sendo um no GE e três no GC. Dois enunciados apresentaram a hesitação '#é' antes da palavra 'vou', os enunciados foram emitidos por um participante do GE na condição inicial e um participante do GC na condição final. Sete enunciados apresentaram duração aumentada de uma sílaba em relação às sílabas adjacentes, seis das quais foram produzidas na última sílaba da palavra e uma produção com ocorrência dentro de palavra (ago_ra). Os sete prolongamentos observados foram produzidos em palavras diversas do enunciado, com coocorrência apenas para a palavra 'seguintes' (n = 2).

Considerando as pausas enquanto disfluências de fala, a maior taxa de descontinuidade de fala observada foi de 9,523%, em um sujeito do GE na condição final. O valor apresentado pelo participante se deveu exclusivamente às pausas produzidas em seu enunciado. Ainda considerando as pausas enquanto disfluências, a menor taxa de descontinuidade de fala observada foi de 0% (n = 33). O grupo com a maior média de descontinuidade de fala foi o GC na condição final (média = 3,653%). O grupo com a menor média de descontinuidade de fala foi o GE na condição inicial (média = 2,999%). O GE na condição final apresentou média de 3,222% de descontinuidade de fala e o GC na condição inicial apresentou média de 3,295% de descontinuidade de fala.

Desconsiderando as pausas enquanto disfluências, a maior taxa de descontinuidade de fala observada foi de 8,695%, em um sujeito do GC na condição inicial. A menor taxa de descontinuidade de fala observada foi de 0% (n = 69). Ainda desconsiderando as pausas enquanto disfluências de fala, o grupo com a maior média de descontinuidade de fala foi o GE na condição final (média = 1,100%). O valor apresentado pelo GE na condição final foi devido aos quatro sujeitos que emitiram prolongamentos típicos e a um sujeito que realizou uma revisão em seu enunciado. O grupo com a menor descontinuidade de fala foi o GE na condição inicial (média = 0,238). O GC apresentou média de 0,819% de descontinuidade de fala na condição inicial e 0,969% na condição final.

O valor médio de sílabas por minuto apresentado pelo GE foi de 305,480 SPM para a condição inicial e de 316,192 SPM para a condição final. Já o GC apresentou média de 329,743 SPM na condição inicial e 316,277 SPM na condição final. Os valores de sílabas por segundo não apresentaram grande variação entre os grupos, sendo o GC na condição inicial o grupo com maior número de emissões médias de sílabas por segundo (5,495) e o GE na condição inicial o grupo com o menor número de emissões médias de sílabas por segundo (5,091).

Em relação aos valores médios de palavras por minuto, o GE apresentou 129,294 PPM na condição inicial e 133,991 PPM na condição final. Já o GC apresentou média de 128,996 PPM na condição inicial e 122,940 PPM na condição final.

Os valores de taxa de articulação de sílabas e taxa de articulação de palavras foram muito próximos aos valores de sílabas por minuto e de palavras por minuto, respectivamente. Não houve variação superior a 0,228 na comparação entre os valores. A média de diferença entre os valores de palavra por minuto e taxa de articulação de palavras foi de 0,159. A média de diferença entre os valores de sílabas por minuto e taxa de articulação de sílabas foi de 0,16.

A análise post-hoc demonstrou que, em média, as variáveis taxa de articulação de palavras e palavras por minuto do grupo experimental na condição final foram diferentes das variáveis taxa de articulação de palavras e palavras por minuto do grupo controle na condição final (**Tabela 6**).

Tabela 6. Resultados da análise de comparações múltiplas.

Variável dependente		(I)	(J)	Diferença média (I-J)	Erro	Sig.	Intervalo de Confiança 95%	
		Grupo	Grupo				Limite inferior	Limite superior
Taxa artic. palavras	DMS	GEf	GCf	11,086657*	4,913387	,027	1,30485	20,86846
		GCf	GEf	-11,086657*	4,913387	,027	-20,86846	-1,30485
Palavras por minuto	DMS	GEf	GCf	11,051867*	4,876291	,026	1,34392	20,75982
		GCf	GEf	-11,051867*	4,876291	,026	-20,75982	-1,34392

*. A diferença média é significativa no nível 0.05.

6.3 Grupo acentual e contornos melódicos

Conforme apresentado anteriormente, quatro enunciados apresentaram erros de emissão e seguinte correção. Em três desses enunciados os falantes procederam à reemissão de todo o grupo acentual. No enunciado que não houve reemissão de todo o grupo acentual, o sujeito reemitiu apenas a palavra acometida pelo erro (ser- _ seguintes), a qual era parte de um grupo acentual composto por três palavras.

Em relação aos valores da diferença em Hz apresentada pelos diferentes contornos melódicos, não houve diferença significativa entre os valores do contorno C0, executado pelos falantes apenas na sílaba saliente da última palavra do enunciado.

Na análise dos valores dos contornos C1 e Cn, executados pelos falantes na vogal saliente da palavra 'vou', houve diferenciação do grupo experimental na condição final, apontando maior variação de F0 (**Gráfico 4**).

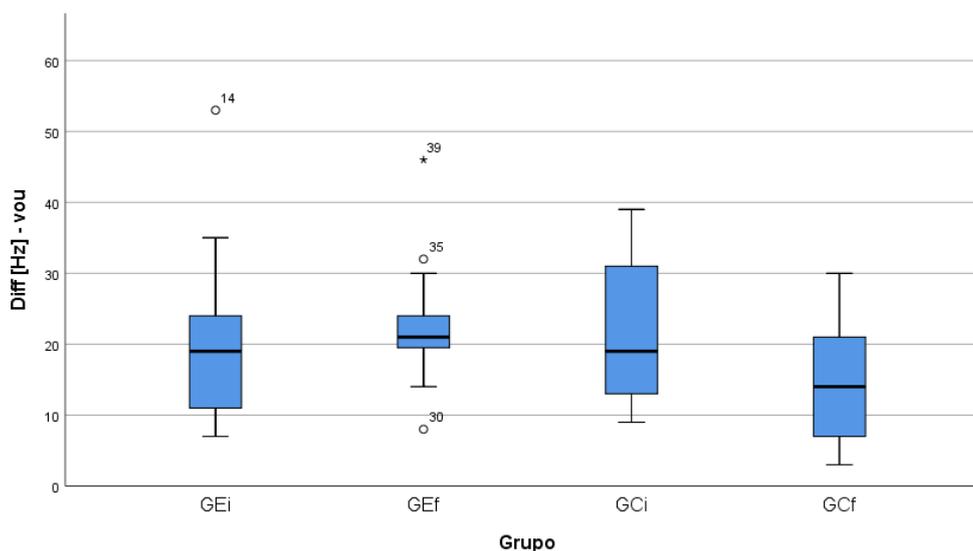
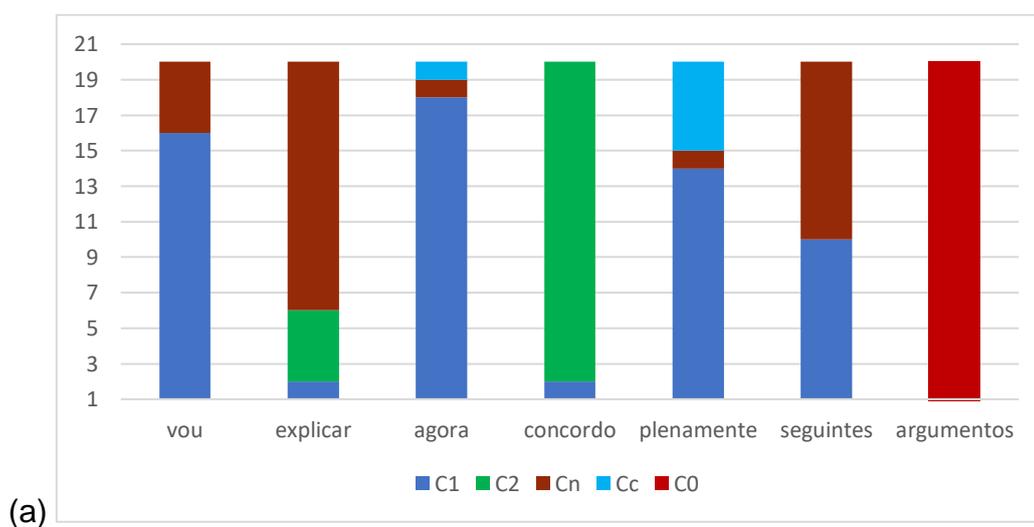


Gráfico 4. Boxplot simples apresentando os dados descritivos em relação aos valores da diferença em Hz apresentada pelos sujeitos na palavra 'vou', por grupo.

A frequência de ocorrência de cada tipo de contorno nas sílabas salientes do enunciado pode ser observada nos gráficos a seguir. Serão apresentados os tipos de contorno expressos pelos grupos experimental (**Gráfico 5**) e controle (**Gráfico 6**) nas condições inicial e final.



(Continuação do Gráfico 5)

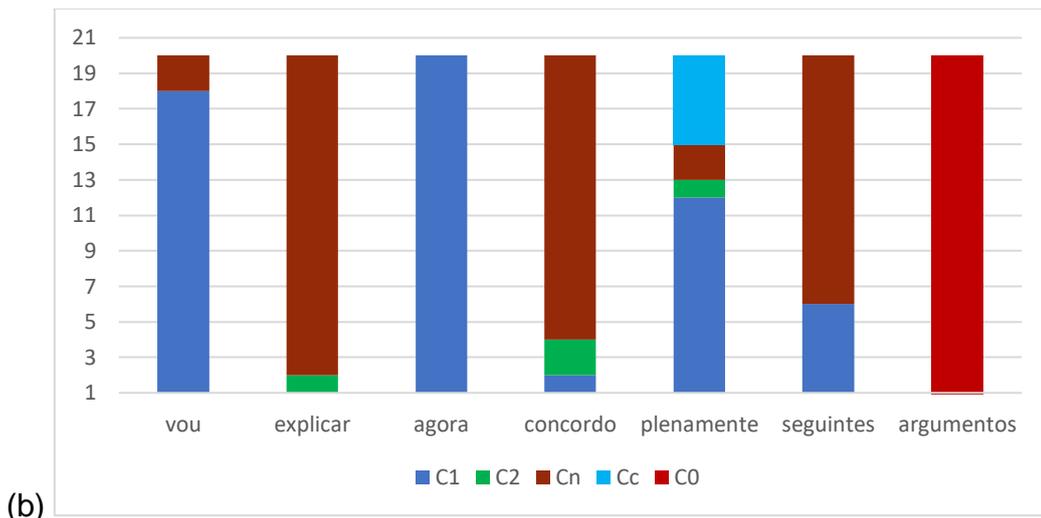


Gráfico 5. Tipos de contornos melódicos apresentados pelo grupo experimental nas condições inicial (a) e final (b).

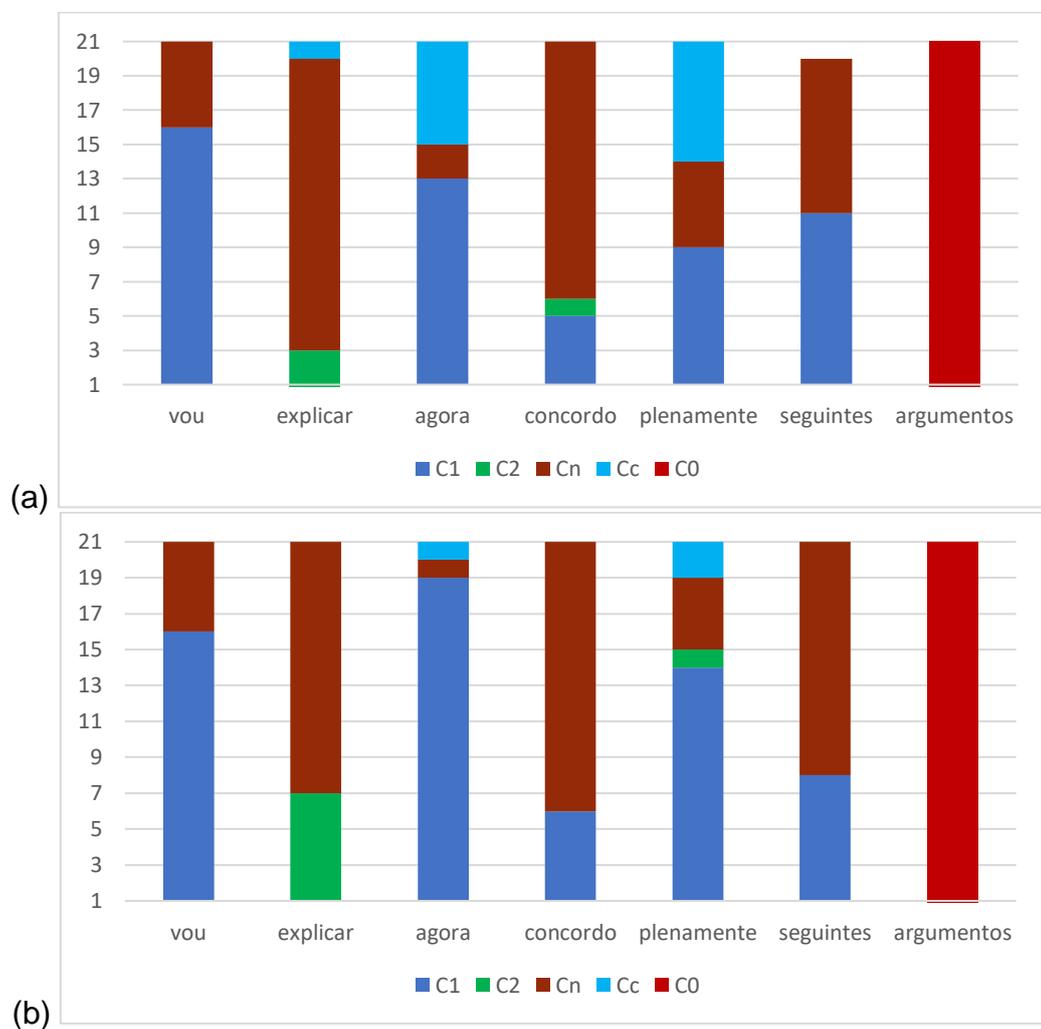
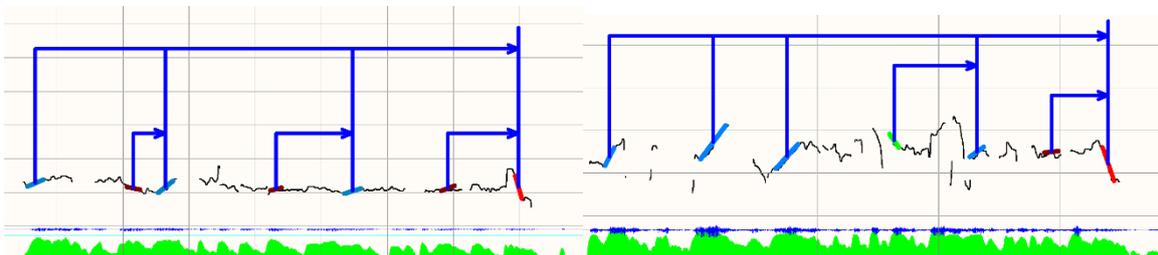


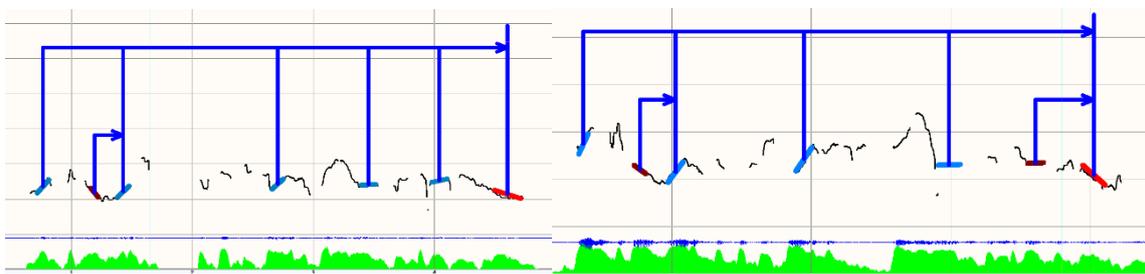
Gráfico 6. Tipos de contornos melódicos apresentados pelo grupo controle nas condições inicial (a) e final (b).

Quanto à análise da estrutura prosódica dos enunciados, os resultados dos sujeitos que obtiveram melhora significativa (aumento de 0,5 ou mais nas médias de notas da APA) no GE estão expostos na **Figura 6**.

Sujeito 2: mudança de organização prosódica do tipo 1 para o tipo 20.



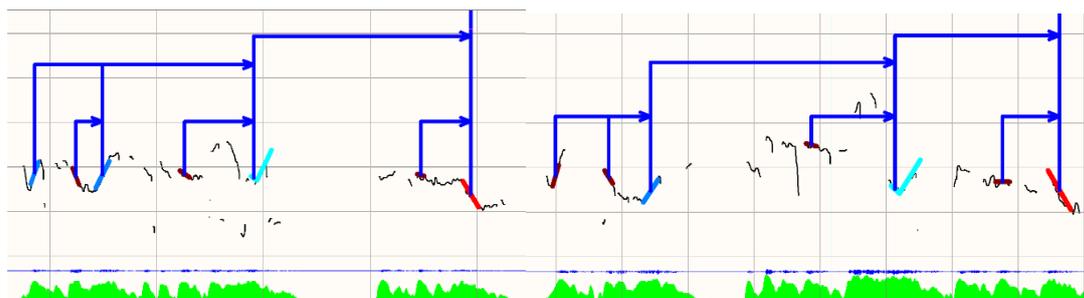
Sujeito 3: mudança de organização prosódica do tipo 3 para o tipo 21.



Sujeito 4: mudança de organização prosódica do tipo 4 para o tipo 1.

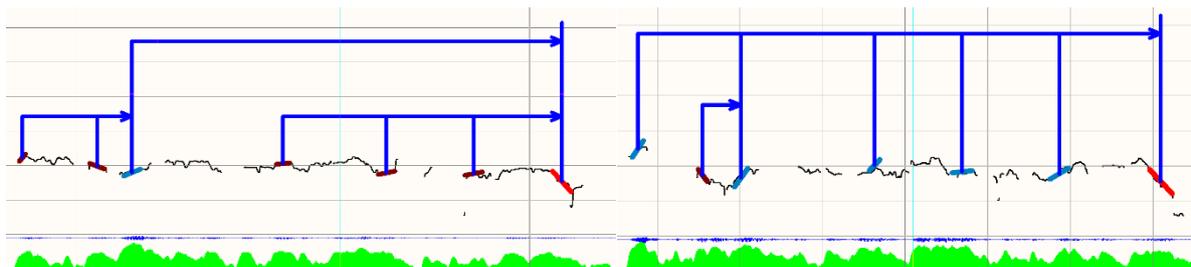


Sujeito 8: mudança de organização prosódica do tipo 2 para o tipo 16.

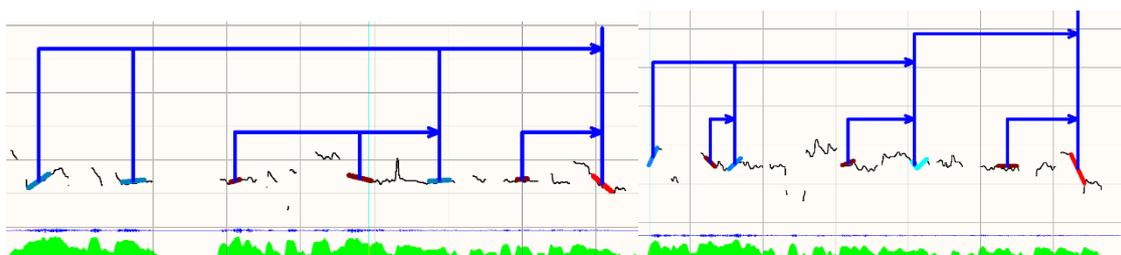


(Continuação da **Figura 6**)

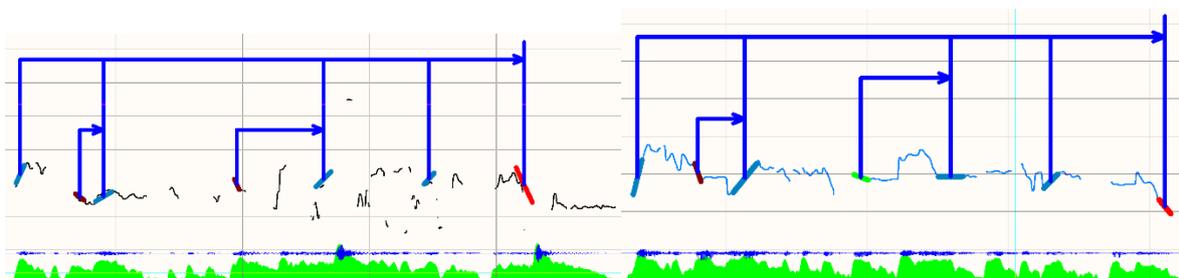
Sujeito 12: mudança de organização prosódica do tipo 10 para o tipo 3.



Sujeito 13: mudança de organização prosódica do tipo 11 para o tipo 2.



Sujeito 19: mudança de organização prosódica do tipo 5 para o tipo 26.



Condição pré treinamento (esquerda) e condição pós treinamento (direita).

Figura 6. Tipos de organização prosódica apresentada pelos sujeitos com melhora significativa no GE nas condições pré e pós treinamento.

A comparação da estrutura prosódica dos enunciados que obtiveram a maior e a menor nota na análise perceptivo-auditiva pode ser visualizada na **Figura 7**.

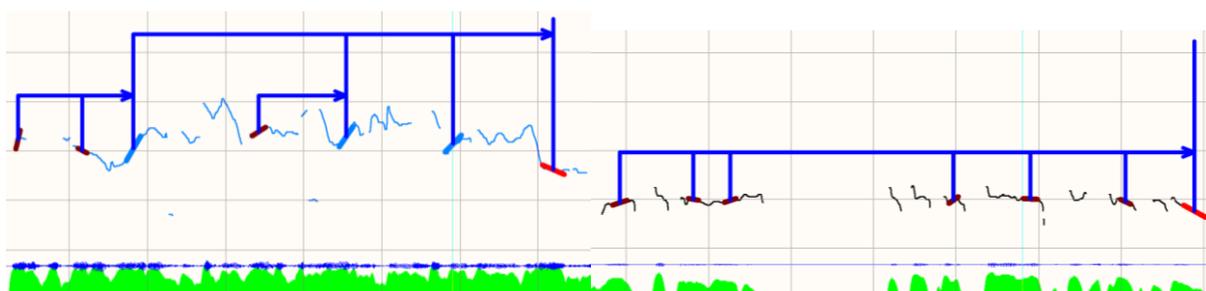


Figura 7. Comparação do tipo de estrutura prosódica dos enunciados que obtiveram a maior (esquerda) e menor (direita) média de notas da APA.

6.4 Pausas

Quarenta e seis pausas foram emitidas entre os quatro grupos da pesquisa. A duração média das pausas produzidas, independente de grupo, foi de 0,283 segundos. Considerando os grupos, a duração média das pausas foi de 0,379 s para o GE na condição inicial e de 0,285 s na condição final, para o GC a duração média das pausas foi de 0,201 s na condição inicial e de 0,263 s na condição final. A maior pausa da amostra (0,694 s), foi emitida por um sujeito do GC na condição final. As quatro maiores pausas da amostra, todas acima de 0,6 s, foram emitidas por sujeitos do GC na condição final ($n = 2$) e sujeitos do GE na condição inicial ($n = 2$).

A duração média da primeira pausa, independente de grupo, foi de 0,300 s. A duração média da segunda pausa, independente de grupo, foi de 0,216 s. O GC na condição final foi o grupo que mais produziu duas pausas no mesmo enunciado ($n = 4$), enquanto o GE na condição final foi o grupo que menos produziu duas pausas no mesmo enunciado ($n = 1$).

As pausas silenciosas tiveram duração média de 0,267 s. Seis pausas tiveram duração inferior a 0,1 s, as quais apresentaram média de duração de 0,084. Dentre as quarenta e seis pausas emitidas, seis continham ruídos respiratórios. As pausas com ruídos respiratórios tiveram duração média de 0,379 s. Três dos seis enunciados que continham pausas com ruídos respiratórios compuseram o grupo dos enunciados que obtiveram as menores notas na análise perceptivo-auditiva.

As pausas respiratórias ocorreram em posições diferentes no enunciado, mas nunca interrompendo um grupo acentual. Houve, porém, uma ocorrência de interrupção de grupo acentual, porém a pausa fez parte de uma revisão, realizada por um sujeito do grupo controle, na condição final.

A duração total média dos enunciados foi de 4,31 s. Em relação aos grupos, o GE na condição final foi o grupo com menor média de duração dos enunciados (4,155 s) e o GC na condição final foi o grupo com maior média de duração dos enunciados (4,527 s). O GE na condição inicial teve 4,280 s de duração média de seus enunciados e o GC na condição inicial teve média de 4,276 s de duração em seus enunciados.

O **Gráfico 7** apresenta o número total de pausas produzidas em relação aos grupos. O grupo experimental na condição final foi o grupo que produziu a menor

quantidade de pausas ($n = 9$), enquanto o grupo controle na condição final apresentou o maior número de pausas ($n = 13$).

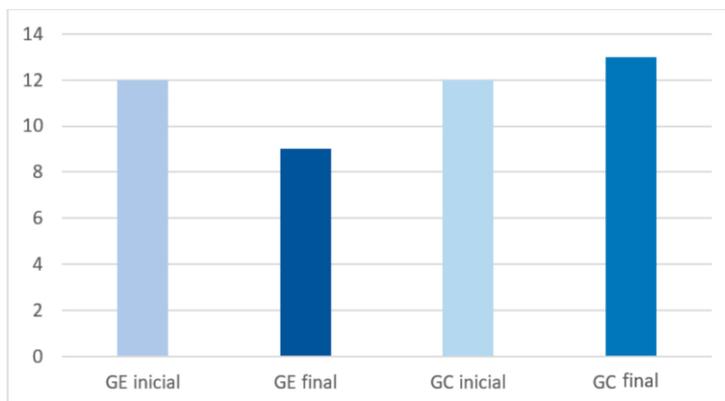


Gráfico 7. Número total de pausas produzidas por grupo.

Quanto à produção de pausas em relação às palavras do enunciado, independente de grupo (**Gráfico 8**), as maiores ocorrências foram observadas após a palavra 'agora' ($n = 23$), seguida da palavra 'plenamente' ($n = 8$). A mesma observação pode ser feita considerando a alocação em grupos (**Gráfico 9**).

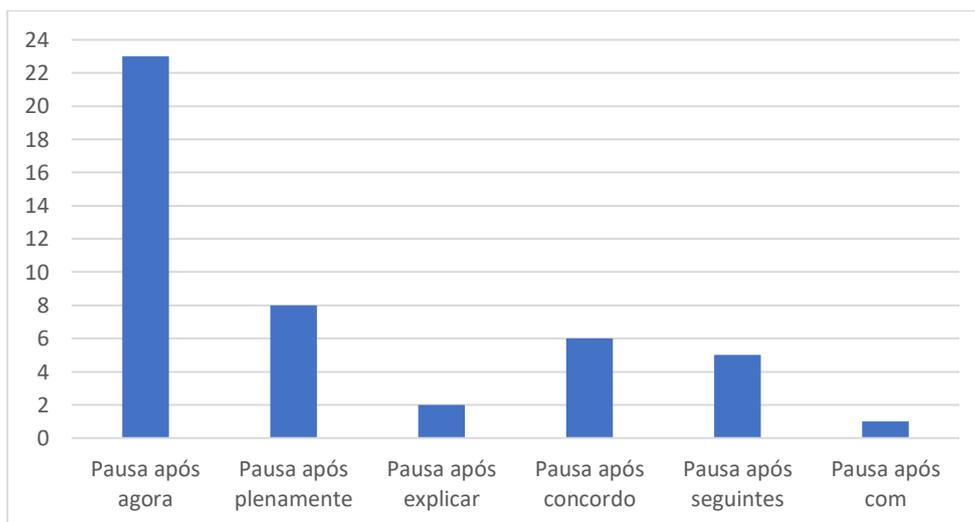


Gráfico 8. Quantitativo de pausas utilizadas e sua localização em relação às palavras do enunciado, independente de grupo.

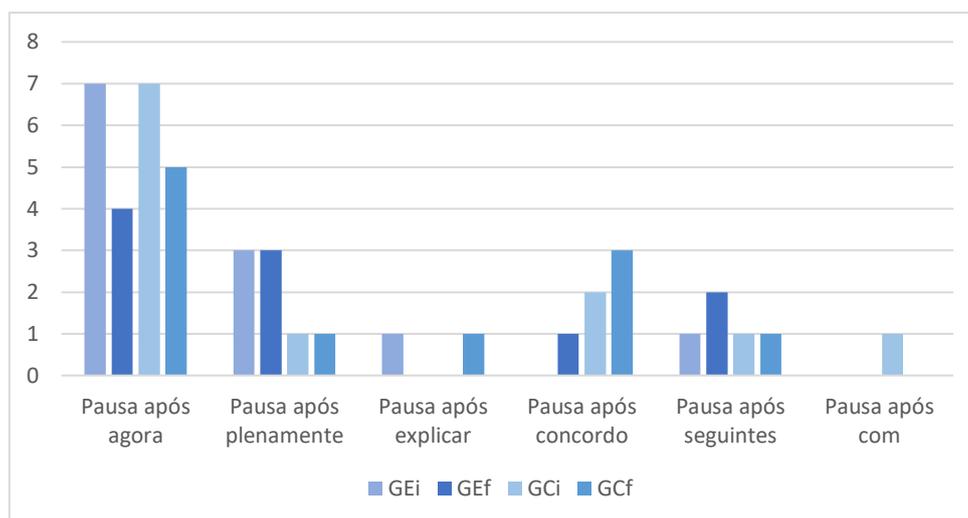


Gráfico 9. Quantitativo de pausas utilizadas e sua localização em relação às palavras do enunciado, considerando os grupos experimental (GE) e controle (GC) nas condições inicial (i) e final (f).

6.5 Correlações

Foram realizadas correlações entre os parâmetros analisados e a média das notas obtidas na análise perceptivo-auditiva (APA) por cada sujeito, independente da alocação em grupos. O teste de correlação de Pearson foi utilizado para os dados com distribuição normal (**Tabela 7**) e o teste de Spearman foi realizado para os dados que não possuíam distribuição normal (**Tabela 8**).

Tabela 7. Resultados das correlações de Pearson.

	Taxa de artic. palavras	Média tôn. enun. (Hz)	Sílabas por segundo	Palavras por minuto
Nota média na APA	,401	,335	,322	,399

Tabela 8. Resultados das correlações de Spearman.

	% Descont. fala (DC+DTG)	%DTG	nº pausas	Duração pausa 1 (s)	Duração pausa 2 (s)
Nota média na APA	-,268	-,074	-,255	-,256	-,194

6.6 Maiores notas vs. menores notas

Para fins de análise da influência das variáveis analisadas sobre as notas obtidas na análise perceptivo-auditiva (APA), os participantes foram alocados em dois grupos: maiores notas na APA e menores notas na APA. Cada grupo foi composto por quinze participantes, independente de alocação inicial em GE ou GC e condição inicial ou final, apenas a nota da APA foi levada em consideração para a formação desses grupos.

Os **Gráficos 10, 11 e 12** apresentam os resultados relativos às pausas para os grupos com as menores e maiores notas na APA. O grupo que obteve as maiores notas apresentou menor número de pausas ($n = 3$) em relação ao grupo com as menores notas ($n = 11$) (**Gráfico 10**). Nenhum indivíduo do grupo com as maiores notas emitiu mais de uma pausa em seu enunciado.



Gráfico 10. Quantitativo de pausas utilizadas nos enunciados dos quinze indivíduos que obtiveram as maiores e menores notas na APA.



Gráfico 11. Duração, sem segundos, da primeira pausa dos enunciados dos quinze indivíduos que obtiveram as maiores e menores notas na APA.



Gráfico 12. Duração, sem segundos, da segunda pausa dos enunciados dos quinze indivíduos que obtiveram as maiores e menores notas na APA.

O **Gráfico 13** apresenta os resultados referentes aos tipos de contornos melódicos produzidos pelos grupos que obtiveram as quinze maiores notas na análise perceptivo-auditiva e as quinze menores notas na análise perceptivo-auditiva.

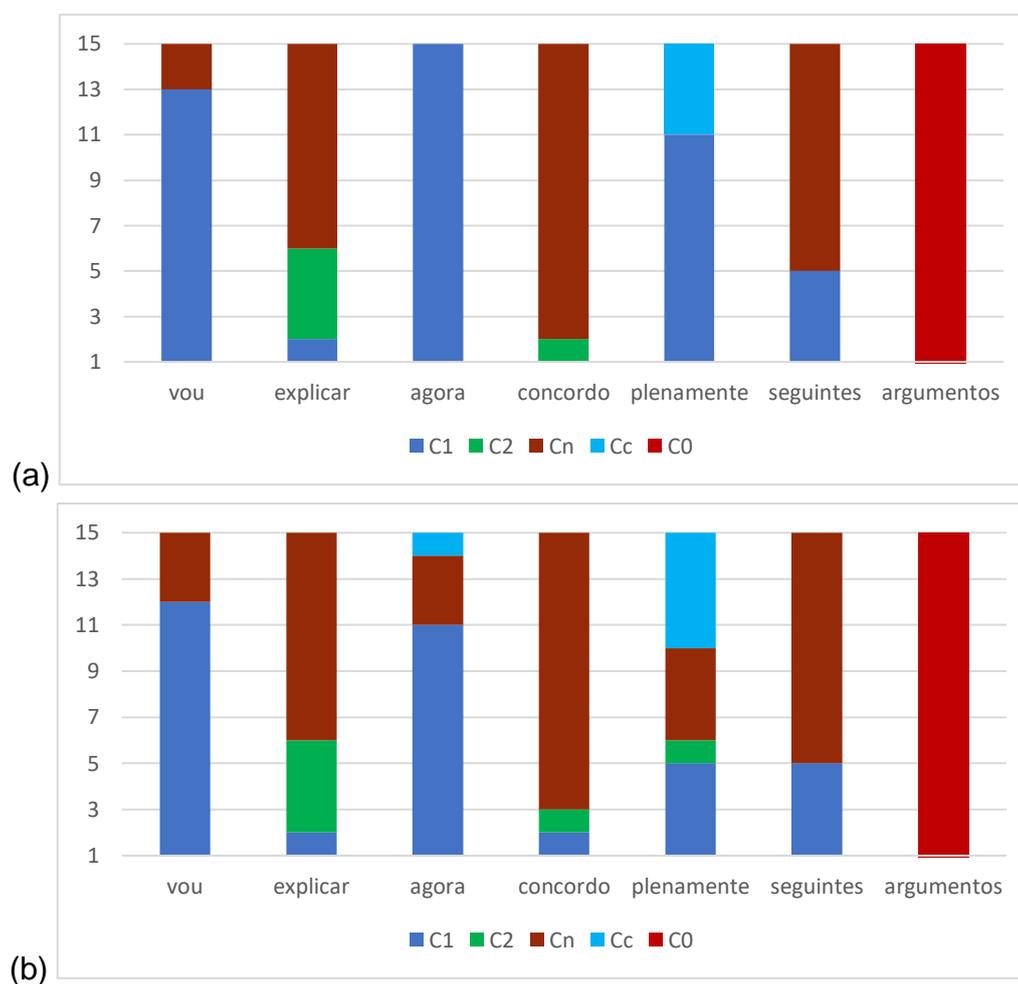


Gráfico 13. Tipos de contornos melódicos apresentados pelos grupos que obtiveram as maiores (a) e menores (b) notas na análise perceptivo-auditiva.

A comparação entre os quinze indivíduos que obtiveram as maiores notas na APA e os quinze indivíduos que obtiveram as menores notas na APA realizada com base nas médias das variáveis analisadas pode ser visualizada na **Tabela 9**.

Tabela 9. Média dos valores das variáveis dos indivíduos que obtiveram as maiores notas e menores notas na APA.

Variável	Média das maiores notas	Média das menores notas
Duração total	4,0286	4,6454
nº pausas	0,2	0,7333
Duração pausa 1 (s)	0,0694	0,1854
Duração pausa 2 (s)	0	0,0362
nº sílabas	22,1333	23,0666
Sílabas por segundo	5,5573	5,0486
Sílabas por minuto	333,4412	302,9206
nº palavras	9,1333	9,1333
Palavras por minuto	137,4162	120,0423
Taxa de articulação de sílabas	333,3718	302,6989
Taxa de articulação de palavras	137,3468	119,8206
% Descontinuidade de fala	1,2554	4,6417
%DTG	0	0,2898
Nota média na APA	3,6695	1,3942
Diferença em Hz - vou	21,2666	15,6
Diferença em Hz - explicar	15,6666	8,3333
Diferença em Hz - agora	22,2	15,7333
Diferença em Hz - concordo	6,9333	5,8666
Diferença em Hz - plenamente	14,6333	9,0333
Diferença em Hz - seguintes	9,7333	9,3333
Diferença em Hz - argumentos	29,6	23
Média tônica do enunciado (Hz)	17,1476	12,4142

O teste t independente apontou diferença entre as médias dos grupos com as maiores e menores notas na análise perceptivo-auditiva para a maioria das variáveis analisadas (**Tabela 10**).

Tabela 10. Teste t de amostras independentes entre os grupos com as maiores e menores notas na análise perceptivo-auditiva.

		Teste de Levene para igualdade de variâncias		teste t para Igualdade de Médias						
		Z	Sig.	t	df	Sig. (2 extremidades)	Diferença média	Erro padrão de diferença	95% Intervalo de Confiança da Diferença	
									Inferior	Superior
Duração total	Variâncias iguais não assumidas			-2,928	24,348	0,007	-0,610733	0,208585	-1,040906	-0,180561
Sílabas por minuto	Variâncias iguais assumidas	0,100	0,754	2,051	28	0,050	29,915539	14,586210	0,037041	59,794036
Sílabas por segundo	Variâncias iguais assumidas	0,100	0,754	2,051	28	0,050	0,498592	0,243104	0,000617	0,996567
Palavras por minuto	Variâncias iguais assumidas	0,385	0,540	3,154	28	0,004	17,114627	5,426649	5,998639	28,230614
Taxa artic. sílabas	Variâncias iguais assumidas	0,100	0,755	2,056	28	0,049	30,067805	14,627861	0,103990	60,031621
Taxa artic. palavras	Variâncias iguais assumidas	0,373	0,546	3,164	28	0,004	17,266893	5,457987	6,086714	28,447072
% Descont. fala	Variâncias iguais assumidas	0,349	0,560	-3,480	28	0,002	-3,386362	0,973197	-5,379866	-1,392858
Nota média na APA	Variâncias iguais assumidas	0,043	0,837	32,377	28	0,000	2,269565	0,070098	2,125976	2,413155
Média tôn. enun. [Hz]	Variâncias iguais assumidas	0,111	0,741	2,790	28	0,009	5,238095	1,877454	1,392305	9,083885
Diff [Hz] - vou	Variâncias iguais assumidas	0,029	0,865	2,072	28	0,048	6,733	3,249	0,078	13,389
Diff [Hz] - explicar	Variâncias iguais assumidas	0,097	0,758	2,438	28	0,021	8,200	3,363	1,311	15,089
Diff [Hz] - agora	Variâncias iguais assumidas	0,066	0,800	2,880	28	0,008	7,3333	2,5466	2,1169	12,5498
Diff [Hz] - concordo	Variâncias iguais assumidas	0,079	0,781	0,847	28	0,404	1,667	1,968	-2,365	5,698
Diff [Hz] - plenamente	Variâncias iguais assumidas	1,538	0,225	1,980	28	0,058	5,2667	2,6602	-0,1826	10,7159
Diff [Hz] - seguintes	Variâncias iguais assumidas	0,007	0,935	0,218	28	0,829	0,867	3,978	-7,283	9,016
Diff [Hz] - argumentos	Variâncias iguais não assumidas			1,380	23,346	0,181	6,600	4,784	-3,288	16,488

Para as variáveis que não apresentaram distribuição normal, foi realizado o teste Mann-Whitney (**Tabela 11**). O teste Mann-Whitney mostrou que a nota na análise perceptivo-auditiva tem efeito sobre as variáveis número de pausas ($U = 64,500$; $p < 0,05$) e duração da pausa 1 ($U = 70,000$; $p < 0,05$), mas não na duração da pausa 2 ($U = 97,500$; $p > 0,05$) e na porcentagem de DTG ($U = 105,000$; $p > 0,05$).

Tabela 11. Teste Mann-Whitney entre os grupos com as maiores e menores notas na análise perceptivo-auditiva.

	nº pausas	Dur. pausa 1 (s)	Dur. pausa 2 (s)	%DTG
U de Mann-Whitney	64,500	70,000	97,500	105,000
Wilcoxon W	184,500	190,000	217,500	225,000
Z	-2,303	-1,990	-1,438	-1,000
Significância Sig. (bilateral)	,021	,047	,150	,317
Sig exata [2*(Sig. de 1 extremidade)]	,045 ^b	,081 ^b	,539 ^b	,775 ^b

a. Variável de Agrupamento: Grupo

b. Não corrigido para vínculos.

6.7 Outros achados

Dentre os demais resultados encontrados na amostra, houve duas ocorrências nas quais os sujeitos acrescentaram uma palavra ao enunciado: ‘vou **me** explicar’ e ‘**eu** vou explicar’, ambas ocorreram no grupo experimental na condição inicial.

Um sujeito do grupo controle, na condição inicial, emitiu a parte final do enunciado no singular: ‘seguinte argumento’. Outro sujeito alocado no grupo controle apresentou rotacismo em suas emissões em ambas as condições (/isprɪkɐh/ e /prenemêtʃɪ/), um de seus enunciados compôs o grupo dos enunciados que obtiveram as menores notas na análise perceptivo-auditiva.

Houve duas ocorrências de acréscimo do fonema /h/ na palavra ‘seguintes’, por sujeitos diferentes: /sehɡĩtʃɪs/. As condições de ocorrência foram: grupo experimental na condição final e grupo controle na condição inicial. Um sujeito do grupo controle, na condição final, iniciou a emissão da palavra ‘seguintes’ com o mesmo acréscimo, porém interrompeu sua produção e corrigiu a emissão (ser- _ seguintes).

Um enunciado apresentou alteração da posição do fonema /h/ na palavra ‘argumentos’: /agʊhmêtʊs/. O enunciado foi emitido por um sujeito do grupo experimental, na condição final. A mesma alteração de posição do fonema /h/ na

palavra 'argumentos' ocorreu em um enunciado emitido por um sujeito do grupo controle na condição final, porém foi seguida de correção.

Um sujeito do grupo controle na condição final emitiu seu enunciado de forma hipernasal, este enunciado compôs o grupo dos enunciados que obtiveram as menores notas na análise perceptivo-auditiva.

As coarticulações mais frequentemente observadas nos enunciados emitidos tanto pelo grupo controle quanto pelo grupo experimental foram as seguintes: /voɪsplɪkeragɔrɐ/ e /kũsegĩfɪzahgɔmẽtus/.

7. Discussão

O presente trabalho possui o objetivo geral de verificar a reorganização dos parâmetros temporais da fala de estudantes de nível superior após um treinamento de aprimoramento do desempenho comunicativo oral. O grupo submetido ao treinamento obteve as melhores notas na análise perceptivo-auditiva (APA) realizada por juízes leigos, se diferenciando estatisticamente do grupo controle. Em relação aos parâmetros da organização temporal analisados, as maiores notas na APA foram atribuídas aos enunciados com menor duração total, menor número de pausas, maiores valores de palavras emitidas por minuto e taxa de articulação de palavras, menor descontinuidade de fala e de disfluências típicas da gagueira e maior variação de F0.

Em relação às escolhas metodológicas para a avaliação de características da fala, a análise perceptivo-auditiva é considerada o padrão ouro(114,116). Quanto à duração da amostra de fala utilizada para a análise perceptivo-auditiva, o número médio de sílabas da frase analisada no presente trabalho (média de 22,46 sílabas de duração) está acima do número de sílabas estipuladas para as amostras de fala encadeada utilizadas para validação de instrumentos de avaliação de características vocais para falantes do português brasileiro (15 sílabas de duração)(167) e utilizadas em estudos de análise semelhante (enunciados de 5 a 11 sílabas)(118). O tempo necessário para o julgamento adequado das características individuais apresentadas no enunciado foi adequado na presente pesquisa, considerando que a média de duração dos enunciados foi de 4,31 segundos e há estudos de percepção auditiva que utilizam estímulos com duração abaixo de 1 segundo(168) e enunciados curtos(108,118,121,123,169–171). Em relação ao número de juízes da APA, o número de juízes está acima do encontrado em estudos recentes que utilizaram esse método(118,123–125), os quais possuem média de 8 juízes em suas análises perceptivo-auditivas.

A emissão padronizada da frase analisada favorece a manutenção das características suprasegmentais do enunciado e viabiliza a manifestação da intenção do falante, tendo em vista que o modelo da frase indica 'o que' deve ser dito, e não 'como' deve ser dito. Nessa situação, não é tão exigido do falante o esforço linguístico

de elaboração do enunciado, mas sim de seus componentes suprasegmentais, como os componentes temporais aqui analisados. Do ponto de vista do ouvinte, a padronização dos enunciados não altera a qualidade de seu julgamento, tendo em vista que a informação lexical não é necessária para o reconhecimento da voz(168,172,173). Seguindo a proposta do presente trabalho, o fato de não haver qualquer conhecimento prévio por parte dos juízes acerca do contexto de produção dos enunciados (e o fato de serem leigos), permitiu que os juízes se baseassem unicamente em pistas acústicas para exprimir seu julgamento do nível de certeza.

Considerando o reconhecimento da validade da análise perceptivo-auditiva aqui empregada, seus resultados indicaram que o grupo submetido ao treinamento obteve as melhores notas na APA, na análise descritiva (**Gráfico 1**) e se diferenciou do grupo controle, segundo estatística inferencial (**Tabela 3**). Os grupos controle e experimental, na condição inicial, por sua vez, não se diferenciaram em relação às notas da APA, segundo estatística inferencial (**Tabela 4**). Os resultados indicaram que há efeito de grupo sobre a nota média obtida na APA (**Tabela 5**). O treinamento aplicado, portanto, foi efetivo em seu objetivo de melhorar o nível de certeza na fala dos sujeitos.

Iniciando a discussão da organização temporal, o GEf apresentou maior taxa de articulação de palavras e palavras por minuto quando comparado ao GCf (**Tabela 6**). Tal diferença se manteve quando todos os grupos foram analisados, sendo que o GE na condição final apresentou a maior média de palavras emitidas por minuto (133,991), enquanto o GC na condição final apresentou a menor média de palavras emitidas por minuto (122,940). Os valores de palavras emitidas por minuto por todos os grupos foram acima do esperado para a faixa etária média de seus componentes(161). A proximidade dos valores de taxa de articulação de sílabas e SPM, assim como os valores de taxa de articulação de palavras e PPM, pode ser explicada pela presença e duração das pausas nos enunciados, as quais ocorreram com baixa frequência e duração. Sabe-se que valores elevados das medidas relacionadas à velocidade de fala estão relacionados a maiores avaliações no nível de certeza(36,108), fato corroborado pelo presente estudo.

Apenas dois enunciados apresentaram DTGs (prolongamento e bloqueio), ambas antes da palavra 'seguintes'. A baixa ocorrência de DTGs era esperada,

considerando que a amostra não foi composta por indivíduos com gagueira. Além deste fato, mesmo que algum indivíduo apresentasse gagueira, o tamanho do enunciado utilizado para análise, assim como o acesso prévio ao seu conteúdo linguístico por parte do falante, provavelmente reduziria as chances de ocorrência de DTGs na emissão(54,55). Quanto à ocorrência das duas DTGs antes da palavra 'seguintes', uma palavra de conteúdo, esperava-se que as disfluências, caso presentes, ocorressem em palavras funcionais(174), as quais exercem funções sintáticas e possuem baixa carga semântica(175,176). O fato de um dos enunciados com DTG ter sido julgado entre os piores da amostra está de acordo com achados anteriores em relação à percepção de disfluências típicas da gagueira e o julgamento do nível de certeza da fala. Os dois enunciados que apresentaram disfluências típicas da gagueira tiveram porcentagem de ocorrência de DTGs acima de 4%, valor o qual está acima do esperado para falantes fluentes(177–179). Não é possível inferir, porém, com base neste achado, que estes sujeitos tenham gagueira, pois há que se considerar que o número de sílabas emitidas não foi adequado para as recomendações de cálculo para o diagnóstico de gagueira(180).

Em relação às disfluências comuns, quatro enunciados apresentaram revisões. Tal ocorrência era esperada, uma vez que o fenômeno de retomar um enunciado (e não o deixar por terminar) após uma auto interrupção disfluente parece ser uma característica frequente dos enunciados orais(181). O fato de que, desconsiderando as pausas enquanto disfluências, o GEf teve a maior média de descontinuidade de fala, pode ser explicado pelo fato de os falantes estarem mais vigilantes em relação à sua fala e, por consequência, apresentarem mais disfluências comuns como forma de aumentar o tempo de emissão do enunciado (prolongamento típico), o que aumentaria seu tempo de planejamento e diminuiria seus erros, e, também, aumentaria seu tempo para a correção de eventuais mensagens inadequadas (revisão).

De forma geral, a presença de disfluências nos enunciados era esperada, considerando que as disfluências são um fenômeno comum na fala espontânea e no discurso em palestra(182) e são representativas de características individuais dos falantes(181). Pode-se concluir, ainda, a partir desse resultado, que a presença de disfluências comuns não interferiu no resultado da análise perceptivo-auditiva de forma significativa ao ponto de causar efeito de grupo no GEf.

No que diz respeito à análise entoacional, a segmentação da curva de F0 com foco em cada uma das sílabas salientes do enunciado permitiu uma análise mais precisa dos focos de variação, a qual poderia ter sido mascarada pela análise da variação de F0 apresentada no enunciado como um todo.

Quando analisada individualmente entre os sujeitos, a forma de uso dos contornos melódicos apresentou grande variação: alguns indivíduos produziram contornos planos em quase todo o enunciado enquanto outros indivíduos empregaram o uso de contornos planos, ascendentes e complexos em suas emissões (**Figura 6** e **Figura 7**). Quando analisados dentro dos grupos GE e GC, porém, os contornos melódicos não apresentaram tendências relacionadas ao treinamento ou ao julgamento na APA. Contudo, quando os sujeitos do GE que obtiveram melhora significativa foram analisados dentro de um grupo próprio, a análise visual da **Figura 6** indica uma tendência para a substituição do contorno Cn pelo contorno C1. Esta afirmação pode ser feita, inclusive, na comparação entre o enunciado que obteve a maior nota e o enunciado que obteve a menor nota na APA (**Figura 7**). Tal substituição está apoiada no fato de que o contorno plano Cn não é perceptível pelo ouvido humano enquanto variação de F0, enquanto o contorno ascendente C1 o é (103,104). Essa substituição de contornos provocaria, então, aumento da percepção do ouvinte em relação à variação melódica do enunciado e, conseqüente, aumento do nível de certeza.

A análise individual das sílabas e vogais tônicas do enunciado revelou que o GEf se diferenciou dos demais grupos em relação à variação de F0 na palavra 'vou' (**Gráfico 4**), sendo a variação de F0 para esta palavra maior neste grupo. Houve, ainda, resultado estatístico inferencial da variação de F0 na palavra 'vou' entre os grupos com as maiores e menores notas (**Tabela 9**). A presença de resultado significativo em relação à variação de F0 na primeira palavra do enunciado é corroborada por estudos anteriores (171,183,184) com achados de alta sensação de confiança relacionados ao aumento da variação de F0 na posição inicial do enunciado (108).

A organização da estrutura prosódica dos enunciados (**Figura 6**), de forma geral, parece estar ligada à individualidade dos falantes. Não houve reorganização de estrutura prosódica sistemática o suficiente entre os sujeitos de forma que um modelo

pudesse ser elaborado para explicar tal reorganização. Ressalta-se que os efeitos das variáveis independentes sobre os parâmetros analisados são ditados não apenas por aspectos linguísticos, mas também por aspectos interpessoais(171), os quais contribuíram para a variedade de estruturas prosódicas apresentadas (n = 36).

Quanto à segmentação dos grupos acentuais, houve reemissão de todo o grupo acentual em caso de erro e conseqüente interrupção do GA. As pausas também não apresentaram tendência para a segmentação do grupo acentual. As pausas, de forma geral, ocorreram na fronteira de GA e foram silenciosas. A maior ocorrência de pausas nas fronteiras do enunciado está de acordo com o padrão esperado de emissões realizadas por falantes fluentes(73,164,185).

O menor número absoluto de pausas emitidas por grupo foi produzido pelo GE na condição final (n=9), enquanto o maior número de pausas foi produzido pelo GC na condição final (n=13) (**Gráfico 6**). Tal achado pode sugerir que maiores avaliações auditivas do nível de certeza da fala podem estar relacionadas a menores números de pausas por emissão.

Em relação às hesitações, houve apenas duas ocorrências, ambas antes da emissão do enunciado, sendo que nenhuma delas foi produzida pelo grupo submetido ao treinamento de desempenho comunicativo oral. A ocorrência de hesitações que precederam o enunciado proposto pode ser associada ao fato de que os falantes estavam elaborando o conteúdo que pretendiam emitir após a frase padronizada, considerando que as hesitações são elementos da fala produzidos de forma consciente(186) que tendem a acontecer durante buscas de memória, antes de uma resposta ser dita(129).

Considerando que os ouvintes são sensíveis à localização das pausas(73), houve maior ocorrência de pausas após a palavra 'agora', seguida da palavra 'plenamente' (ambas palavras de conteúdo), ambas seguidas de palavras funcionais ('porque' e 'com', respectivamente). Este achado não está de acordo com estudos anteriores que relatam maior frequência de pausas antes de palavras de conteúdo(164,187). Tal localização das pausas, quando consideradas as características linguísticas das palavras, pode ser explicada por demais fatores envolvidos na emissão do falante, como a intencionalidade de transmitir certeza. Uma

explicação possível para tal sugestão seria o fato de essa localização das pausas segmentar a emissão da frase padronizada em enunciados afirmativos.

A duração das pausas respiratórias foi, em média, maior do que a duração das pausas silenciosas. Este achado está em conformidade com estudo prévio que indica que as pausas respiratórias são mais longas quando comparadas às pausas não respiratórias(68) e podem exercer função de planejamento(171).

A duração média das pausas produzidas em geral e em relação aos grupos foi em sua maioria acima de 0,250 s, a única exceção foi o grupo controle na condição inicial (0,201 s). A duração média das pausas está de acordo com os limiares perceptivos dos ouvintes, pois as pausas são bem discriminadas por ouvintes em durações acima de 200 ms(188). A duração apresentada pode ser resultado do tempo de planejamento por parte do falante, já que as pausas silenciosas maiores do que 250 ms podem refletir o tempo necessário para o planejamento linguístico (hesitação), enquanto pausas menores do que 250 ms podem ser consideradas mais relacionadas com a articulação do que com o funcionamento cognitivo durante a produção da fala(189).

Ressalta-se, sobre a análise das pausas, que estas não podem ser extraídas de seu contexto de fala e estudadas isoladamente sem o descarte de informações úteis(188) e, portanto, o ambiente de ocorrência das pausas e sua relação com os demais aspectos envolvidos no processo da fala foram considerados, respeitando a proposta de análise do presente trabalho.

No que diz respeito às correlações entre as variáveis temporais e as notas na análise perceptivo-auditiva, nenhuma correlação foi mais forte do que .4 (**Tabela 7 e Tabela 8**). As correlações positivas ocorreram para a taxa de articulação de palavras, média tônica do enunciado, SPS e PPM. Já as correlações negativas ocorreram para a descontinuidade de fala, %DTG, pausas e duração da primeira e segunda pausa. A correlação mais forte encontrada foi entre a taxa de articulação de palavras e a nota média obtida na APA. Apesar disso, os dados apontam tendências de correlação entre os julgamentos na APA e as variáveis analisadas, conforme pode ser observado nos resultados encontrados entre os grupos com as maiores e menores notas na APA, os quais serão descritos a seguir.

Na avaliação dos enunciados que obtiveram as maiores notas e as menores notas na APA, as maiores diferenças observadas foram as relacionadas às variáveis de duração (**Tabela 9** e **Tabela 10**). Os enunciados que obtiveram as maiores notas na APA foram mais curtos em sua duração total, possuíam menos pausas, suas pausas foram mais curtas, apresentaram maior número de SPM e PPM, menor descontinuidade de fala, menor %DTG, maior variação de F0 nas palavras 'vou', 'explicar' e 'agora' e maior média de variação de F0 para as sílabas salientes do enunciado como um todo.

Observou-se que quanto maiores as notas, menos pausas foram emitidas e menor a sua duração. Nenhum indivíduo do grupo com as maiores notas emitiu mais de uma pausa em seu enunciado. Dois dos cinco enunciados que obtiveram as maiores notas na análise perceptivo-auditiva apresentaram apenas uma pausa. A presença de pausas nos melhores enunciados da amostra não é inesperada, pois as pausas auxiliam na compreensão auditiva do enunciado(109), fator este que pode ter compelido os ouvintes a julgarem os enunciados de forma positiva, em consideração ao uso adequado das pausas para a facilitação de seu entendimento.

O teste Mann-Whitney (**Tabela 11**) mostrou que a nota na análise perceptivo-auditiva tem efeito sobre as variáveis número de pausas e duração da primeira pausa, mas não na duração da segunda pausa e na porcentagem de DTG. Considerando que todas as variáveis avaliadas por este teste apresentaram menor média no grupo que obteve as maiores notas na APA (**Tabela 9**), é possível concluir que a relação entre as variáveis número de pausas e duração da primeira pausa é inversamente proporcional às notas na APA.

Além dos achados referentes aos parâmetros propostos para análise na pesquisa, demais achados surgiram na amostra e foram relatados. A presença de rotacismo por parte de um sujeito da pesquisa provavelmente pode ser explicada por questões sociolinguísticas individuais do falante e de seu contexto, assim como a inserção e alteração do fonema /h/ nas palavras 'seguintes' e 'argumentos'.

As coarticulações mais frequentemente observadas na amostra foram as seguintes: /voɪsplɪkɐragɔrɐ/ e /kũsegĩtʃɪzahgumẽtus/. A coarticulação das palavras 'vou' e 'explicar' resultou na alteração das vogais /u/ e /e/, os quais foram substituídas pela vogal /ɪ/, provavelmente devido ao efeito do arredondamento dos lábios durante

a produção da vogal /o/, alterando a produção da vogal /e/ subsequente. Na coarticulação das palavras ‘explicar’ e ‘agora’, ocorreu a alteração do último fonema da palavra ‘explicar’ (o qual seria produzido conforme dialeto local, de forma isolada, como /h/), sendo sua produção substituída pelo fonema /r/. A coarticulação das palavras ‘com’, ‘os’ e ‘seguintes’ foi realizada de tal forma que a palavra ‘os’ foi omitida e incorporada entre a produção nasal do fonema /ũ/ e o primeiro fonema (/s/) da palavra seguinte. Na coarticulação entre as palavras ‘seguintes’ e ‘argumentos’, pode-se observar que o fonema seguinte (/a/) levou seu traço de vozeamento ao fonema anterior (o fonema /s/ foi produzido como /z/). A presença de coarticulação é esperada e explicada pelo fato de o contexto fonético na fala espontânea ser variado, sendo a coarticulação mais intensa nesta tarefa de fala(16), além de influenciada por características individuais do falante.

Em relação às limitações do estudo, por tratar-se de um *corpus* auditivo, perde-se as características visuais que contribuem para o julgamento da sensação de certeza e incerteza. Quanto à análise das disfluências de fala, ressalta-se que as amostras de fala não possuíram o componente visual. Sabe-se que a análise em vídeo é essencial para a análise da fluência de fala, principalmente para os eventos típicos da gagueira. Por outro lado, a adição de aspectos visuais à amostra implicaria na possível interferência destes elementos nos julgamentos perceptivos, desviando o propósito de análise temporal da fala aqui proposta e realizada. A qualidade das gravações pode ter sido afetada por questões individuais do falante, como a intensidade vocal e a amplitude articulatória. Sabe-se que uma articulação ampla e precisa aumenta a intensidade da fala(190–192). A influência da intensidade, porém, não foi avaliada no presente estudo. Apesar dos cuidados tomados em relação ao ambiente das gravações, alguns áudios da amostra apresentaram ruídos externos eventuais e eco. A presença desses fatores eventuais, porém, não pareceu alterar os julgamentos na análise perceptivo-auditiva.

Dentre as direções futuras sugeridas, ressalta-se que um estudo da influência de fatores sociolinguísticos no julgamento dos enunciados pode ser benéfico para a compreensão dos julgamentos do nível de certeza. Além disso, um estudo de análise dos componentes extralinguísticos da comunicação também poderia trazer benefícios para o entendimento da percepção de certeza.

8. Conclusão

No presente trabalho, a análise da organização temporal da fala dos sujeitos submetidos ao treinamento de desempenho comunicativo oral apontou que houve reorganização em relação à taxa de articulação de palavras e ao número de palavras emitidas por minuto.

O treinamento de desempenho comunicativo oral aumentou o nível de certeza na fala dos sujeitos submetidos a ele, conforme resultados da avaliação perceptivo-auditiva realizada por juízes leigos.

O julgamento dos ouvintes leigos indicou que uma organização temporal na qual o enunciado é emitido com menor duração, com menos pausas, com pausas mais curtas, com maior número de sílabas emitidas por segundo e por minuto, maior número de palavras emitidas por minuto, maior taxa de articulação de palavras, menor descontinuidade de fala, ausência de disfluências típicas da gagueira e maior variação de F0 nas sílabas salientes do enunciado transmite maior sensação de certeza.

Os resultados apresentados possuem implicações clínicas para a prática fonoaudiológica na medida em que fornecem conhecimentos baseados em evidências a respeito dos parâmetros temporais da fala necessários para o treinamento do desempenho comunicativo oral visando o aumento da sensação de certeza na fala.

Referências

1. De Saussure F. Curso de Linguística Geral. 28th ed. São Paulo: Editora Cultrix; 2012. 58 p.
2. Laver J. Towards an integrated theory of nonverbal communication. In: Proceedings of the 14th International Congress of Phonetic Sciences - ICPHS99 [Internet]. San Francisco. Berkeley: University of California; 1999. p. 2433–6. Available from: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14_2433.pdf
3. Menegon P, Madureira S. Metáforas no Ensino de Canto e seus Efeitos na Qualidade Vocal: um Estudo Acústico-Perceptivo. In: Sonoridades: a expressividade da fala, no canto e na declamação. 2016. p. 62–89.
4. Laver J. The phonetic description of voice quality. Cambridge: Cambridge University Press; 1980.
5. Laver J. Principles of Phonetics. Cambridge: Cambridge University Press; 1994.
6. Celeste LC, Lima-Gregio AM. Aspectos teóricos para elaboração de programa de treinamento da performance comunicativa. In: Sonoridades: a expressividade da fala, no canto e na declamação. 2016. p. 111–8.
7. Fontes MA de S. Os Papéis das Prosódias Vocal e Visual na Expressão de Emoções na Fala. In: Sonoridades: a expressividade da fala, no canto e na declamação. 2016. p. 99–110.
8. Freud D, Ezrati-Vinacour R, Amir O. Speech rate adjustment of adults during conversation. J Fluency Disord [Internet]. 2018;57:1–10. Available from: <https://doi.org/10.1016/j.jfludis.2018.06.002>
9. Hazlett DE, Duffy OM, Moorhead SA. Review of the impact of voice training on the vocal quality of professional voice users: Implications for vocal health and recommendations for further research. J Voice [Internet]. 2011;25(2):181–91. Available from: <http://dx.doi.org/10.1016/j.jvoice.2009.08.005>

10. Borrego MC de M, Behlau M. Mapeamento do eixo condutor da prática fonoaudiológica em expressividade verbal no trabalho de competência comunicativa. *CoDAS*. 2018;30(6):e20180054.
11. Rodero E, Diaz-Rodriguez C, Larrea O. A Training Model for Improving Journalists' Voice. *J Voice* [Internet]. 2018;32(3):386.e11-386.e19. Available from: <http://dx.doi.org/10.1016/j.jvoice.2017.05.006>
12. Celeste LC, Lima AM, Seixas JMA, Silva MA da, Silva EM da. Treinamento da performance comunicativa em universitários da área da saúde TT - Communicative performance training in university health students. *Audiol, Commun res* [Internet]. 2018;23:e1879–e1879. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2317-64312018000100303
13. Cagliari LC. *Elementos de Fonética do Português Brasileiro*. Universidade Estadual de Campinas; 1981.
14. Zellner B. Pauses and the temporal structure of speech. *Fundam speech Synth speech Recognit*. 1994;41–62.
15. Barbosa PA. “Syllable-timing in Brazilian Portuguese”: uma crítica a Roy Major. *DELTA Doc Estud em Lingüística Teórica e Apl* [Internet]. 2000;16(2):369–402. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502000000200006
16. Daniloff RG, Hammarberg RE. On defining coarticulation. *J Phon* [Internet]. 1973;1(3):239–48. Available from: [https://doi.org/10.1016/S0095-4470\(19\)31388-9](https://doi.org/10.1016/S0095-4470(19)31388-9)
17. Kent RD, Read C. *Análise acústica da fala*. São Paulo: Cortez Editora; 2015. 545 p.
18. Browman CP, Goldstein L. Articulatory Phonology: An Overview. *Phonetica*. 1992;49(3–4):155–80.
19. Barbosa PA, Madureira S. *Manual de fonética acústica experimental: aplicações a dados do português*. São Paulo: Cortez Editora; 2015. 591 p.
20. Poeppel D, Assaneo MF. Speech rhythms and their neural foundations. *Nat*

- Rev Neurosci [Internet]. 2020;21(6):322–34. Available from:
<http://dx.doi.org/10.1038/s41583-020-0304-4>
21. Abrams DA, Bhatara A, Ryali S, Balaban E, Levitin DJ, Menon V. Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cereb Cortex*. 2011;21(7):1507–18.
 22. Boyes Braem P. Rhythmic temporal patterns in the signing of deaf early and late learners of Swiss German sign language. *Lang Speech*. 1999;42(2–3):177–208.
 23. Ghitza O. On the role of theta-driven syllabic parsing in decoding speech: intelligibility of speech with a manipulated modulation spectrum. *Front Psychol*. 2012;3:238.
 24. Shannon R V., Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. *Science (80-)*. 1995;270(5234):303–4.
 25. Doelling KB, Arnal LH, Ghitza O, Poeppel D. Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage*. 2014;85:761–8.
 26. Poeppel D, Idsardi WJ, Van Wassenhove V. Speech perception at the interface of neurobiology and linguistics. *Philos Trans R Soc B Biol Sci*. 2008;363(1493):1071–86.
 27. McMurray B, Kovack-Lesh KA, Goodwin D, McEchron W. Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition*. 2013;129(2):362–78.
 28. Falk S, Kello CT. Hierarchical organization in the temporal structure of infant-direct speech and song. *Cognition [Internet]*. 2017;163:80–6. Available from:
<http://dx.doi.org/10.1016/j.cognition.2017.02.017>
 29. Elliott TM, Theunissen FE. The modulation transfer function for speech intelligibility. *PLoS Comput Biol*. 2009;5(3).
 30. Luque J, Luque B, Lacasa L. Scaling and universality in the human voice. *J R Soc Interface*. 2015;12(105):1–6.

31. Torre IG, Luque B, Lacasa L, Kello CT, Hernández-Fernández A. On the physical origin of linguistic laws and lognormality in speech. *R Soc open Sci*. 2019;6(8).
32. Qin Y, Carlini N, Cottrell G, Goodfellow I, Raffel C. Imperceptible, robust, and targeted adversarial examples for automatic speech recognition. In: *Proceedings of the 36th International Conference on Machine Learning*, PMLR. 2019. p. 5231–40.
33. Henderson A, Goldman-Eisler F, Skarbek A. Sequential Temporal Patterns in Spontaneous Speech. *Lang Speech*. 1966;9(4):207–16.
34. Moniz H, Batista F, Mata AI, Trancoso I. Speaking style effects in the production of disfluencies. *Speech Commun [Internet]*. 2014;65(May):20–35. Available from: <http://dx.doi.org/10.1016/j.specom.2014.05.004>
35. Schachter S, Christenfeld N, Ravina B, Bilous F. Speech disfluency and the structure of knowledge. *J Pers Soc Psychol*. 1991;60(3):362–7.
36. Brennan SE, Williams M. The Feeling of Another's Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers. *J Mem Lang*. 1995;34(3):383–98.
37. Yairi E. Disfluencies of normally speaking two-year-old children. *J Speech Hear Res*. 1982;24(December 1981):490–5.
38. Riper C Van. *The nature of stuttering*. Prentice Hall; 1982.
39. Oliveira CMC de, Bernardes APL, Broglio GAF, Capellini SA. Perfil da fluência de indivíduos com taquifemia. *Pró-Fono Rev Atualização Científica*. 2010;22(4):445–50.
40. Ambrose NG, Yairi E. Normative Disfluency Data for Early Childhood Stuttering. *J Speech, Lang Hear Res*. 1999;42:895–909.
41. Yaruss JS, LaSalle LR, Conture EG. Evaluating Stuttering in Young Children: Diagnostic Data. *Am J Speech-Language Pathol*. 1998;7(4):62–76.
42. Ambrose NG, Yairi E. The Role of Repetition Units in the Differential Diagnosis of Early Childhood Incipient Stuttering. *Am J Speech-Language Pathol*.

- 1995;4(3):82–8.
43. Yairi E, Lewis B. Disfluencies at the onset of stuttering. *J Speech Hear Res.* 1984;27:154–9.
 44. Cordes AK, Ingham RJ. Stuttering includes both within-word and between-word disfluencies. *J Speech Hear Res.* 1995;38(2):382–6.
 45. Blacfkmer ER, Mitton JL. Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition.* 1991;39(3):173–94.
 46. Bortfeld H, Leon SD, Bloom JE, Schober MF, Brennan SE. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Lang Speech.* 2001;44(2):123–47.
 47. Maclay H, Osgood CE. Hesitation Phenomena in Spontaneous English Speech. *WORD.* 1959;15(1):19–44.
 48. Finlayson IR, Corley M. Disfluency in dialogue: An intentional signal from the speaker? *Psychon Bull Rev.* 2012;19(5):921–8.
 49. Clark HH, Wasow T. Repeating Words in Spontaneous Speech. *Cogn Psychol.* 1998;37(3):201–42.
 50. Beattie GW, Butterworth BL. Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. *Lang Speech.* 1979;22(3):201–11.
 51. Levelt WJ. Monitoring and self-repair in speech. *Cognition.* 1983;14(1):41–104.
 52. Schnadt MJ, Corley M. The influence of lexical, conceptual and planning based factors on disfluency production. 28th Annu Conf ... [Internet]. 2006;8–13. Available from: <http://csjarchive.cogsci.rpi.edu/Proceedings/2006/docs/p750.pdf>
 53. Merlo S, Mansur LL. Descriptive discourse: topic familiarity and disfluencies. *J Commun Disord.* 2004;37(6):489–503.
 54. Shriberg E. Disfluencies in switchboard. In: *Proceedings of International Conference on Spoken Language Processing.* Philadelphia, PA; 1996. p. 11–4.
 55. Oviatt S. Predicting spoken disfluencies during human-computer interaction.

- Comput Speech Lang. 1995;9(1):19–35.
56. Tree JEF. The Effects of False Starts and Repetitions on the Processing of Subsequent Words in Spontaneous Speech. Vol. 34, *Journal of Memory and Language*. 1995. p. 709–38.
 57. Deng H, Lin Y, Utsuro T, Kobayashi A, Nishizaki H, Hoshino J. Automatic Fluency Evaluation of Spontaneous Speech Using Disfluency-Based Features. In: *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing*. 2020. p. 9239–43.
 58. Fox Tree JE. Interpreting Pauses and Ums at Turn Exchanges. *Discourse Process*. 2002;34(1):37–55.
 59. Andrade CRF de. Protocolo para avaliação da fluência da fala. *Pro-Fono*. 2000;12(2):131–4.
 60. Taschenberger L, Tuomainen O, Hazan V. Disfluencies in spontaneous speech in easy and adverse communicative situations: the effect of age. 2019;(July):55–8.
 61. Themistocleous C, Eckerström M, Kokkinakis D. Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. *PLoS One*. 2020;15(7).
 62. Yan X, Kim HR, Kim JY. Dimensionality of speech fluency: Examining the relationships among complexity, accuracy, and fluency (CAF) features of speaking performances on the Aptis test. *Lang Test*. 2020;
 63. Trouvain J, Werner R, Möbius B. An acoustic analysis of inbreath noises in read and spontaneous speech. In: *Proceedings of the International Conference on Speech Prosody*. 2020. p. 789–93.
 64. Goldman-Eisler F. The predictability of words in context and the length of pauses in speech. *Lang Speech*. 1958;1(3):226–31.
 65. Navarretta C. The Functions of Fillers , Filled Pauses and Co-Occurring Gestures in Danish Dyadic Conversations. In: *3rd European Symposium on Multimodal Communication*. 2015. p. 55–61.

66. Reynolds A, Paivio A. Cognitive and emotional determinants of speech. *Can J Psychol Can Psychol*. 1968;22(3):164–175.
67. Himmelmann NP, Sandler M, Strunk J, Unterladstetter V. On the robustness of intonational phrases in spontaneous speech: A crosslinguistic interrater study. *Phonology*. 2018;35:207–45.
68. Trouvain J, Fauth C, Möbius B. Breath and non-breath pauses in fluent and disfluent phases of German and French L1 and L2 read speech. *Proc Int Conf Speech Prosody*. 2016;31–5.
69. Gyarmathy D, Horváth V. Pausing strategies with regard to speech style. In: *The 9th Workshop on Disfluency in Spontaneous Speech*. 2019. p. 27–30.
70. Duez D. Perception of silent pauses in continuous speech. *Lang Speech*. 1985;28(4):377–89.
71. Goldman-Eisler F. Pauses, Clauses, Sentences. *Lang Speech*. 1972;15(2):103–13.
72. Viola IC, Madureira S. The roles of pause in speech expression. *Proc 4th Int Conf Speech Prosody, SP 2008*. 2008;721–5.
73. Kahng J. The effect of pause location on perceived fluency. *Appl Psycholinguist*. 2018;39(3):569–91.
74. Krivokapić J. Prosodic planning: Effects of phrasal length and complexity on pause duration. *J Phon*. 2007;35(2):162–79.
75. Jeung-Yoon Choi. Pause length and speech rate as durational cues for prosody markers. *J Acoust Soc Am*. 2003;114(4):2395–2395.
76. de Jong NH, Bosker HR. Choosing a threshold for silent pauses to measure second language fluency. In: *DiSS 2013 Proceedings of the 6th Workshop on Disfluency in Spontaneous Speech* [Internet]. 2013. p. 17–20. Available from: http://www.isca-speech.org/archive/diss_2013/dis6_017.html
77. Hieke AE, Kowal S, O'Connell DC. The trouble with “articulatory” pauses. *Lang Speech*. 1983;26(3):203–14.
78. Campione E, Véronis J. A large-scale multilingual study of pause duration. In:

- Speech Prosody 2002 Proceedings of the 1st International Conference on Speech Prosody [Internet]. 2002. p. 199–202. Available from: http://www.isca-speech.org/archive/sp2002/sp02_199.html
79. Juste FS, Andrade CRF. Perfil da fluência da fala em diferentes tarefas para indivíduos com Doença de Parkinson. *CoDAS*. 2017;29(4).
 80. Brabo NC, Minett TSC, Ortiz KZ. Fluency in Parkinson's disease: disease duration, cognitive status and age. *Arq Neuropsiquiatr*. 2014;72(5):349–55.
 81. Moura RBO. Efeitos da indução da fala rápida e lenta na fluência de adultos com e sem gagueira. Universidade Estadual Paulista - UNESP - Câmpus de Marília; 2019.
 82. de Andrade CRF, de Oliveira Martins V. Variação da fluência da fala em idosos. *Pro-Fono*. 2010;22(1):13–8.
 83. Shriberg E, Stolcke A, Hakkani-Tur D, Tur G. Prosody-based automatic segmentation of speech into sentences and topics. *Speech Commun*. 2000;32(1–2):127–54.
 84. Morgan N, Fosler-Lussier E. Combining multiple estimators of speaking rate. In: *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98*. 1998. p. 729–32.
 85. Siegler M. Measuring and compensating for the effects of speech rate in large vocabulary continuous speech recognition. Carnegie Mellon University; 1995.
 86. Wang D, Narayanan SS. Robust Speech Rate Estimation for Spontaneous Speech. *IEEE Trans Audio Speech Lang Processing*. 2007;15(8):2190–201.
 87. Raneri D, Von Holzen K, Newman R, Bernstein Ratner N. Change in maternal speech rate to preverbal infants over the first two years of life. *J Child Lang*. 2020;47(6):1263–75.
 88. Polyanskaya L, Ordin M, Busa MG. Relative Saliency of Speech Rhythm and Speech Rate on Perceived Foreign Accent in a Second Language. *Lang Speech*. 2017;60(3):333–55.
 89. Versfeld NJ, Dreschler WA. The relationship between the intelligibility of time-

- compressed speech and speech in noise in young and elderly listeners. *J Acoust Soc Am*. 2002;111(1):401–8.
90. Logan KJ, Conture EG. Length, grammatical complexity, and rate differences in stuttered and fluent conversational utterances of children who stutter. *J Fluency Disord*. 1995;20(1):35–61.
 91. Brundage SB, Ratner NB. Measurement of stuttering frequency in children's speech. *J Fluency Disord*. 1989;14(5):351–8.
 92. Simantiraki O, Cooke M. Listeners' speech rate preferences in stationary and modulated maskers. In: *Proceedings of the 23rd International Congress on Acoustics*. Aachen, Germany; 2019. p. 5771–3.
 93. Müller JA, Wendt D, Kollmeier B, Debener S, Brand T. Effect of speech rate on neural tracking of speech. *Front Psychol*. 2019;10(MAR):1–15.
 94. Wingfield A, Ducharme JL. Effects of age and passage difficulty on listening-rate preferences for time-altered speech. *Journals Gerontol - Ser B Psychol Sci Soc Sci*. 1999;54(3):1967–70.
 95. Decruy L, Lesenfants D, Vanthornhout J, Francart T. Top-down modulation of neural envelope tracking: the interplay with behavioral, self-report and neural measures of listening effort [Internet]. *bioRxiv*. 2019. 815365 p. Available from: <http://biorxiv.org/content/early/2019/11/14/815365.abstract>
 96. Kenny OP, Nelson DJ, Bodenschatz JS, McMonagle HA. Separation of non-spontaneous and spontaneous speech. In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. 1998. p. 573–6.
 97. Laczkó M. Temporal characteristics of teenagers' spontaneous speech and topic based narratives produced during school lessons. In: *The 9th Workshop on Disfluency in Spontaneous Speech*. 2019. p. 63.
 98. O'Shaughnessy D. A study of French vowel and consonant durations. *J Phon* [Internet]. 1981;9(4):385–406. Available from: [https://doi.org/10.1016/S0095-4470\(19\)31016-2](https://doi.org/10.1016/S0095-4470(19)31016-2)
 99. Martin P. Intonation, structure prosodique et ondes cérébrales: introduction à

- l'analyse prosodique. 1st ed. ISTE Group; 2018.
100. Cagliari LC. Prosódia: algumas funções dos supra-segmentos. *Cad Estud Linguist.* 1992;23:137–51.
 101. Martin P. Regional Variations of Sentence Intonation in French-The Continuation Contour in Parisian French. In: *Speech Prosody 2002, International Conference.* 2002.
 102. Martin P. *The Structure of Spoken Language: Intonation in Romance.* Cambridge University Press; 2015.
 103. Martin P. Remanence of sentence prosody in Romance languages. In: *ExLing 2012, the 5th Tutorial and Research Workshop on Experimental Linguistics.* Athens; 2016.
 104. D'Alessandro C, Rosset S, Rossi J-P. The pitch of short-duration fundamental frequency glissandos. *J Acoust Soc Am.* 1998;104(4):2339–48.
 105. Kleinschmidt DF, Jaeger TF. Robust speech perception: recognize the familiar, generalize to the similar, and adapt to the novel. *Psychol Rev.* 2015;122:148–203.
 106. Caporali SA, Arieta AM. Reconhecimento de fala: estudo comparativo entre grupos com e sem queixa de percepção da fala. *Rev da Soc Bras Fonoaudiol.* 2004;9(3):129–35.
 107. Magalhães LA, Cimonari PM, Novaes BCAC. Avaliação de percepção de fala em crianças com deficiência auditiva usuárias de aparelho de amplificação sonora: a questão do instrumento e seus critérios. *Rev da Soc Bras Fonoaudiol.* 2007;12(3):221–32.
 108. Jiang X, Pell MD. The sound of confidence and doubt. *Speech Commun* [Internet]. 2017;88:106–26. Available from: <http://www.sciencedirect.com/science/article/pii/S0167639316301509>
 109. Lee FF. Time compression and expansion of speech by the sampling method. *J audio Eng Soc.* 1972;20(9):738–42.
 110. Tree JEF, Clark HH. Pronouncing “the” as “thee” to signal problems in

- speaking. *Cognition* [Internet]. 1997;(62):151–67. Available from: [https://doi.org/10.1016/S0010-0277\(96\)00781-0](https://doi.org/10.1016/S0010-0277(96)00781-0)
111. Clark HH, Fox Tree JE. Using uh and um in spontaneous speaking. *Cognition*. 2002;84(1):73–111.
 112. Fox Tree JE. Listeners' uses of um and uh in speech comprehension. *Mem Cogn*. 2001;29(2):320–6.
 113. Overath T, Paik JH. From acoustic to linguistic analysis of temporal speech structure: Acousto-linguistic transformation during speech perception using speech quilts. *bioRxiv*. 2019;1–24.
 114. Shrivastav R, Sapienza CM, Nandur V. Application of psychometric theory to the measurement of voice quality using rating scales. *J Speech, Lang Hear Res*. 2005;48:323–35.
 115. Narasimhan S V., Rashmi R. Multiparameter Voice Assessment in Dysphonics: Correlation Between Objective and Perceptual Parameters. *J Voice* [Internet]. 2020; Available from: <https://doi.org/10.1016/j.jvoice.2020.06.009>
 116. Eadie T, Sroka A, Wright DR, Merati A. Does knowledge of medical diagnosis bias auditory-perceptual judgments of dysphonia? *J Voice*. 2011;25:420–9.
 117. Eadie TL, Baylor CR. The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *J Voice*. 2006;20:527–44.
 118. Jiang X, Gossack-Keenan K, Pell MD. To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Q J Exp Psychol*. 2020;73(1):55–79.
 119. Mauchand M, Caballero JA, Jiang X, Pell MD. Immediate online use of prosody reveals the ironic intentions of a speaker: neurophysiological evidence. *Cogn Affect Behav Neurosci*. 2020;
 120. Pinto I, Firmino-Machado J, Castro E, Alves S, Helena D, Condé A. The effects on the acoustic parameters and auditory-perceptive characteristics of voice in children submitted to adenoidectomy with or without tonsillectomy. *Int J Pediatr Otorhinolaryngol* [Internet]. 2019;125:51–5. Available from: <https://doi.org/10.1016/j.ijporl.2019.06.021>

121. Englert M, Madazio G, Gielow I, Lucero J, Behlau M. Learning factor influence on the perceptual-auditory analysis. *CoDAS*. 2018;30(3).
122. Celeste LC. A prosódia na expressão de atitudes na fala de indivíduos com e sem gagueira. Universidade Federal de Minas Gerais; 2010.
123. Englert M, Barsties v. Latoszek B, Maryn Y, Behlau M. Validation of the acoustic breathiness index to the Brazilian Portuguese language. *Logop Phoniatr Vocology* [Internet]. 2020;0(0):1–7. Available from: <https://doi.org/10.1080/14015439.2020.1864467>
124. Fernandes DE, Andrada e Silva MAD. Superagudos: análise perceptivo-auditiva da voz e autorrelato em sopranos profissionais. *CoDAS*. 2020;32(4).
125. Marino VCDC, Dutka JDCR, Manicardi FT, Gifalli G, Silva PP, Pegoraro-Krook MI. Influência de estímulos de fala na identificação perceptivo-auditiva da hipernasalidade em indivíduos com fissura labiopalatina. *CoDAS*. 2020;32(6).
126. Overath T, McDermott JH, Zarate JM, Poeppel D. The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nat Neurosci*. 2015;18(6):903–11.
127. Shtyrov Y, Kujala T, Palva S, Ilmoniemi RJ, Näätänen R. Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *Neuroimage*. 2000;12(6):657–63.
128. Arnold JE, Tanenhaus MK, Altmann RJ, Fagnano M. The old and thee, uh, new: Disfluency and reference resolution. *Psychol Sci*. 2004;15(9):578–82.
129. Swerts M, Kraemer E. Audiovisual prosody and feeling of knowing. *J Mem Lang*. 2005;53(1):81–94.
130. Corley M, MacGregor LJ, Donaldson DI. It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition*. 2007;105(3):658–68.
131. Brennan SE, Schober MF. How Listeners Compensate for Disfluencies in Spontaneous Speech. *J Mem Lang*. 2001;44(2):274–96.
132. MacGregor LJ, Corley M, Donaldson DI. Not all disfluencies are equal: The

- effects of disfluent repetitions on language comprehension. *Brain Lang* [Internet]. 2009;111(1):36–45. Available from: <http://dx.doi.org/10.1016/j.bandl.2009.07.003>
133. MacGregor LJ, Corley M, Donaldson DI. Listening to the sound of silence: Disfluent silent pauses in speech have consequences for listeners. *Neuropsychologia*. 2010;48(14):3982–92.
 134. Fraundorf SH, Watson DG. The disfluent discourse: Effects of filled pauses on recall. *J Mem Lang* [Internet]. 2011;65(2):161–75. Available from: <http://dx.doi.org/10.1016/j.jml.2011.03.004>
 135. Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am*. 1994;95(2):1053–64.
 136. Rose RL. The structural signaling effect of silent and filled pauses. In: *The 9th Workshop on Disfluency in Spontaneous Speech*. 2019. p. 19.
 137. Trouvain J. On the comprehension of extremely fast synthetic speech. *Saarl Work Pap Linguist*. 2007;1(1):5–13.
 138. Black JW. The effect of delayed side-tone upon vocal rate and intensity. *J Speech Hear Disord*. 1951;16(1):56–60.
 139. Moein N, Mohamadi R, Rostami R, Nitsche M, Zomorodi R, Ostadi A, et al. Investigation of the effect of Delayed Auditory Feedback and transcranial Direct Current Stimulation (DAF-tDCS) treatment for the enhancement of speech fluency in adults who stutter: a Randomized Controlled Trial. *Res Sq*. 2020;Preprint:1–16.
 140. Romeiser SA, Kiley SJ, Nocella NJ. The Effects of Altered Auditory Feedback (AAF) on Fluency in Adults Who Stutter: A Systematic Review. *Commun Sci Disord Syst Rev Publ*. 2019;
 141. Perrucini DS, Cardoso AC V, Moura RBOD, Lorena MCMD, Buzzeti PBMDM, Oliveira CMCD. Efeito da retroalimentação auditiva atrasada na fala e leitura de taquifêmicos. *Audiol - Commun Res*. 2017;22(e1795).
 142. Buzzeti PBMDM, Oliveira CMCD. Efeito imediato do atraso da retroalimentação auditiva nas disfluências típicas da gagueira. *Rev CEFAC*.

- 2018;20(3):281–90.
143. Fernandes LDC, Bomfim DAS, Machado GC, Andrade CLD. Influência da retroalimentação auditiva nos parâmetros acústicos vocais de indivíduos sem queixas vocais. *Audiol - Commun Res.* 2018;e-1785.
 144. Ferreira AB de H. *Novo Aurélio Século XXI: o dicionário da língua portuguesa.* Rio de Janeiro: Nova Fronteira; 1999.
 145. Smith VL, Clark HH. On the Course of Answering Questions. *J Mem Lang.* 1993;32:25–38.
 146. Bellinghausen, C. Fangmeier T, Schröder B, Keller J, Drechsel S, Birkholz P, Elst L, et al. On the role of disfluent speech for uncertainty in articulatory speech synthesis. In: *The 9th Workshop on Disfluency in Spontaneous Speech.* 2019. p. 39–42.
 147. Arnold J, Hudson Kam C, Tanenhaus M. If you say Thee uh you are describing something hard: The online attribution of disfluency during reference comprehension. *J Exp Psychol Learn Mem Cogn.* 2007;33:914–30.
 148. Barr DJ. Paralinguistic correlates of conceptual structure. *Psychon Bull Rev* [Internet]. 2003;10(2):462–7. Available from: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-0346902186&doi=10.3758%2F03196507&partnerID=40&md5=3f508980746f4f9a8acebc555f81e4a1>
 149. McAleer P, Todorov A, Belin P. How do you say ‘Hello’? Personality impressions from brief novel voices. *PLoS One.* 2014;9(3):e90779.
 150. Belin P, Boehme B, McAleer P. The sound of trustworthiness: Acoustic-based modulation of perceived voice personality. *PLoS One.* 2017;12(10):e0185651.
 151. Pon-Barry H, Shieber S. Assessing self-awareness and transparency when classifying a speaker’s level of certainty. In: *Proceedings of the International Conference on Speech Prosody.* 2010. p. 1–4.
 152. Pon-Barry H, Shieber S. The importance of sub-utterance prosody in predicting level of certainty. In: *The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics: Proceedings of*

- NAACL-HLT 2009. Boulder, Colorado; 2009. p. 105.
153. Gomes DA do V. Aspectos acústicos prosódicos de atitudes comunicativas ligadas à confiança: revisão sistemática exploratória e proposta de inclusão de parâmetros prosódicos no treinamento fonoaudiológico. Universidade de Brasília - UnB; 2019.
 154. Martin P. WinPitch W10. Pitch Instruments Inc;
 155. Costa LMO, Martins-Reis V de O, Celeste LC. Methods of analysis speech rate: A pilot study. *Codas*. 2016;28(1):41–5.
 156. Cangı ME, İŞILDAR A, Tekin A, Saraç AB. A preliminary study of normative speech rate values of Turkish speaking adults. *ENT Updat*. 2020;10(3):381–9.
 157. Chang AC-S. Speech Rate in Second Language Listening. *TESOL Encycl English Lang Teach*. 2018;1–6.
 158. Akira H, Vogel C, Luz S, Campbell N. Speech rate calculations with short utterances: A study from a speech-to-speech, machine translation mediated map task. In: *LREC 2018 - 11th International Conference on Language Resources and Evaluation*. 2019. p. 3176–83.
 159. Tivadar H. Speech Rate in Phonetic -Phonological Anal Ysis of Public Speech (Using the Example of Political and Media Speech). *Jazykoved Cas*. 2017;68(1):37–56.
 160. Damhoureyeh MA, Darawsheh WB, Qa'dan WN, Natour YS. Preliminary Speech Rate Normative Data in Adult Jordanian Speakers. *J Lang Teach Res*. 2020;11(2):204.
 161. Martins V de O, Andrade CRF de. Perfil evolutivo da fluência da fala de falantes do português brasileiro. *Pró-Fono Rev Atualização Científica*. 2008;20(1):7–12.
 162. Mussalim F, Bentes AC. Introdução à linguística—domínios e fronteiras. Cortez; 2009.
 163. Marchal A, Reis C. Produção da fala. Editora UFMG; 2012.
 164. Yang IY. Pause—Tonic Stress Interaction in English L2 Speech of Korean

- Talkers with Different Proficiency Levels. *Korean J English Lang Linguist*. 2018;18(1):1–29.
165. Kang O. Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System* [Internet]. 2010;38(2):301–15. Available from: <http://dx.doi.org/10.1016/j.system.2010.01.005>
 166. Zeches JT, Yorkston KM. Pause Structure in Narratives of Neurologically Impaired and Control Subjects. *Clin Aphasiology*. 1995;23:155–63.
 167. Englert M, Lima L, Constantini AC, Latoszek BB v., Maryn Y, Behlau M. Acoustic Voice Quality Index - AVQI for brazilian portuguese speakers: Analysis of different speech material. *Codas*. 2019;31(1):1–7.
 168. Drozdova P, van Hout R, Scharenborg O. L2 voice recognition: The role of speaker-, listener-, and stimulus-related factors. *J Acoust Soc Am*. 2017;142(5):3058–68.
 169. Lavan N, Burston LFK, Ladwa P, Merriman SE, Knight S, McGettigan C. Breaking voice identity perception: Expressive voices are more confusable for listeners. *Q J Exp Psychol*. 2019;72(9):2240–8.
 170. Mauchand M, Vergis N, Pell MD. Ironic tones of voices. In: *Proceedings of the International Conference on Speech Prosody*. 2018. p. 443–7.
 171. Fuchs S, Petrone C, Krivokapić J, Hoole P. Acoustic and respiratory evidence for utterance planning in German. *J Phon*. 2013;41(1):29–47.
 172. Winters SJ, Levi S V, Pisoni DB. Identification and discrimination of bilingual talkers across languages. *J Acoust Soc Am*. 2008;123(6):4524–38.
 173. Levi S V, Schwartz RG. The development of language-specific and language-independent talker processing. *J Speech, Lang Hear Res*. 2013;56:913–20.
 174. Juste F, Andrade CRF de. Typology of speech disruptions and grammatical classes in stuttering and fluent children. *Pró-Fono Rev Atualização Científica*. 2006;18:129–40.
 175. Weber-Fox C, Neville HJ. Sensitive periods differentiate processing of open- and closed-class words. *J Speech, Lang Hear Res*. 2001;44(6):1338–53.

176. Grela B, Rashiti L, Soares M. Dative prepositions in children with specific language impairment. *Appl Psycholinguist*. 2004;25:467–80.
177. Conture EG. *Stuttering, its nature, diagnosis, and treatment*. Boston: Allyn and Bacon; 2001.
178. Yairi E, Ambrose NG. Early childhood stuttering. For *Clinicians by Clinicians*. Pro-Ed, editor. Austin, TX; 2005.
179. Yairi E. Applications of disfluencies in measurements of stuttering [Letter to the editor]. *J Speech Hear Res*. 1996;39:402–4.
180. Riley GD. *Stuttering severity instrument for children and adults*. 4th ed. Austin, TX: Pro-Ed; 2009.
181. Pallaud B, Bertrand R, Blache P, Prévot L, Rauzy S. Suspensive and Disfluent Auto-Breaks in Spoken French. (Submitted). 2016;(February 2019).
182. Medeiros H, Moniz H, Batista F, Trancoso I, Nunes L. Disfluency detection based on prosodic features for university lectures. In: *INTERSPEECH*. 2013. p. 2629–33.
183. Pell MD. Influence of emotion and focus location on prosody in matched statements and questions. *J Acoust Soc Am*. 2001;109(4):1668–80.
184. Ladd DR, Johnson C. 'Metrical' factors in the scaling of sentence-initial accent peaks. *Phonetica*. 1987;44(4):238–45.
185. Hawkins PR. The syntactic location of hesitation pauses. *Lang Speech*. 1971;14(3):277–88.
186. Tian Y, Maruyama T, Ginzburg J. Self Addressed Questions and Filled Pauses: A Cross-linguistic Investigation. *J Psycholinguist Res*. 2017;46(4):905–22.
187. Goldman-Eisler F. Speech Production and the Predictability of Words in Context. *Q J Exp Psychol*. 1958;10(2):96–106.
188. Boomer DS, Dittmann AT. Hesitation Pauses and Juncture Pauses in Speech. *Lang Speech*. 1962;5(4):215–20.
189. Goldman-Eisler F. *Psycholinguistics: Experiments in spontaneous speech*. 1968.

190. Carrara AE, Mourão LF, Ferraz HB, Behlau MS, Pontes PAL, Andrade LAF. Effect of voice rehabilitation on oral communication of Parkinson's disease patients. *Acta Neurol Scand.* 1997;96:199–205.
191. Ramig LO, Bonitati C, Lemke J, Horii Y. Voice treatment for patients with Parkinson's disease: development of an approach and preliminary efficacy data. *J Med Speech Lang Pathol.* 1994;3:191–209.
192. Coutinho SB, Diaféria G, Oliveira G, Behlau M. Voz e fala de Parkinsonianos durante situações de amplificação, atraso e mascaramento. *Pro-Fono.* 2009;21(3):219–24.

Anexos

Anexo 1 – Parecer Consubstanciado do CEP

UNB - FACULDADE DE
CEILÂNDIA DA UNIVERSIDADE
DE BRÁSÍLIA



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: EFICÁCIA DE TREINAMENTO FONOAUDIOLÓGICO DAS HABILIDADES COMUNICATIVAS COM FOCO EM SEMINARISTAS DO SEMINÁRIO MAIOR DE BRÁSÍLIA - SMAB

Pesquisador: Dianete do Valle

Área Temática:

Versão: 6

CAAE: 64822717.6.0000.8093

Instituição Proponente: Faculdade de Ceilândia - FUNDAÇÃO UNIVERSIDADE DE BRÁSÍLIA

Patrocinador Principal: Financiamento Próprio

DADOS DA NOTIFICAÇÃO

Tipo de Notificação: Envio de Relatório Parcial

Detalhe:

Justificativa: Envio de relatório parcial após mudança no desenho do estudo.

Data do Envio: 11/06/2019

Situação da Notificação: Parecer Consubstanciado Emitido

DADOS DO PARECER

Número do Parecer: 3.420.515

Apresentação da Notificação:

Segundo os autores: "Introdução: O meio acadêmico, em especial a sala de aula, é um ambiente extremamente profícuo de situações de exposição do indivíduo, sendo a necessidade de falar em público um dos motivos da busca do treinamento para oratória. Objetivo: O presente trabalho tem como objetivo, aferir a eficácia da aplicação de treinamento da oratória, em seminaristas do Seminário Maior de Brasília, se valendo para tal da análise dos padrões prosódicos de velocidade de fala, pausa e ênfase no discurso em apresentações orais, do acompanhamento das técnicas fonoaudiológicas da performance comunicativa e treinamento fonoaudiológico, da análise dos padrões prosódicos reorganizados pós-treinamento fonoaudiológico e da comparação com os padrões prosódicos pré-tratamento. Métodos: Trata-se de um projeto de estudo longitudinal prospectivo baseado na análise do discurso de 50 seminaristas do Seminário Maior de Brasília, os

Endereço: UNB - Prédio da Unidade de Ensino e Docência (UED), Centro Metropolitano, conj. A, lote 01, Sala AT07/66
Bairro: CEILÂNDIA SUL (CEILÂNDIA) **CEP:** 72.220-900
UF: DF **Município:** BRÁSÍLIA
Telefone: (61)3107-8434 **E-mail:** cep.fce@gmail.com

UNB - FACULDADE DE
CEILÂNDIA DA UNIVERSIDADE
DE BRASÍLIA



Continuação do Parecer: 3.420.515

quais deverão cursar a disciplina "Aprimoramento da Performance Comunicativa: Atividade Prática". Os procedimentos deste projeto consistirão em cinco etapas distintas: gravação inicial, treinamento fonoaudiológico, gravação final, análise prosódica, análise acústica através do programa PRAAT e análise dos resultados. Resultado esperado: O presente estudo espera apresentar ao seu final, uma validação do treinamento fonoaudiológico para a comunicação de seminaristas."

Objetivo da Notificação:

Apresentar o relatório parcial.

Avaliação dos Riscos e Benefícios:

Os autores referem não ter transcorrido riscos, mesmo aqueles previstos inicialmente no projeto. Segundo os autores não há clareza sobre os possíveis benefícios, contudo, após a realização da devolução de dados, houve mudança da conduta dos participantes em relação a oratória.

Comentários e Considerações sobre a Notificação:

O relatório foi adequadamente elaborado.

Considerações sobre os Termos de apresentação obrigatória:

Todos os documentos foram adequadamente apresentados.

Recomendações:

Não há.

Conclusões ou Pendências e Lista de Inadequações:

Não há pendências.

Considerações Finais a critério do CEP:

Relatório parcial aprovado.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Envio de Relatório Parcial	relatorio_parcial.doc	11/06/2019 10:01:09	Dianete do Valle	Postado

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Endereço: UNB - Prédio da Unidade de Ensino e Docência (UED), Centro Metropolitano, conj. A, lote 01, Sala AT07/66
Bairro: CEILÂNDIA SUL (CEILÂNDIA) CEP: 72.220-900
UF: DF Município: BRASÍLIA
Telefone: (61)3107-8434 E-mail: cep.fce@gmail.com

UNB - FACULDADE DE
CEILÂNDIA DA UNIVERSIDADE
DE BRASÍLIA



Continuação do Parecer: 3.420.515

Não

BRASILIA, 27 de Junho de 2019

Assinado por:
Danielle Kaiser de Souza
(Coordenador(a))

Endereço: UNB - Prédio da Unidade de Ensino e Docência (UED), Centro Metropolitano, conj. A, lote 01, Sala AT07/66
Bairro: CEILANDIA SUL (CEILANDIA) **CEP:** 72.220-900
UF: DF **Município:** BRASILIA
Telefone: (61)3107-8434 **E-mail:** cep.fce@gmail.com

Anexo 2 – Textos argumentativos apresentados aos participantes

Você está recebendo 1 texto com duas opiniões antagônicas sobre o mesmo tema. Escolha um dos posicionamentos apresentados e argumente com o objetivo de convencer o seu interlocutor sobre a tese que você concorda.

“Graves dificuldades são enfrentadas pelos jovens chamados ao sacerdócio. Tais dificuldades são devidas à vulnerabilidade e ao enfraquecimento da identidade espiritual dos candidatos, como também ao impacto de alguns dos atuais modelos culturais e à frágil situação das famílias de origem.”

“Por outro lado, há os que defendem que a vocação sacerdotal nada tem a ver com vulnerabilidade ou enfraquecimento espiritual. A vocação é dom de Deus e sendo assim é inquestionável. As dificuldades que os jovens chamados ao sacerdócio enfrentam em nada diferem-se daquelas enfrentadas pelos jovens na sociedade atual. Todos os jovens são bombardeados pelos mesmos modelos sociais atuais, que parecem ser contrários aos da igreja.”

A Formação Sacerdotal nos Seminários – Recomendações pastorais da Assembleia Plenária da Pontifícia Comissão para a América Latina. Fevereiro de 2009. Disponível em: http://www.vatican.va/roman_curia/congregations/cbishops/pcal/documents/rc_cbishops_pcal_20090220_pastorale_po.html.

1) Você deverá introduzir a sua fala com a frase abaixo:

Vou explicar agora porque concordo plenamente com os seguintes argumentos...