

O impacto do contexto da disfluência na organização temporal de consoantes na gagueira

The impact of a dysfluency environment on the temporal organization of consonants in stuttering

Leticia Correa Celeste¹, Vanessa de Oliveira Martins-Reis²

RESUMO

Objetivo: Analisar e comparar o *voice onset time* (VOT) em falantes do português brasileiro com e sem gagueira, com foco em três momentos diferentes de discurso: fala fluente, pré-disfluência e pós-disfluência.

Métodos: Foram feitas gravações da fala de 20 participantes (n=10 com gagueira e n=10 sem gagueira). Os dados foram transcritos e segmentados para análise acústica e segmentos de oclusivas não vozeadas do Português Brasileiro, /p/, /t/ e /k/ foram extraídos. Os segmentos foram classificados por grupo - se foram produzidos por pessoas que gaguejam (PG) ou por pessoas que não gaguejam (PnG) - e de acordo com o contexto/ambiente de fala (ou seja, na fala fluente, na pré-disfluência, e na pós-disfluência). Os testes Friedman e Wilcoxon foram utilizados para comparação dentro dos grupos e o teste de Mann-Whitney, em comparações intergrupos. As análises estatísticas foram executadas usando SPSS 14.0, com nível de significância de $\alpha=0,05$. **Resultados:** O VOT de falantes com e sem gagueira diferiu mais no ambiente de pré-disfluência, durante o qual, falantes com gagueira apresentaram VOT significativamente mais longo do que falantes que não gaguejavam. Depois de passar por um momento de disfluência, no entanto, o VOT dos participantes com gagueira retornou a medidas semelhantes às pessoas que não gaguejavam. **Conclusão:** Nos ambientes de pré-disfluência e pós-disfluência, as PG produzem VOT mais longos do que as PnG. Já no discurso fluente de PG, as oclusivas se comportam de forma diferente. São discutidas as implicações desses resultados para o controle motor da fala.

Descritores: Gagueira; Acústica; Planejamento; Destreza motora; Fala

ABSTRACT

Purpose: To analyze and compare the voice onset time (VOT) in Brazilian Portuguese speakers who stutter and those who do not, focusing on three different moments of speech: fluent, pre-dysfluent and post-dysfluent environments. **Methods:** Twenty participants (n=10 with stuttering and n=10 without stuttering) were recorded. The data were transcribed and segmented for acoustic analysis, and it was extracted tokens of Brazilian Portuguese voiceless stops /p/, /t/ and /k/. Tokens were classified according to whether they were produced by people who stutter (PWS) or by people who do not stutter (PWnS), and according to their environment in speech (i.e., in fluent speech, pre-dysfluency, and post-dysfluency). For comparisons within groups it were used the Friedman and Wilcoxon tests, and the Mann-Whitney test was used in intergroup comparisons. Statistical analyses were executed using SPSS 14.0 with the significance level set at $\alpha=0.05$. **Results:** VOT in stuttering and non-stuttering speakers differed most in the environment of pre-dysfluencies, during which stuttering speakers show significantly longer VOT than speakers who do not stutter. After passing through a moment of dysfluency, however, stuttering speakers' VOT returns to measures similar to non-stuttering speakers'. **Conclusion:** In pre-dysfluent and post-dysfluent speech, PWS produces longer VOT than PWnS. In the fluent speech of PWS, the stops behave differently. The implications of these results for speech motor control are discussed.

Keywords: Stuttering; Acoustics; Planning; Motor skills; Speech

Trabalho realizado no Laboratório de Fonética, Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brasil.

(1) Universidade de Brasília – UnB – Brasília (DF), Brasil.

(2) Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brasil.

Financiamento: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Conflito de interesses: Não

Contribuição dos autores: LCC pesquisador principal, elaboração da pesquisa, elaboração do cronograma, levantamento da literatura, coleta e análise dos dados, redação do artigo, submissão e trâmites do artigo; VOMR pesquisadora, elaboração da pesquisa, análise dos dados, redação do artigo.

Endereço para correspondência: Leticia Correa Celeste. Centro Metropolitano, Conjunto A, lote 01, Brasília (DF), Brasil, CEP: 72220-900.

E-mail: leticiaceleste@unb.br

Recebido em: 3/9/2014; **Aceito em:** 9/3/2015

INTRODUÇÃO

Acusticamente, a produção de consoantes oclusivas envolve três etapas: (1) o intervalo de oclusão; (2) a explosão (*burst*) e (3) o intervalo entre a explosão e o início da fonação. O intervalo de tempo entre o início da explosão e o início da fonação é chamado de *voice onset time* (sem tradução oficial para o português brasileiro, mas que poderia ser interpretado como tempo de latência até o início do vozeamento), ou VOT^(1,2). Para uma fala fluente, os falantes necessitam coordenar eficientemente a musculatura oral-facial e a vibração das pregas vocais, sendo que essas habilidades muitas vezes falham em pessoas que gaguejam (PCG)⁽³⁾. Essa incoordenação pode resultar em rupturas da fala, comprometendo a estabilidade articulatória. Esse fenômeno pode ser observado em medições acústicas e por meio de uma análise mais aprofundada de consoantes oclusivas dentro de certos contextos. As medidas de VOT podem esclarecer algumas questões sobre a instabilidade articulatória em PCG.

De forma geral, a produção da fala exige três fases: elaboração, preparação e execução. Essas fases devem ser controladas com precisão e a sua base temporal, coordenada⁽⁴⁾. Cada gesto articulatório da fala pode ser visto como uma ‘estrutura coordenada’ de diferentes fatores⁽¹⁾, quer com base em teorias que defendem exclusivamente o controle central, ou em teorias que defendem a fala como uma tarefa dinâmica de movimentos articulatórios coordenados⁽⁴⁾. Cada som requer uma coordenação dos movimentos musculares específicos, que devem ser modificados ao longo de um período de tempo.

Para produzir as consoantes oclusivas, é necessário, inicialmente, que o falante decida se as pregas vocais irão participar, isto é, se o vozeamento estará envolvido.

Independentemente da vibração das pregas vocais, o palato deve ser levantado, de modo que o trato nasal seja bloqueado e o ar impedido de escapar através da cavidade oral⁽¹⁾, gerando um acúmulo de ar e aumento da pressão na cavidade intraoral. Então, os articuladores envolvidos liberam a corrente de ar⁽⁵⁾.

A definição de quais articuladores estão envolvidos é importante, pois há uma relação entre o ponto de articulação e a duração da liberação de pressão de ar intraoral⁽⁶⁾. Em um estudo com diversas línguas⁽⁷⁾ os autores mostraram que as oclusivas velares apresentam VOT maior do que as oclusivas anteriores. A complexidade dos movimentos articulatórios envolvidos na produção das oclusivas confirma a necessidade de estudos com pessoas que têm distúrbios de comunicação.

Devido à complexidade articulatória das oclusivas, o VOT foi estudado em diferentes distúrbios da comunicação, como a dislexia⁽⁸⁾, a gagueira^(3,9), afasia e disartria⁽¹⁰⁾, entre outros. Estudos recentes sugerem que os gânglios basais têm um papel importante na temporização da fala⁽⁴⁾ e, quando estes são afetados por qualquer desordem, os falantes podem apresentar grandes variações na duração das oclusivas. Um estudo de indivíduos com doença de Parkinson mostrou que esses pacientes

tendem a misturar padrões de VOT, como, por exemplo, através do tensionamento das consoantes anteriores, porque quanto mais anterior a oclusiva, menor o VOT⁽⁷⁾. Em pesquisa sobre a relação entre a Coreia de Huntington e as oclusivas não vozeadas⁽¹⁾, houve uma diminuição da média do VOT ao longo do tempo. Pesquisas também sugerem que a gagueira pode estar relacionada a uma disfunção dos gânglios da base⁽¹¹⁻¹³⁾.

A gagueira é considerada um distúrbio da fluência da fala⁽¹⁴⁾, com foco em dificuldades com o controle motor da fala^(11,12,15-17). PCG apresentam perturbação na sua capacidade de processamento de discurso com estabilidade temporal⁽¹⁸⁾ e, por essa razão, mais e mais pesquisadores têm estudado o VOT entre PCG^(3,9,19,20), uma vez que o VOT é uma metodologia padrão para investigar características temporais da fala. Um estudo mostrou⁽²¹⁾ que PCG produzem um VOT mais longo nas oclusivas não vozeadas, enquanto em outro⁽²²⁾, foi observado que crianças com gagueira apresentam um VOT maior tanto para as oclusivas vozeadas, quanto para as oclusivas não vozeadas, quando comparadas a crianças sem gagueira.

Em um estudo em falantes com e sem gagueira nativos do alemão, os participantes foram instruídos a produzir sílabas isoladas (/papapa/, /tatatas/ e /kakakas/), com ênfase na segunda sílaba. Os resultados mostraram que, mesmo no discurso fluente, os participantes com gagueira apresentaram maior variação na duração da primeira sílaba enunciada⁽²⁰⁾. Em uma pesquisa com /p/ em frases, analisando a diferença entre as produções de participantes com e sem gagueira por meio de eletroglotografia e medições acústicas, os resultados sugeriram uma diferença entre os dois grupos na duração dos intervalos de eventos do subsistema orolaringeo⁽⁹⁾.

Um estudo brasileiro comparou como gagos e não gagos produzem as oclusivas precedidas por uma vogal (por exemplo, [ap]). Os resultados indicaram que os segmentos das PCG eram mais longos e a diferença aumentou com a gravidade da gagueira⁽²³⁾.

A fim de verificar o uso de VOT como parâmetro de naturalidade de fala, por meio da comparação dos resultados de dois tratamentos de gagueira, autores analisaram a produção de /b/ em sentenças proferidas por falantes com e sem gagueira⁽³⁾. Eles concluíram que o VOT é um parâmetro satisfatório: as PCG apresentaram VOT mais longos, o que foi interpretado como uma disfunção do processamento neuromotor envolvido na fala⁽³⁾.

Pesquisas sobre VOT e seu papel na gagueira oferecem recursos importantes para a compreensão de processos articulatórios. Neste sentido, o presente estudo contribui oferecendo uma descrição acústica de VOT de falantes do Português Brasileiro com e sem gagueira, considerando o ambiente de disfluência como principal ferramenta na compreensão de erro de planejamento motor. O objetivo deste estudo foi verificar como o *voice onset time* (VOT) é produzido por pessoas que gaguejam, com foco em três momentos diferentes de discurso: ambientes fluentes, pré-disfluentes e pós-disfluentes.

MÉTODOS

O estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais (UFMG), sob protocolo CAAE – 0308.0.203.000-11. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido, conforme resolução 196/96 do Conselho Nacional de Saúde.

Participaram do estudo 20 indivíduos, divididos em dois grupos: dez no grupo experimental (GE) e dez no grupo controle (GC). Todos os indivíduos eram do gênero masculino com idades entre 20 e 45 anos, já que gênero e idade parecem influenciar as medidas de VOT⁽²⁴⁾. Os critérios de inclusão no estudo foram: ausência de problemas gerais de saúde e resultados negativos em triagem para distúrbios da comunicação (de linguagem, auditivos, neurológicos, cognitivos, etc). Todos os participantes passaram por triagem realizada por uma fonoaudióloga, para determinar a inclusão ou não no estudo. Todos os participantes do GE apresentavam gagueira do desenvolvimento com gravidade de moderada a grave, de acordo com a Escala de Iowa⁽²⁵⁾, sem outro distúrbio sensorial ou problemas de fala e audição. Nenhum deles realizou qualquer tipo de tratamento para melhora da fala antes de participar do estudo. O grupo controle apresentou fluência normal, a partir do Protocolo para Avaliação da Fluência de Fala⁽²⁶⁾.

Dois conjuntos de dados foram coletados para o presente estudo. Para coleta das amostras de fala autoexpressiva, foi utilizada a metodologia proposta pelo Protocolo para Avaliação da Fluência de Fala⁽²⁶⁾. Desta forma, uma figura foi apresentada a cada participante, com a seguinte instrução: “Por favor, olhe esta figura e fale o que quiser sobre ela”. As amostras de fala foram gravadas.

Para eliciar a produção das consoantes oclusivas não vozeadas do Português Brasileiro, solicitou-se que os participantes lessem dez sentenças em voz alta. Inicialmente, as sentenças foram lidas silenciosamente. As sentenças caracterizaram-se pela presença das oclusivas não vozeadas /p/, /t/ e /k/, em diferentes posições (e.g., início de palavra *Eu gosto da Carol*; entre vogais *eu deixei o recado*). Posteriormente, os participantes foram instruídos a ler as sentenças três vezes, em voz alta. Em seguida, em um ambiente tratado acusticamente, os participantes foram instruídos a ler as frases em voz alta três vezes, em um microfone conectado a um computador com Praat instalado (versão 5.1.02, 1992-2010, disponível gratuitamente em www.praat.org). Os dados foram armazenados para edição e análise posterior.

Todas as amostras de fala foram transcritas ortograficamente, a fim de identificar as sílabas fluentes e disfluentes e todas as disfluências foram classificadas (hesitação, interjeição, revisão, palavra não terminada, repetição de palavra, repetição de segmento, repetição de frase, repetição de sílaba, repetição de som, prolongamento, bloqueio, pausa e intrusão de sons ou segmentos). As sílabas foram contadas, excluindo-se interjeições, revisões, palavras não terminadas, repetições de palavras,

repetições de segmentos, repetições de frases e repetições de sílabas. Esses valores finais foram utilizados para calcular as taxas de elocução e frequências das rupturas da fala.

A variação da taxa de elocução tem se mostrado um fator importante na análise do VOT⁽²¹⁾. No entanto, as disfluências da fala dos participantes não representam, necessariamente, o tempo de articulação de cada sílaba. Por essas razões, foram calculadas as medidas de velocidade de fala para cada enunciado gravado. A velocidade de fala foi calculada dividindo-se o número de sílabas pelo tempo articulação (duração das pausas e disfluências subtraídos da duração total do enunciado), resultando na taxa de articulação. A taxa média de articulação foi de 5,5 sílabas/segundo. Frases cuja taxa de articulação obtiveram um desvio padrão mais elevado que 1,5 foram eliminadas.

Os dados transcritos foram segmentados para análise acústica e trechos contendo as consoantes não vozeadas do Português Brasileiro - /p/, /t/ e /k/ - foram identificados e extraídos. As consoantes incluídas na análise ocorreram em diferentes posições da frase, mas deveriam estar no contexto VCV (vogal-consoante-vogal). Além disso, a análise foi realizada em sílabas em posição átona. Sílabas tônicas e sílabas encontradas no final dos enunciados foram excluídas, porque diferem, em duração, das sílabas átonas.

Foram avaliados três parâmetros das oclusivas não vozeadas: o intervalo ou silêncio de oclusão, o VOT e a duração total. O intervalo de oclusão foi marcado entre a extremidade da vogal precedente e o início da explosão (*burst*) da oclusiva. As fronteiras do VOT foram marcadas a partir do início da explosão (*burst*) da oclusiva até o início da vogal seguinte (Figura 1). A triangulação da onda sonora, espectrograma e curva de F0 foram utilizados para determinar essas fronteiras.

Para determinar o efeito do ambiente das disfluências nas oclusivas não vozeadas, cada parâmetro das oclusivas foi investigado nos seguintes contextos de fala:

1. Fluência: não houve disfluência na palavra que precedia ou seguia a palavra que continha o segmento analisado;
2. Pré-disfluência: a oclusiva ocorreu na sílaba que precedia a disfluência;
3. Pós-disfluência: a oclusiva ocorreu em uma sílaba imediatamente após a disfluência.

O número de consoantes analisadas foi maior para PCG que para pessoas sem gagueira (PSG). Vários PCG pediram para reler as frases e esses segmentos adicionais foram incluídos nos dados.

As medidas de confiabilidade interjuízes foram obtidas para determinar momentos de disfluência na fala de PCG. A fim de identificar momentos de disfluência, um teste perceptivo foi administrado a três falantes nativos brasileiros. Nenhum deles apresentou gagueira ou qualquer outro distúrbio de fala, linguagem ou auditivo. O teste foi aplicado individualmente, em um lugar silencioso, na casa de cada juiz. Os juízes receberam um pedaço de papel, listando as frases lidas pelos participantes com gagueira. Enquanto ouviam as gravações, cada juiz foi instruído

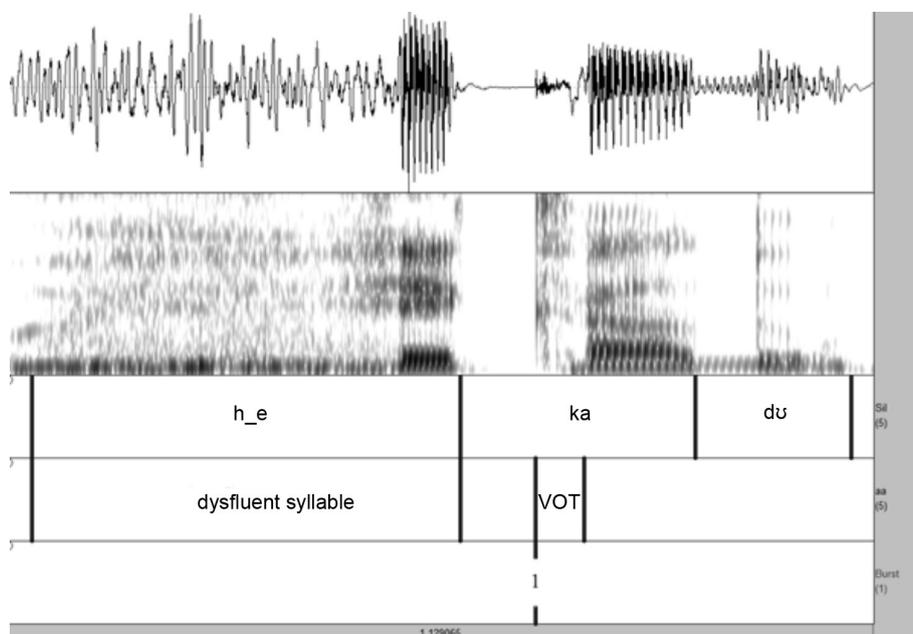


Figura 1. Marcação de fronteira. Primeira camada (ortografia) = transcrição do enunciado; segunda camada (aa) = análise acústica da oclusiva; terceira camada (*burst*) = explosão

a sublinhar os momentos de disfluência na lista. O volume das gravações foi ajustado individualmente, de acordo com o nível de conforto de cada ouvinte e a gravação das sentenças foi repetida várias vezes, para que os avaliadores pudessem ter certeza sobre a rotulagem. Somente marcas identificadas por todos os três juízes e as pesquisadoras foram incluídas na análise.

A análise estatística foi desenvolvida separando-se os grupos (PCG e PSG) e os contextos de fala da seguinte forma: (a) comparação entre as consoantes; (b) as consoantes na fala de PCG, em comparação com PSG; (c) contextos de fala comparados uns com os outros. Medidas descritivas e inferenciais foram calculadas. Os testes de Friedman e Wilcoxon foram utilizados para comparação dentro dos grupos e o teste de Mann-Whitney, nas comparações intergrupos. As análises estatísticas foram realizadas utilizando o *software* SPSS 14.0, com nível de significância de 5%.

RESULTADOS

O efeito do contexto de fala sobre as consoantes oclusivas e suas explosões foi marcante. Como descrito anteriormente, a análise do GE compreendeu três categorias: fluência, pré-disfluências e pós-disfluências. Os efeitos do contexto de fala em cada consoante, tanto para PCG quanto para PSG, estão demonstrados nas Tabelas 1 e 2.

Para o GE, as medidas acústicas de intervalo de oclusão e VOT diferenciaram as consoantes não vozeadas em todos os contextos de fala. No entanto, as consoantes não se diferiram pela medida de duração total na fala pós-disfluência. No grupo controle, as consoantes se diferiram quanto ao intervalo de oclusão e ao VOT.

A análise dos efeitos do contexto de fala (GC, GE/fluente,

GE/pré-disfluência e GE/pós-disfluência) mostrou que a maioria dos valores de p demonstrou diferença significativa. Na análise intergrupo (GC *versus* ambientes de fala de GE), o VOT das pessoas que gaguejam foi maior em cinco das nove comparações e foi menor para /t/ na fala fluente e /p/ na pré-disfluência. As medidas acústicas de /p/ na fala fluente não mostraram diferença entre os grupos. Os tempos de oclusão das consoantes oclusivas não vozeadas foram mais longos para GE, em quatro das nove comparações e mais curto para /p/ e /k/ na pós-disfluência. Os valores das consoantes foram mais elevados em cinco das nove comparações. No entanto, as durações de /p/ e /k/ foram menores em contextos pós-disfluências (Tabela 3).

A Tabela 3 também mostra uma comparação de contextos de fala para GE. Para esses falantes, o VOT é, geralmente, mais longo na fala pós-disfluência, enquanto o tempo de oclusão e a duração total são mais longos na fala fluente. Na fala pré-disfluência, todas as medidas acústicas das consoantes oclusivas não vozeadas apresentaram valores mais elevados do que na fala fluente. Os segmentos retirados dos ambientes pré-disfluente geralmente apresentaram valores mais elevados do que aqueles no discurso pós-disfluente, com exceção de VOT para /p/, que é mais curto em ambientes pré-disfluente.

DISCUSSÃO

Este estudo verificou a influência do ambiente de fala (fluente, pré-disfluência e pós-disfluência) na produção de consoantes oclusivas não vozeadas por indivíduos com gagueira e falantes do Português Brasileiro. Na fala pré-disfluente, o GE apresentou medidas superiores a GC em todas as medidas realizadas para cada consoante. Podem-se levantar três possíveis hipóteses que justificariam tal diferença, conforme segue.

Tabela 1. Medidas descritivas das oclusivas não vozeadas produzidas por pessoas com gagueira, em todos os três ambientes experimentais e no grupo controle

Contexto de fala	Medidas	Número	Varição	Média	Mediana
/p/ Fluente	Tempo de oclusão	31	0,06 – 0,99	0,13	0,11
	VOT	31	0,01 – 0,03	0,02	0,02
	Duração total	31	0,08 – 1,02	0,15	0,13
/p/ Pré-disfluência	Tempo de oclusão	33	0,27 – 0,30	0,28	0,28
	VOT	33	0,01 – 0,02	0,01	0,01
	Duração total	33	0,28 – 0,32	0,29	0,29
/p/ Pós-disfluência	Tempo de oclusão	38	0,07 – 0,08	0,07	0,07
	VOT	38	0,03 – 0,05	0,04	0,04
	Duração total	38	0,10 – 0,12	0,11	0,11
/t/ Fluente	Tempo de oclusão	45	0,02 – 0,13	0,08	0,08
	VOT	45	0,01 – 0,03	0,02	0,02
	Duração total	45	0,03 – 0,14	0,09	0,1
/t/ Pré-disfluência	Tempo de oclusão	27	0,09 – 0,30	0,13	0,11
	VOT	27	0,01 – 0,04	0,03	0,04
	Duração total	27	0,12 – 0,31	0,17	0,14
/t/ Pós-disfluência	Tempo de oclusão	37	0,07 – 0,09	0,08	0,08
	VOT	37	0,02 – 0,04	0,03	0,03
	Duração total	37	0,10 – 0,12	0,11	0,11
/k/ Fluente	Tempo de oclusão	29	0,10 – 0,99	0,12	0,1
	VOT	29	0,03 – 0,05	0,04	0,05
	Duração total	29	0,14 – 1,04	0,21	0,15
/k/ Pré-disfluência	Tempo de oclusão	28	0,09 – 0,33	0,31	0,32
	VOT	28	0,03 – 0,08	0,08	0,08
	Duração total	28	0,12 – 0,40	0,38	0,39
/k/ Pós-disfluência	Tempo de oclusão	58	0,04 – 0,08	0,06	0,06
	VOT	58	0,01 – 0,08	0,05	0,06
	Duração total	58	0,07 – 0,15	0,11	0,12
/p/ PSG	Tempo de oclusão	20	0,01 – 0,17	0,1	0,11
	VOT	20	0,01 – 0,04	0,02	0,02
	Duração total	20	0,02 – 0,21	0,12	0,13
/t/ PSG	Tempo de oclusão	30	0,03 – 0,13	0,07	0,07
	VOT	30	0,02 – 0,05	0,03	0,03
	Duração total	30	0,05 – 0,16	0,1	0,1
/k/ PSG	Tempo de oclusão	50	0,01 – 0,13	0,08	0,08
	VOT	50	0,03 – 0,07	0,05	0,05
	Duração total	50	0,05 – 0,17	0,13	0,13

Legenda: VOT = *voice onset time*; PSG = fala de pessoas sem gagueira

Diferentes teorias identificam a formulação da linguagem como fator primário para produção das disfluências da fala, tanto em falantes fluentes, quanto em falantes gogos. Uma das teorias afirma que as disfluências são consequências dos erros detectados durante a preparação do plano fonético⁽²⁷⁾. Os resultados encontrados na pré-disfluência podem refletir um ajuste dos gestos articulatorios na identificação do erro no planejamento do plano fonético, confirmando tais ideias.

Já em termos neuromotores, a articulação mais lenta em pessoas com gagueira pode ser interpretada como uma disfunção no processamento da fala⁽³⁾. A hipótese de que a gagueira resulta de uma disfunção nos gânglios da base reforça a ideia de que o distúrbio é diretamente relacionado ao tempo de produção da fala^(11,12). Essa disfunção resulta de uma anormalidade estrutural que afeta o fluxo de informações entre a área de Broca e o córtex motor, ou seja, entre a programação do planejamento

Tabela 2. Comparação dos valores de p: intervalo de oclusão, VOT e duração total das consoantes para cada contexto de fala

Ambiente	Medidas	Teste Friedman	Valor de p	Teste Wilcoxon
PCG-Fluente	Intervalo de oclusão	18.000	<0,001	/t/</p/; /t/</k/
	VOT	43.630	<0,001	/t/</p/</k/
	Duração total	40.963	<0,001	/t/</p/</k/
PCG-Pré-disfluência	Intervalo de oclusão	44.308	<0,001	/t/</p/</k/
	VOT	52.000	<0,001	/p/</t/</k/
	Duração total	44.308	<0,001	/t/</p/</k/
PCG-Pós-disfluência	Intervalo de oclusão	34.108	<0,001	/k/</t/</p/
	VOT	48.054	<0,001	/t/</p/</k/
	Duração total	1.552	0,460	-
PSG	Intervalo de oclusão	9.300	0,010	/t/</p/
	VOT	25.848	<0,001	/p/</t/</k/
	Duração total	5.200	0,074	-

Legenda: PCG-Fluente = fala fluente de pessoas com gagueira; PCG-Pré-disfluência = fala pré-disfluências de pessoas com gagueira; PCG-Pós-disfluência = fala pós-disfluências de pessoas com gagueira; PSG = fala de pessoas sem gagueira; VOT = *voice onset time*

Tabela 3. Comparação dos valores de p: grupos e contextos de fala

Medidas	PSG	PSG	PSG	PCG fluente	PCG fluente	PCG pré-disf
	vs PCG-fluente	vs PCG-pré-disf	vs PCG pós-disf	vs PCG pré-disf	vs PCG pós-disf	vs PCG pós-disf
/p/	Intervalo de oclusão	0,787	<0,001	0,003	<0,001	<0,001
	VOT	0,512	<0,001	<0,001	<0,001	<0,001
	Duração total	0,689	<0,001	0,054	<0,001	<0,001
/t/	Intervalo de oclusão	0,348	0,001	0,016	<0,001	0,694
	VOT	<0,001	<0,001	0,118	<0,001	<0,001
	Duração total	0,299	<0,001	0,009	<0,001	0,004
/k/	Intervalo de oclusão	<0,001	<0,001	<0,001	0,003	<0,001
	VOT	0,046	<0,001	0,032	<0,001	0,098
	Duração total	<0,001	<0,001	<0,001	0,003	<0,001

Teste Mann-Whitney ($p=0,005$)

Legenda: PSG = fala de pessoas sem gagueira; PCG fluente = fala fluente de pessoas com gagueira; PCG pré-disf = fala pré-disfluências de pessoas com gagueira; PCG pós-disf = fala pós-disfluências de pessoas com gagueira; VOT = *voice onset time*; vs = *versus*

motor da fala e a execução do movimento⁽¹⁵⁾.

Uma terceira hipótese que justificaria os resultados é o efeito coarticulatório. O fenômeno da coarticulação tem sido tema central de estudos atuais sobre produção de fala e pode ser definido, de forma geral, como a sobreposição de sons durante a produção de fala. Isso quer dizer que a produção do som /b/ é diferente se produzido isoladamente, ou nas sílabas /bu/ ou /bi/, sendo que, na primeira produção há um arredondamento dos lábios, enquanto na segunda, há um estiramento. Esse é apenas um exemplo de coarticulação nos sons produzidos no Português Brasileiro.

É esperado, então, que a fala apresente alta variabilidade e que o sinal da fala seja de difícil segmentação, dada a característica de influência contínua e recíproca dos segmentos da fala, em grande parte universal⁽²⁸⁾. Ressalta-se que, em estudos sobre coarticulação na fala de indivíduos com gagueira, o efeito nessa população é idêntico ou similar aos fenômenos coarticulatórios

da fala de pessoas sem essa desordem de fala⁽²⁹⁾, ou seja, o efeito da coarticulação está presente na fala de pessoas com gagueira. Além disso, a coarticulação pode ser subdividida em dois grupos: esquerda-para-direita (LR) e direita-para-esquerda (RL). O primeiro grupo, LR, é chamado também de progressivo, pois determinadas propriedades de um segmento perseveram, a ponto de influenciar o segmento seguinte. O segundo grupo, RL, também chamado de regressivo, é relacionado aos efeitos antecipatórios, ou seja, um segmento influencia aqueles que vêm antes dele⁽²⁸⁾. O grupo RL, mais comum na fala espontânea, pode ser explicado pela intervenção cognitiva nos processos biológicos e biomecânicos⁽²⁸⁾.

A identificação articulatória é esperada na disfluência. No presente estudo, foi avaliado o efeito da disfluência na consoante das sílabas pré-disfluência e pós disfluência. Da mesma forma que a coarticulação regressiva é mais comum que a progressiva, os efeitos das disfluências foram muito mais

marcantes nas pré-disfluências do que nas pós-disfluências. Tal fato é evidenciado nas Tabelas 1, 2 e 3, onde se vê que o VOT do GE é maior que o do GC, porém, o tempo de oclusão tende a ser menor no grupo com gagueira. Desta forma, os efeitos das disfluências parecem ser bidirecionais, com predomínio nas pré-disfluências.

Com relação à análise da fala fluente do GE, foi observado que os fones comportam-se de maneiras diversas, confirmando os resultados de um estudo⁽²⁰⁾ que demonstrou maior variabilidade nas medidas de duração de oclusivas não vozeadas em falantes com gagueira. Entretanto, os resultados encontrados não concordam com os achados de outros estudos^(3,21), pois os resultados de duração do VOT na fala de pessoas com e sem gagueira foram iguais para o fonema bilabial aqui estudado, enquanto foram diferentes nos estudos acima citados. A diferença encontrada com relação a esses estudos pode ser explicada pelo tipo de ambiente de fala analisado. Aqui, a fala sem disfluências foi analisada em três momentos diferentes, enquanto outros estudos parecem considerar a fala fluente como um todo^(3,21).

É importante ressaltar que na produção fluente do /k/ o tempo de oclusão do GE é muito maior do que o de falantes fluentes (GC), compensando o VOT reduzido na determinação da duração total do fone. Isso confirma a grande variabilidade no controle motor de indivíduos com gagueira. Essa grande variabilidade, evidenciada pelas medidas, pode reafirmar a instabilidade do controle motor de indivíduos com gagueira^(11,12).

Muitos estudos têm mostrado que as mudanças na duração do VOT seguem uma ordem hierárquica para as oclusivas: velar > alveolar > labial^(1,2). Neste estudo, esperava-se que os grupos de PCG e PSG iriam manter essa tendência, embora com valores mais elevados para o primeiro grupo. Para PSG, essa relação continuou a ser clara, com diferenças significativas entre as consoantes. No entanto, os dados para PCG não confirmaram essa hipótese. Esses resultados, portanto, concordam com os de uma pesquisa da organização temporal em distúrbios da fala, nos quais a duração das consoantes produzidas por pessoas que gaguejam não segue as tendências típicas determinadas pelo ponto e modo de articulação⁽³⁰⁾.

CONCLUSÃO

Na pré-disfluência, o grupo com gagueira apresentou duração superior em todas as medidas realizadas para cada oclusiva não vozeada.

Considerando que nenhum dos participantes do presente estudo fez terapia fonoaudiológica anteriormente, os resultados apontam para a necessidade de pesquisas futuras com o objetivo de investigar se a terapia fonoaudiológica pode afetar as características temporais das consoantes descritas aqui. Os dados foram interpretados a partir de uma abordagem conservadora e, de maneira geral, a gagueira dos participantes foi classificada como grave. Trabalhos futuros deverão considerar

diferentes níveis de gravidade da gagueira. Além disso, pode ser interessante a aplicação de metodologia semelhante em amostras de fala espontânea.

REFERÊNCIAS

- Hertrich I, Ackermann PH. Exploration of orofacial speech movements. In: Auzou P, Rolland V, Pinto S, Ozsancak C. Les dysarthries. Marseille: Solal; 2007. p. 132-9.
- Lousada ML. Estudo da produção de oclusivas do português europeu [dissertação]. Aveiro: Secção Autónoma de Ciências da Saúde; 2006.
- Sassi FC, Andrade FC. Acoustic analyses of speech naturalness: a comparison between two therapeutic approaches. *Pró Fono*. 2004;16(1):31-8.
- Pinto S. Bases anatomofisiologiques de l'articulation supralaryngée. In: Auzou P, Rolland V, Pinto S, Ozsancak C. Les dysarthries. Marseille: Solal; 2007. p 53-6.
- Zemlin WR. Princípios de anatomia e fisiologia em fonoaudiologia. 4a ed. Porto Alegre: Artes Médicas Sul; 2000.
- Abdelli-Beruh NB. Influence of place of articulation on some acoustic correlates of the stop voicing contrast in Parisian French. *J Phonetics*. 2009;37:66-78. <http://dx.doi.org/10.1016/j.wocn.2008.09.002>
- Lisker L, Abramson AS. A cross-language study of voicing in initial stops: acoustical measurements. *Word*. 1964;20(3):384-422.
- Giraud K, Trébuchon-DaFonseca A, Démonet JF, Habib M, Liégeois-Chauvel C. Asymmetry of voice onset time-processing in adult developmental dyslexics. *J Clin Neurophysiol*. 2008;119(7):1652-3. <http://dx.doi.org/10.1016/j.clinph.2008.02.017>
- Max L, Gracco VL. Coordination of oral and laryngeal movements in the perceptually fluent speech of adults who stutter. *J Speech Lang Hear Res*. 2005;48(3):524-42. [http://dx.doi.org/10.1044/1092-4388\(2005/036\)](http://dx.doi.org/10.1044/1092-4388(2005/036))
- Auzou P, Özşancak C, Morris RJ, Jan M, Eustache F, Hannequin D. Voice onset time in aphasia, apraxia of speech and dysarthria: a review. *Clin Ling Phonet*. 2000;14(2):131-50.
- Alm PA. Stuttering and the basal ganglia circuits: a critical review of possible relations. *J Commun Disord*. 2004;37(4):325-69. <http://dx.doi.org/10.1016/j.jcomdis.2004.03.001>
- Alm PA. Stuttering and sensory gating: a study of acoustic startle prepulse inhibition. *Brain Lang*. 2006;97(3):317-21. <http://dx.doi.org/10.1016/j.bandl.2005.12.001>
- Smits-Bandstra S, De Nil LF. Sequence skill learning in persons who stutter: implications for cortico-striato-thalamo-cortical dysfunction. *J Fluency Disord*. 2007;32(4):251-78. <http://dx.doi.org/10.1016/j.jfludis.2007.06.001>
- Arcuri CF, Osborn E, Schiefer AM, Chiari BM. Taxa de elocução de fala segundo a gravidade da gagueira. *Pró Fono*. 2009;21(1):45-50. <http://dx.doi.org/10.1590/S0104-56872009000100008>
- Giraud AL, Neumann K, Bachoud-Levi A, Gudenberg AW, Euler HA, Lanfermann H et al. Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. *Brain Lang*. 2008;104(2):190-9. <http://dx.doi.org/10.1016/j.bandl.2007.04.005>

16. Wu JC, Maguire G, Riley G, Fallon J, LaCasse L, Chin S et al. A positron emission tomography [18F] deoxyglucose study of developmental stuttering. *Neuroreport*. 1995;6(3):501-5. <http://dx.doi.org/10.1097/00001756-199502000-00024>
17. Rosenberger PB. Dopaminergic systems and speech fluency. *J Fluency Disord*. 1980;5(3):255-67. [http://dx.doi.org/10.1016/0094-730X\(80\)90031-5](http://dx.doi.org/10.1016/0094-730X(80)90031-5)
18. Andrade CRF, Cervone LM, Sassi FC. Relationship between the stuttering severity index and speech rate. *Sao Paulo Med J*. 2003;121(2):81-4. <http://dx.doi.org/10.1590/S1516-31802003000200010>
19. Arenas RM, Zebrowski PM, Moon JB. Phonetically governed voicing onset and offset in preschool children who stutter. *J Fluency Disord*. 2012;37(3):179-87. <http://dx.doi.org/10.1016/j.jfludis.2012.04.001>
20. Jäncke L. Variability and duration of voice onset time and phonation in stuttering and non-stuttering adults. *J Fluency Disord*. 1994;19(1):21-7. [http://dx.doi.org/10.1016/0094-730X\(94\)90012-4](http://dx.doi.org/10.1016/0094-730X(94)90012-4)
21. Hillman RE, Gilbert HR. Voice onset time for voiceless stop consonants in the fluent reading of stutterers and nonstutterers. *J Acoust Soc Amer*. 1977;61(2):610-66. <http://dx.doi.org/10.1121/1.381308>
22. Adams MR. Voice onsets and segment durations of normal speakers and beginning stutterers. *J Fluency Disord*. 1987;12(2):133-9. [http://dx.doi.org/10.1016/0094-730X\(87\)90019-2](http://dx.doi.org/10.1016/0094-730X(87)90019-2)
23. Arcuri CF, Osborn E, Schiefer AM, Chiari BM. Duração do onset vocálico da fala fluente de gagos. *Rev Soc Bras Fonoaudiol*. 2010;15(1):108-14. <http://dx.doi.org/10.1590/S1516-80342010000100018>
24. Celeste LC, Teixeira EG P. Efeito do sexo e idade na produção do VOT. *Rev Letras*. 2009;2(1):28-39.
25. Yairi E, Ambrose NG. Early childhood stuttering I: persistency and recovery rates. *J Speech Lang Hear Res*. 1999;42(5):1097-112. <http://dx.doi.org/10.1044/jslhr.4205.1097>
26. Andrade CRF. Protocolo para avaliação da fluência da fala. *Pró Fono*. 2000;12(2):131-4.
27. Potmas A, Kolk H. The covert repair hypothesis: prearticulatory repair processes in normal and stuttered disfluências. *J Speech Hear Res*. 1993;3:472-87.
28. Tatham M, Morton K. *Speech production and perception*. New York: Palgrave Macmillan; 2006.
29. Sussman HM, Byrd CT, Guitar B. The integrity of anticipatory coarticulation in fluent and non-fluent tokens of adults who stutter. *Clin Linguist Phon*. 2011;25(3):169-86. <http://dx.doi.org/10.3109/02699206.2010.517896>
30. Cardoso BAS, Reis C. Variables for the study of the temporal organization in speech disorders. In: *Proceedings of IV Speech Prosody*; 6-9 maio 2008; Campinas, Brasil. [Local, editor e data desconhecidos] p. 195-8.