



ANÁLISE ACÚSTICA DA FALA PARA DETECÇÃO DE FADIGA E SONOLÊNCIA EM AVIADOR

Vasconcelos, Carla Aparecida de¹; Vieira, Maurílio Nunes²; Kecklund, Göran³; Yehia, Hani Camille⁴.

- (1) Programa de Pós-Graduação em Neurociências da Universidade Federal de Minas Gerais. Avenida Antônio Carlos, 6627, Belo Horizonte - MG, Brasil. E-mail: carlavfono@yahoo.com.br.
- (2) CEFALA - Centro de Estudos da Fala, Acústica, Linguagem e Música da Universidade Federal de Minas Gerais. Avenida Antônio Carlos, 6627, Belo Horizonte - MG, Brasil. E-mail: maurilionunesv@cpdee.ufmg.br.
- (3) Stress Research Institute - Stockholm University and Department of Clinical Neuroscience - Karolinska Institute - Stockholm - Sweden. E-mail: goran.kecklund@su.se.
- (4) CEFALA - Centro de Estudos da Fala, Acústica, Linguagem e Música da Universidade Federal de Minas Gerais. Avenida Antônio Carlos, 6627, Belo Horizonte - MG, Brasil. E-mail: hani@cefala.org.

RESUMO

A fadiga mental e a sonolência são reconhecidas como causas determinantes de acidentes e incidentes relacionados ao erro humano na aviação. No Brasil, de acordo com o CENIPA (Centro de Investigação e Prevenção de Acidentes Aeronáuticos), o aumento de acidentes no modal aéreo entre 2006 e 2013 foi de 158%, atingindo uma taxa de 1 acidente a cada 2 dias, sendo que os fatores humanos estavam presentes em 90% dos casos. Este artigo descreve um estudo retrospectivo da comunicação oral entre piloto e torre de controle de tráfego aéreo em um caso de acidente e teve como objetivo a detecção da fadiga e sonolência por meio da análise de voz e fala. A análise detalhada do acidente realizada pelo CENIPA indicou que a sonolência e a fadiga provavelmente contribuíram para o acidente. Os dados da voz e da fala foram analisados em duas situações: dados registrados trinta e cinco horas antes do acidente aéreo (condição não sonolenta), que foram comparados com amostras coletadas cerca de 1 hora antes do acidente e também durante o acidente (condição de sono). As gravações foram analisadas acusticamente para a extração de medidas objetivas da organização temporal do discurso como: hesitações, pausas silenciosas, prolongamento de sílabas finais e taxa de articulação. Os resultados sugerem que os métodos adotados neste estudo são viáveis para a detecção de fadiga e sonolência por meio da análise de fala. A avaliação da fala é apontada como um método não invasivo promissor para a detecção de fadiga e sonolência em operações da vida real, como o transporte aéreo. No entanto, são necessários estudos de validação em grande escala para determinar se os resultados obtidos no presente estudo de caso podem ser replicados em outros indivíduos e para outras línguas.

Palavras-chave: Acústica da Fala, Fadiga Mental, Sonolência.

ABSTRACT

Mental fatigue and sleepiness are recognized as determining causes of accidents and incidents related to human error in aviation. In Brazil, according to CENIPA (Center for Investigation and Prevention of Aeronautical Accidents), the increase in accidents in the aerial modal between 2006 and 2013 was 158%, reaching a rate of 1 accident every 2 days, and human factors were present in 90% of these accidents. This paper describes a retrospective study of oral communication between pilot and air traffic control tower just before an accident and aimed to detect fatigue and sleepiness based on voice and speech analysis. The detailed analysis of the accident carried out by CENIPA indicated that sleepiness and fatigue probably contributed to the accident. Voice and speech data were analyzed in two situations: data recorded thirty-five hours before the air crash (non-sleepy condition), which were

compared with samples collected about 1 hour before the accident and also during the accident (sleep condition). The recordings were analyzed acoustically for the extraction of objective measures from the temporal organization of speech such as hesitations, silent pauses, prolongation of final syllables and articulation rate. The results suggest that the methods adopted in this study are feasible for the detection of fatigue and sleepiness through speech analysis. Speech evaluation has been suggested as a promising non-invasive method for detection of fatigue and sleepiness in real life operations such as aviation transport. However, large-scale validation studies are needed to determine whether the results obtained in the present case study can be replicated in other individuals and for other languages.

Keywords: Speech Acoustics, Mental Fatigue, Sleepiness.

1. INTRODUÇÃO

A fadiga pode, de uma perspectiva comportamental, ser definida como o estado de comprometimento do desempenho após um período de esforço mental ou físico [1]. Sonolência refere-se à transição entre vigília e sono, e é definida como a tendência a adormecer mesmo quando o indivíduo supostamente devia estar ativo, trabalhando, por exemplo. No entanto, tanto a fadiga quanto a sonolência caracterizam-se pela diminuição da capacidade de trabalho, da operação segura de máquinas, por falhas de atenção, lentidão cognitiva e problemas de memória [2-4].

Os fatores etiológicos mais comumente citados para fadiga são atividade física intensa, esforço prolongado ou excessivo, sono insuficiente, vigília prolongada e doenças físicas e emocionais [3, 5]. Além disso, vários fatores ambientais e relacionados ao trabalho em uma aeronave podem causar ou contribuir para a fadiga em pilotos. Estes incluem alta carga de trabalho mental, trabalho em turnos, espaço reduzido no *cockpit*, manobras diversificadas, forças de aceleração alternadas, fluxo de ar deficiente, baixa luminosidade, ruído de fundo e vibração contínuos. É provável que as consequências de fatores que contribuem para a fadiga e a sonolência sejam potencializadas quando o indivíduo está, ao mesmo tempo, realizando atividades críticas de segurança que exigem atenção e concentração sustentadas [6, 7].

Ressalta-se que a fadiga e a sonolência são frequentemente usadas como sinônimos, mas diferem porque a sonolência é mitigada pelo sono, enquanto a mitigação da fadiga exige repouso [8].

A avaliação da fala tem sido sugerida como um método promissor para detecção de fadiga e sonolência em operações da vida real, como a aviação. É um desafio obter uma avaliação confiável e precisa da fadiga e da sonolência dos pilotos, especialmente porque métodos invasivos não são opções viáveis. Vários estudos têm proposto que a análise acústica da voz e da fala pode ser um método adequado, tendo em vista que não é invasivo e não interfere na rotina de trabalho dos pilotos [5, 9, 10].

2. DESENVOLVIMENTO

Este estudo teve como objetivo verificar se havia ocorrência de variações na voz e na fala de um piloto que apresentava queixa de fadiga e sonolência. Para tanto, foram comparadas amostras de áudio do piloto coletadas no dia em que ele apresentava tais queixas (dia em que sofreu um acidente aeronáutico) com amostras coletadas em dias anteriores. O piloto era um homem de 45 anos, sem histórico de doenças preexistentes. O acidente ocorreu às 10 horas da

manhã e foi fatal. O tempo total de amostragem de áudio foi de aproximadamente 16 minutos para o dia de controle (antes do acidente) e 30 minutos para o dia da colisão.

Arquivos de áudio contendo diálogos do piloto: (i) em dias de folga (contexto privado), (ii) com o centro do plano de voo no dia anterior ao acidente aéreo (35 horas antes) e (iii) durante o voo do acidente aeronáutico foram recebidos do CENIPA para análise.

As análises acústicas foram realizadas com o PRAAT® versão 5.3.85, enquanto o MS-Excel e o MINITAB versão 16 foram utilizados para análise estatística. O objetivo foi comparar os mesmos parâmetros acústicos gravados durante o dia anterior ao acidente (condição controle) com aqueles gravados no dia do acidente (dia com alto nível de fadiga e sonolência). Assim, o desenho do estudo envolveu um caso-controle retrospectivo para uma comparação individual de voz e fala.

Considerando as características dos arquivos de áudio recebidos - como qualidade de gravação e razão sinal-ruído - somente parâmetros robustos, extraídos sob essas condições, foram analisados. A extração dos parâmetros habituais da fala, incluindo padrões de frequência fundamental e formantes, sofre deterioração nas condições de baixa razão sinal-ruído.

Neste estudo, a classificação linguística foi usada para identificar sílabas, enquanto a divisão silábica foi fonética. No português brasileiro, a estrutura mínima de uma sílaba é o núcleo, que pode ser composto exclusivamente por uma vogal. Um esboço consonantal isolado (que não configura coda) não compõe uma sílaba. As consoantes podem ocupar apenas as posições de ataque (início) ou coda.

A seguir será apresentada a classificação para sílaba.

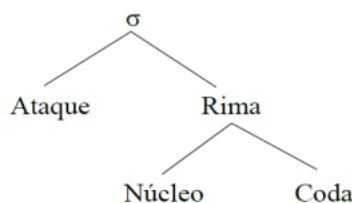


Figura 1 – Estrutura silábica universal. **Fonte:** Adaptado de SILVA [11].

2.1 Parâmetros analisados

A segmentação de todo o material de áudio do estudo de caso (relativo ao dia anterior ao acidente e ao dia do acidente) foi realizada manualmente, e os seguintes parâmetros de análise da organização temporal do discurso foram adotados [12]:

- Tempo Total de Articulação (TTA) - composto pelas sílabas produzidas;
- Tempo Total de Pausa (TTP) - composto por pausas silenciosas e preenchidas como: hesitação, repetição, falsos começos, dentre outros conforme a classificação de Duez [12];

- Tempo de Elocução (TE) - definido por $TE = TTA + TTP$;
- Número de Pausas e Duração Média das Pausas;
- Taxa de Elocução (TxE) - definida por NS / TE , onde NS = Número de Sílabas;
- Taxa de Articulação (TxA) - definida por NS / TTA ;
- Porcentagem de disfluência - definida por $ND / NS \times 100$, onde ND = Número de disfluências.

Neste estudo, as pausas preenchidas chamadas 'prolongamento' não foram subtraídas do número de sílabas fluentes produzidas, embora elas tenham sido consideradas no cálculo da disfluência. Isso resulta em uma diminuição na porcentagem de disfluência. Tal procedimento foi escolhido porque a subtração do que seria produção silábica normal do que já era prolongamento seria arbitrária.

Salienta-se que a disfluência oral é a ruptura do fluxo de fluência das emissões.

2.2 Métodos estatísticos

As análises estatísticas foram baseadas no teste *t* de *Student* pareado para variáveis com distribuição normal, com nível de significância de 5%. Para as variáveis não normalmente distribuídas, foi utilizado o teste pareado de *Wilcoxon*. Este teste, ao contrário do teste *t*, usa a mediana para análise de dados. O nível de significância de 5% também foi adotado.

Medidas de correlação também foram realizadas entre algumas variáveis. O coeficiente de correlação de *Pearson* foi aplicado para as medidas com distribuição paramétrica e o coeficiente de correlação de *Spearman* para casos não paramétricos.

Antes de escolher os testes estatísticos mais adequados para a comparação das amostras (arquivos de áudio extraídos antes e no dia do acidente), o teste de *Kolmogorov-Smirnov* foi realizado para verificar se as variáveis seguiam uma distribuição normal.

Especificamente para as variáveis de Taxa de Elocução e Taxa de Articulação, o valor de *p* foi maior que 0,05 e a hipótese nula de normalidade não pôde ser rejeitada. Devido ao fato de as amostras terem sido extraídas do mesmo informante, a dependência dessas também foi assumida e, portanto, foi necessário aplicar o teste *t* de *Student* para amostras pareadas.

3. RESULTADOS E DISCUSSÕES

Os principais resultados relacionados à organização temporal da fala são apresentados na Tabela 1 e na Figura 2. Eles mostram uma diminuição da taxa de elocução e da taxa de articulação no dia do acidente em comparação com o dia anterior ao acidente. Em comparação com os valores médios da fala espontânea em homens de mesma variação dialetal (Minas Gerais), que é, em média, 5,5 sílabas/s (taxa de elocução) e 6,0 sílabas/s (taxa de articulação), também houve decréscimo.

Houve também um aumento significativo na disfluência do falante no dia do acidente em relação aos dias anteriores, ou seja, nos dias de folga em casa e 35 horas antes do acidente de

avião nas mesmas condições de trabalho. Isto é adicionalmente evidenciado por um aumento das pausas silenciosas e das preenchidas.

Dois testes foram utilizados para a análise da correlação entre as variáveis: i) *Pearson*, para aqueles que seguiram uma distribuição normal (por exemplo, Tempo Total de Disfluência x Taxa de Elocução) e ii) *Spearman* para aqueles que apresentaram distribuição não paramétrica (por exemplo, Número de Pausas x Número de Sílabas).

Para a interpretação dos resultados utilizou-se a seguinte classificação [13]:

- $|r| \geq 0,9$: correlação muito forte;
- $0,7 \leq |r| < 0,9$: correlação forte;
- $0,5 \leq |r| < 0,7$: correlação moderada;
- $0,3 \leq |r| < 0,5$: correlação fraca;
- $0,0 \leq |r| < 0,3$: correlação insignificante.

Exemplos de análise da taxa de articulação e da taxa de elocução são apresentados a seguir.

Tabela 1: Média, desvio padrão e significância da comparação entre o dia anterior ao acidente com o dia do acidente.

Parâmetro (teste <i>t</i>)	Antes do acidente média (DP)	Dia do acidente média (DP)	Valor <i>p</i>
Taxa de Elocução (síl/seg)	5,2 (0,8)	3,8 (0,5)	0,01
Taxa de Articulação (síl/seg)	5,6 (0,3)	4,6 (0,3)	< 0,001

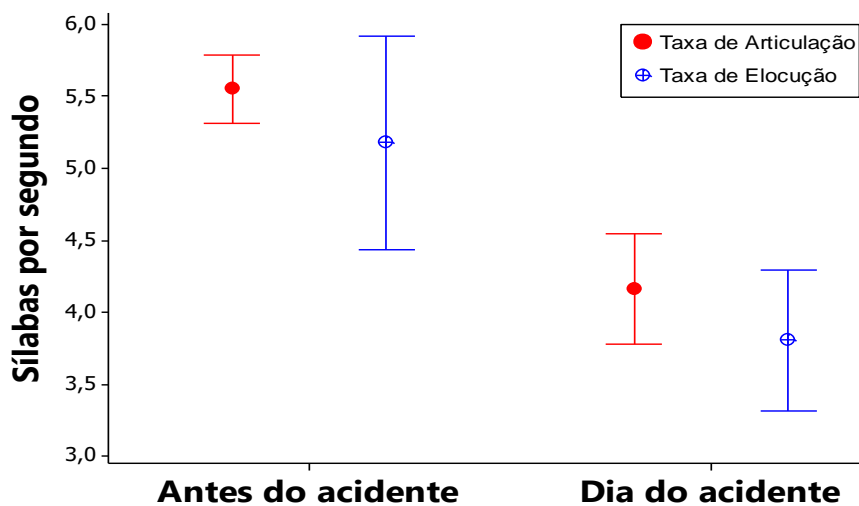


Figura 2: Representação dos intervalos de confiança estimados da Taxa de Articulação (em sílabas por segundo) e da Taxa de Elocução (em sílabas por segundo) para o dia anterior ao acidente e o dia do acidente (teste *t*).

Os resultados para a análise correlacional são apresentados na Tabela 2 e na Figura 3. As variáveis Taxa de Elocução e o Tempo Total de Disfluência mostraram uma forte correlação negativa ($r = -0,73$, $p < 0,001$). O coeficiente de determinação (r^2 , o quadrado do coeficiente de correlação de *Pearson*) é uma medida da proporção da variabilidade em uma variável que é explicada pela variabilidade da outra. Um $r = -0,73$ foi obtido, de modo que $r^2 = 0,54$, significando que 54% da Taxa de Elocução foi explicada pelo Tempo Total de Disfluência. Isso demonstra que não foi apenas a disfluência que interferiu na Taxa de Elocução, com 46% da variabilidade sendo determinada por outros fatores. A Taxa de Articulação encontrada foi menor no dia do acidente em comparação ao dia anterior ao acidente e isso também impactaria na Taxa de Elocução.

O Número de Pausas e Número de Sílabas mostrou um coeficiente de correlação de *Spearman* fraco ($x = 0,47$), indicando que o aumento no Número de Sílabas foi parcialmente associado com um aumento no Número de Pausas neste estudo de caso. Isso nos permite concluir que as pausas não ocorreram devido ao tamanho do enunciado, mas devido a outros motivos e nos fez inferir que o aumento da disfluência é uma consequência marcante do estado de fadiga e sonolência neste indivíduo estudado.

O falante também apresentou, além das alterações de organização temporal do discurso e fluência, alterações vocais e do comportamento vocal como crepitação semelhante ao *vocal-fry*. Ressalta-se que a intensidade no vocal na crepitação é muito débil [14]. É importante considerar que quando os níveis de ruído de fundo são elevados, os falantes aumentam automaticamente a intensidade da voz, fato conhecido como Efeito *Lombard*. A intensidade da voz é primeiramente regulada pelo aumento da pressão subglótica. Quando ocorre aumento da intensidade vocal, geralmente, a frequência fundamental (F_0) também aumenta. Como consequência do aumento dos valores da intensidade e da frequência fundamental, há o declínio do *vocal-fry* [15]. O Efeito *Lombard* é uma resposta vocal involuntária e é o resultado de um *feedback* entre o sistema da produção vocal e a percepção auditiva que permite a correção no desempenho do discurso, ou seja, uma melhora na relação fala/ruído para que o indivíduo consiga ouvir a si mesmo. Outro dado importante a ser considerado é o controle executivo. Esse se refere à capacidade de se regular os processos perceptuais e motores para o comportamento. Os seres humanos, por meio do controle executivo, são capazes de adequar quase qualquer resposta para quase qualquer estímulo recebido, mesmo quando não há nem conexões inatas nem adquiridas entre o estímulo e a resposta. Para maiores informações, ver *Lombard* [16]; *Pearsons* [17]; *Södersten* [15].

É importante ressaltar que o ruído de fundo para o falante analisado era considerável tanto no dia anterior ao acidente quanto no dia do acidente sendo perceptível nas análises dos arquivos de áudio.

Considerando a resposta incompatível entre a emissão vocal do falante analisado nesse estudo e os achados da literatura científica - Efeito *Lombard* e Controle Executivo, é necessário considerar a possibilidade/hipótese de baixa atividade responsiva do falante para controle do *feedback* (produção vocal-percepção auditiva) no dia do acidente.

De acordo com a literatura científica, mudanças na organização temporal da fala, que a tornam mais lenta, são desencadeadas pelo estado de fadiga [5]. Essas alterações ocorrem na fadiga/sonolência, pois nessa situação ocorrem mudanças nos padrões de contração muscular e nos comandos neurológicos da fala [5].

Tabela 2: Coeficiente de correlação de *Pearson* e significância da comparação entre a Taxa de Elocução e o Tempo Total de Disfluência no dia do acidente. Coeficiente de correlação de *Spearman* e significância da comparação entre o Número de Pausas e o Número de Sílabas no dia do acidente.

Variável	r (<i>Pearson</i>)	p -valor	x (<i>Spearman</i>)	p -valor
Taxa de Elocução × TTD	-0,73	0,00	-	-
N. Pausas × N. Sílabas	-	-	0,47	< 0,001

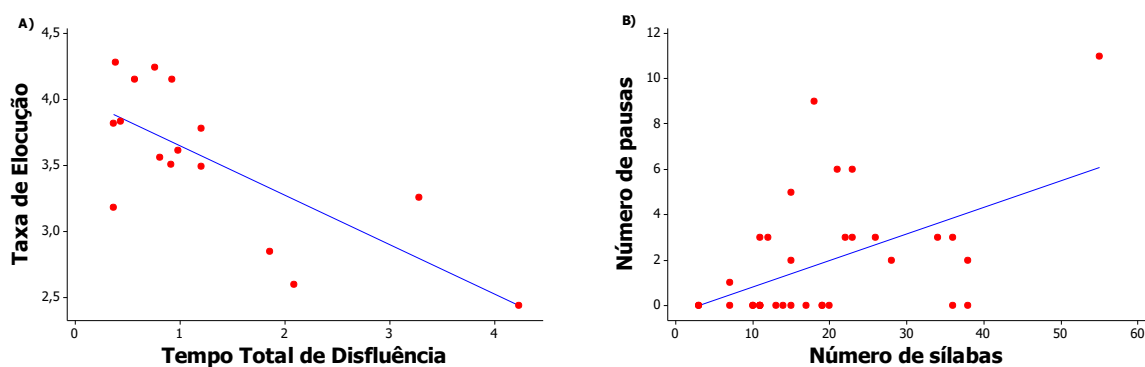


Figura 3: A) Representação da correlação entre as variáveis: Taxa de Elocução x Tempo Total de Disfluência. B) Representação da correlação entre as variáveis: Número de Pausas x Número de Sílabas.

4. CONCLUSÕES

Considerando-se os resultados encontrados neste estudo sobre as variações nos parâmetros fala do piloto no dia do acidente - como as mudanças na organização temporal da fala, mudanças no perfil de fluência de fala - e também variações nos parâmetros de voz como mudanças no comportamento vocal, ocorrência frequente de crepitação na fala semelhante ao *vocal-fry* – em comparação ao dia anterior ao acidente e aos padrões de normalidade - conclui-se que os tipos de medidas de voz e fala adotadas neste estudo são promissoras para a detecção da fadiga e da sonolência. Ressalta-se que a atividade de voo do piloto no dia do acidente ocorreu sem intercorrências e que a colisão se deu próxima ao momento de pouso, sendo que não houve situação de estresse nem variações vocais compatíveis com tal expressão emocional por parte do piloto.

Os resultados obtidos contêm variação significativa entre a fala produzida nas condições basais e pouco antes do acidente. Reconhecidamente, a informação extraída da fala é uma medida indireta de fadiga e sonolência. No entanto, outros métodos baseados em medidas fisiológicas relacionadas à avaliação dos movimentos oculares e atividade elétrica cerebral, além de serem indiretos, são invasivos, mais difíceis de realizar em situações reais com os pilotos e menos viáveis que o registro e a análise da fala.

O uso de gravações e análises de fala para detecção de fadiga e sonolência em pilotos de linhas aéreas é um método promissor e adequado, tendo em vista que situações de comunicação oral entre membros da tripulação e controladores de tráfego aéreo ocorrem com

frequência nas suas rotinas de trabalho. No entanto, estudos de validação em larga escala são necessários para determinar se os resultados obtidos no presente estudo de caso podem ser replicados sobre indivíduos e em outros idiomas.

AGRADECIMENTOS

Os autores agradecem ao Com. Paulo Licati, da Comissão Nacional de Fadiga Humana da ABRAPAC, por sugerir a investigação sobre o impacto da fadiga mental na fala dos pilotos; ao CENIPA por fornecer os dados para estudo; e a CAPES, CNPq e FAPEMIG pelo apoio financeiro.

REFERÊNCIAS

- [1] Shen, J.; Barbera, J. e Shapiro, C. M. Distinguishing sleepiness and fatigue: focus on definition and measurement. *Sleep Medicine Reviews*, v. 10, p. 63-76, 2006.
- [2] Dinges, D. F. An overview of sleepiness and accidents. *Journal of Sleep Research*, v. 4, p. 4-14, 1995.
- [3] Duc, B. D. e Li, X. Functional neuroimaging of circadian fatigue. *International Journal of Computer Applications in Technology*, v. 45, p. 156-163, 2012.
- [4] Fatigue Risk Management System (FRMS): DOC 9966. International Civil Aviation Organization (ICAO). Retrieved 12 January 2017 from <http://www.icao.int/safety/fatiguemanagement/Pages/DocumentsandToolkits.aspx>.
- [5] Vogel, A. P.; Fletcher, J. e Maruff, P. Acoustic analysis of the effects of sustained wakefulness on speech. *Journal of the Acoustical Society of America*, v. 128, p. 3747-3756, 2010.
- [6] Caldwell, J. A.; Mallis, M. M.; Caldwell, J. L.; Paul, M. A.; Miller, J. C. e Neri, D. F. Fatigue countermeasures in aviation. *Aviation, Space, and Environmental Medicine*, v. 80, p. 29-59, 2009.
- [7] Greeley, H. P.; Berg, J.; Friets, E.; Wilson, J.; Greenough, G.; Picone, J.; Whitmore, J. e Nesthus, T. Fatigue estimation using voice analysis. *Behavior Research Methods*, v. 39, p. 610-619, 2007.
- [8] Jepsen, J. R.; Zhao, Z. e Leeuwen, W. M. A. V. Seafarer fatigue: a review of risk factors, consequences for seafarers health and safety and options for mitigation. *International Maritime Health*, v. 66, p. 1-12, 2015.
- [9] Krajewski, J. e Kroger, B. Using prosodic and spectral characteristics for sleepiness detection. *Interspeech Proceedings*, v. 8, p. 1841-1844, 2007.
- [10] Shiomi, K.; Itano, K. e Suzuki, A. Development and evaluation of the fatigue and drowsiness predictor. *Archives of 28th International Congress of the Aeronautical Sciences*, p. 1-7, 2012.
- [11] Silva, T. C. Fonética e Fonologia do Português: roteiro de estudos e guia de exercícios. São Paulo: Contexto, 1998.
- [12] Duez, D. Organisation temporelle de la parole et dysarthrie parkinsonienne. In: Ozsancak, C.; Auzou, P. Les troubles de la parole et de la déglutition dans la maladie de Parkinson. Marseille: Solal, p. 195-211, 2005.
- [13] Mukaka, M. M. Statistics Corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, v. 24, p. 69-71, 2012.
- [14] Colton, R. H., Casper, J. K. Compreendendo os problemas de voz: uma perspectiva fisiológica ao diagnóstico e ao tratamento. Porto Alegre: Artes Médicas, 1996.

- [15] Södersten, M.; Ternström, S.; Bohman, M. Loud Speech in Realistic Environmental Noise: Phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *Journal of Voice*, vol. 19, no. 1, p. 29–46, 2005.
- [16] Lombard, E. Le signe de l'élévation de la voix. *Annales des Maladies de L'Oreille et du Larynx*. v. 27, n. 2, p. 101–9, 1911.
- [17] Pearsons, K.; Bennett, K.; Fidell, S. Speech levels in various noise environments. *U.S. Environmental Protection Agency Report EPA-600/1-77-025*. Springfield, Virginia, 1977.