



Softwares e hardwares de análise acústica da voz e da fala

Softwares and hardwares of acoustic analysis of voice and speech

Software y hardware de análisis acústico de la voz y el habla

Ana P. Mendes*

Lénia J.L. Ferreira**

Eugénia Castro***

Resumo

Este artigo consiste numa revisão sobre as funcionalidades de softwares e hardwares de análise acústica da voz e da fala. Numa primeira fase foi feita a pesquisa e seleção daqueles que o mercado disponibiliza e que são mais utilizados pelos profissionais de saúde destas áreas. Em seguida, reuniram-se dados sobre os parâmetros mais pertinentes, no que respeita à análise da voz e da fala, síntese da fala e intervenção terapêutica. Recolheram-se também informações sobre as especificidades técnicas de cada um. Numa última fase foram reportados os softwares e hardwares mais indicados para as práticas clínica, pedagógica e científica.

Palavras-chave: análise acústica da voz e da fala, software e hardware, reabilitação vocal, otimização vocal, habilitação vocal.

Abstract

This article consists in a review of the software and hardware features of acoustic analysis of voice and speech. Initially, it was reported the systems that are available in the market and the ones that are most commonly used by health care professionals in these areas. Then the most relevant parameters were gathered, for the analysis of voice and speech, speech synthesis and therapeutic intervention. It was also collected information about the technical specifications of each. Finally, the software and hardware that best suited pedagogical, clinical and scientific practices were reported.

Key-words: acoustic analysis of voice and speech, software and hardware, vocal rehabilitation, vocal optimization, vocal enabling.

** Terapeuta da Fala, Doutorada em Speech Science; Professora Adjunta da Escola Superior de Saúde do Instituto Politécnico de Setúbal e investigadora no Instituto de Engenharia Eletrónica e Telemática de Aveiro, Universidade de Aveiro. ** Terapeuta da Fala no Centro Helen Keller, Licenciada em Terapia da Fala. *** Assistente Hospitalar Graduada de Otorrinolaringologia e Coordenadora da Consulta da Voz do Serviço de Otorrinolaringologia do Centro Hospitalar de Vila Nova de Gaia/Espinho, E.P.E.; Coordenadora da Unidade da Voz do Hospital Cuf Porto.*





Resumen

Este artículo es una revisión de la funcionalidad del software y hardware acústico análisis del habla y voz. Inicialmente la investigación se llevó a cabo y seleccionar los que el mercado ofrece y que son utilizados principalmente por programas de salud profesionales en estas áreas. A continuación, los datos recogidos en los parámetros más relevantes en relación con el análisis del enfoque de voz y el habla terapéutica, síntesis del habla y intervención terapéutica. También se recopiló información sobre las características técnicas de cada uno. En la fase final se informó de software y hardware más adecuado para la práctica clínica, educativa y científica.

Palabras-claves: análisis acústico de la voz y el habla, software y hardware, rehabilitación vocal, optimización vocal, habilitación vocal.

Introdução

Ao longo do século XX foram vários os progressos da análise acústica do sinal da voz e da fala, através do avanço da tecnologia, desde o oscilograma obtido pelo oscilógrafo até ao espectrograma obtido pelo espectrógrafo de som, na década de 40. Na década de 60, Gunnar Fant¹ descreveu a Teoria Acústica de Produção da Fala. Esta teoria baseia-se num modelo que possibilita o cálculo das frequências de ressonância do trato vocal, através da localização da constrição feita pela língua, do grau de protrusão labial e do diâmetro transversal do trato vocal. Na década de 70, houve um avanço que se prolongou até aos dias de hoje: o processamento digital de sinais. O mesmo refere-se à análise e/ou alteração das características dos sinais da fala, utilizando técnicas computacionais. Possibilita operações como a disposição e edição da forma de onda, análise da frequência fundamental (F_0), *jitter*, *shimmer*, computação do índice sinal-ruído (S/R), obtenção do espectrograma digital, a sua computação utilizando o coeficiente de predição linear (*linear predictive code* (LPC)) e a síntese da fala^{1,2,3}.

A análise acústica por meio de *softwares* e *hardwares* tem vindo a tornar-se uma realidade frequente para os profissionais de saúde das áreas da voz e da fala. Esta apresenta não só utilidade na vocologia, mas também na área da fala (articulação e fluência), permitindo uma coleta de dados não invasiva, objetiva e precisa².

Considerando-se os avanços da análise acústica e o crescente interesse e necessidade de informação sobre a mesma, o presente estudo tem como objetivos: 1) divulgar os *softwares* e *hardwares* de análise acústica da voz e da fala mais conhecidos e

disponíveis no mercado; 2) fazer um levantamento das funcionalidades mais pertinentes que estes disponibilizam, no que se refere à análise da voz e da fala, síntese da fala, intervenção terapéutica e especificidades técnicas (requisitos de sistema, possibilidade de exportação de dados); e 3) divulgar aqueles que respondem às necessidades do profissional de saúde das áreas da voz e da fala nos contextos: clínico, pedagógico e científico. Este artigo destina-se aos profissionais de saúde que se dedicam à intervenção clínica, pedagógica e investigação da voz e da fala, como o terapeuta da fala/fonoaudiólogo, o otorrinolaringologista especializado em voz, o vocologista e o linguista. São estes os profissionais que manuseiam e contactam mais frequentemente com os *softwares* e *hardwares* e com as informações que estes fornecem. Consequentemente, demonstram maior preocupação em deter conhecimentos sobre os mesmos.

No contexto clínico, através da avaliação acústica da voz e da fala é possível inferir a fisiologia do aparelho fonador em geral e das pregas vocais (PV) em específico, tendo por base os resultados obtidos relativos aos vários parâmetros acústicos^{3,4}. Os mais pertinentes para os destinatários deste artigo são: 1) parâmetros espectrais: F_0 , variabilidade de F_0 , quebras de frequência e extensão máxima de frequência vocal (EMFV); 2) parâmetros de intensidade: intensidade objetiva vocal falada, variabilidade da amplitude, extensão dinâmica, índice de fonação suave (*soft phonation index* (SPI)); 3) parâmetros de perturbação: *jitter*, *shimmer*; 4) parâmetros acústicos espectrais de ruído: índice harmónico-ruído (H/R), índice sinal-ruído (S/R) e 5) parâmetros temporais: tempo máximo de fonação (TMF), índice S/Z, *voice-onset-time* (VOT), débito verbal/taxa de elocução, diadococinésia (ver Tabela 1).





Tabela 1 - Definição de parâmetros acústicos

	Parâmetros Acústicos	Siglas	Definição
Espectrais	Extensão máxima de frequência vocal	EMFV	Determina o intervalo de frequências que o utente pode produzir, desde a frequência mais grave do registo modal até à mais aguda do registo falsete. Também designada de extensão fonatória. Para a sua medição usam-se as vogais /a/ e /e/ ⁵ .
	Frequência fundamental vocal	F ₀	Som produzido pela vibração periódica das PV por unidade de tempo. O sexo masculino apresenta uma F ₀ entre 100 e 150 Hz e o sexo feminino entre 180 e 250 Hz ^{4,5,6,7} .
	Quebras de Frequência		Modificações abruptas e descontroladas da frequência da voz ⁴ .
	Variabilidade de F ₀		Desvio padrão da F ₀ . Oscila entre 2 e 4 semitons (ST) para o sexo masculino e feminino durante o discurso ^{4,5} .
Intensidade	Extensão dinâmica		Variação da intensidade, desde a mais fraca à mais forte produção. Obtém-se durante a produção de vogais sustentadas. É dependente do sexo, da idade, da frequência vocal e dos procedimentos de medição. Valores mínimos de 50 dB e máximos de 115dB ⁵ .
	Índice de fonação suave	SPI	Rácio do nível de energia média entre as baixas e as altas frequências. Relaciona-se com o grau de encerramento das PV, pelo que valores elevados deste parâmetro relacionam-se com a incompleta adução das PV ⁸ .
	Intensidade objetiva vocal falada		Correlato físico da intensidade subjetiva (<i>loudness</i>), que varia em função da pressão subglotal e da amplitude de vibração das PV. Mede-se em níveis de intensidade (<i>intensity level</i> , IL) ou em níveis de pressão sonora (<i>sound pressure level</i> , SPL), cujas unidades são os decibéis (dB). A intensidade vocal média falada em adultos aquando da leitura em voz alta é de 70 dB SPL ^{2,5} .
	Variabilidade da amplitude		Refere-se ao desvio-padrão da amplitude durante a fala ou a leitura em voz alta de adultos. A variabilidade da amplitude durante a fala ronda os 5,4 dB SPL ⁵ .
Perturbação	Jitter		Variação involuntária da frequência fundamental. Os valores médios de Jitter para vogais sustentadas no registo modal são de 1% ⁵ .
	Shimmer		Variação involuntária da intensidade objetiva. No registo modal, os dados normativos de Shimmer para vogais sustentadas são de 0,5% ⁵ .
Ruído	Índice sinal - ruído	S/R	Medida que compara o sinal acústico da voz com o ruído de fundo, permitindo extrair a energia aperiódica na voz (componente ruído) ao sinal acústico ^{2,4} .
	Índice harmónico - ruído	H/R	Obtido através da divisão da amplitude média da onda sonora pela amplitude média do ruído. Os valores médios desta medida rodam os 11.9 dB, com um desvio padrão de 2.32 dB ⁵ .
Temporais	Débito verbal/ taxa de elocução		Número de unidades linguísticas por unidade de tempo. Frequentemente é feita a contagem do número de palavras que um sujeito produz num minuto ⁹ .
	Diadococinésia		Capacidade de realizar movimentos rápidos, sequenciais ou alternados (por exemplo, pedindo ao utente para produzir a sequência /tatata/ ou /pataka/ ¹⁰ na diadococinésia da fala; ou /a/ /a/ /a/ para a diadococinésia laríngea ²).
	Índice S/Z		Rácio entre a duração máxima de /s/ e a duração máxima de /z/. O índice S/Z é igual a 1, segundo dados normativos ⁵ .
	Tempo máximo de fonação	TMF	Tempo máximo que o utente consegue sustentar um som numa respiração. Indica a coordenação entre os sistemas respiratório e fonatório. Em média adultos sem patologia do sexo masculino e feminino apresentam um TMF de 20 segundos e 15 segundos, respetivamente. Uma criança apresenta um TMF de 10 segundos ⁵ .
	Voice-onset-time	VOT	Intervalo de tempo entre a produção de uma consoante oclusiva e o início da vogal que se segue. O valor desta medida é positivo se a produção da consoante se inicia primeiro que a produção da vogal e negativo se o contrário suceder ⁷ .





A análise acústica constitui uma ferramenta complementar de avaliação da função vocal, que engloba também a endoscopia (laringoscopia, estroboscopia), glotografia, eletrolaringografia, eletromiografia e filtragem inversa. Deve ainda realizar-se uma avaliação áudio-percetual da voz, mioesquelética, da motricidade orofacial e laríngea, assim como exames audiométricos (i.e., audiometria, timpanometria). Todas as avaliações podem ser fundamentais para uma posterior intervenção adequada, no que respeita às necessidades do utente².

A fase de intervenção compreende não só a reabilitação, como também a otimização e habilitação das capacidades vocais e da fala. Estes *softwares* e *hardwares* incorporam ferramentas úteis para a modificação das características vocais do utente. Permitem ainda verificar a eficácia e a eficiência da intervenção terapêutica, através da comparação de resultados obtidos nas várias etapas da terapia.

Em contexto pedagógico, os *softwares* e *hardwares* constituem equipamentos usados nos métodos de ensino ativo, especificamente no ensino superior. Permitem a aquisição de competências de avaliação vocal acústica e a compreensão das características da voz do utente. Alguns dos *softwares* e *hardwares* fornecem também feedback auditivo e visual, em tempo real ou tardio, das medições vocais do utente. Assim, estes instrumentos podem ser utilizados na intervenção terapêutica e pedagógica, pois aumentam a sua consciencialização vocal².

Na investigação, a análise acústica do sinal da voz e da fala contribui para o desenvolvimento e aprofundamento de conhecimentos nas áreas da voz falada, cantada, laríngea, alaríngea, transexual e na área da fala. Uma das vantagens consiste na identificação de padrões associados com patologias, que podem ser aplicados na intervenção clínica, na medida em que ajudam a perceber como auxiliar na otimização, habilitação ou reabilitação da voz e da fala^{4,5}.

Em suma, os *softwares* e *hardwares* de análise acústica da voz e da fala providenciam feedback visual e auditivo, fornecem parâmetros quantitativos, objetivos, válidos e fiáveis, além de que melhoram a perceção vocal e da fala do utente. Deste modo, podem ser utilizados na prevenção, avaliação, diagnóstico e intervenção na patologia da voz e da fala, assim como na otimização das competências vocais e da fala. Têm ainda um papel na pedagogia e investigação.

Funcionalidades dos *softwares* e *hardwares*

Para este trabalho foi feita a revisão de dezessete *softwares* e *hardwares*, sete gratuitos e dez pagos (ver Tabela 2). A escolha destes *softwares* e *hardwares* teve por base uma revisão bibliográfica de artigos de periódicos, livros e *websites* da área de terapia da fala/fonoaudiologia, de forma a verificar quais aqueles que são mais utilizados pelos investigadores e profissionais de saúde das áreas da voz e da fala na atualidade. Foram também selecionadas as funcionalidades mais utilizadas pelos mesmos, relativas à análise acústica da voz e da fala, síntese da fala, intervenção terapêutica e especificidades técnicas. Quanto à análise da voz e da fala, foi feita uma pesquisa dos *softwares* e *hardwares* que permitem: 1) análise dos parâmetros acústicos espectrais, de intensidade, perturbação, ruído e temporais referidos na Tabela 1; 2) obtenção da forma de onda; 3) análise espectral (espectros e espectrogramas), de formantes e obtenção do LPC; 4) obtenção do primeiro e segundo formantes das vogais (F1 e F2, respetivamente); 5) filtragem do som; 6) transcrição fonética e ortográfica; 7) feedback visual para o utente. Foram ainda recolhidas informações sobre os *softwares* e *hardwares* que permitem a síntese da fala (i.e., capacidade de produção de fala partindo da determinação de parâmetros acústicos, formantes ou do desenho de espectrogramas^{3,11,12}) (ver Tabela 3).

As funcionalidades para a intervenção terapêutica consistiram em: 1) fornecimento de feedback visual e auditivo; 2) atividades para a modulação dos parâmetros acústicos F_0 (quando a modulação de F_0 pode ocorrer no domínio temporal e das frequências, intitula-se de *Frequency Altered Feedback* - FAF¹³), intensidade objetiva vocal falada, modulação da duração de segmentos da fala; 3) obtenção do *Delayed Auditory Feedback* (DAF) (i.e., atraso no feedback auditivo fornecido ao utente, sabendo a sua própria previsão¹³); 4) treino do VOT, início da fonação suave e produção de ataques vocais bruscos; 5) treino da articulação de fonemas; 6) obtenção de F_0 em semitons (ST) (a sua importância prende-se com a facilidade por parte do cantor em compreender como atingir uma determinada frequência vocal, já que se encontra geralmente mais familiarizado com os ST) (ver Tabela 3).





Tabela 2 – Softwares e hardwares de análise acústica da voz e da fala

Softwares		Hardwares
Gratuitos	Pagos	Pagos
<ul style="list-style-type: none"> • Audacity 2.0.0 • EMU Speech Database System 2.3.0 • WaveSurfer 1.8.5 • Praat 5.3.04 • Speech Filing System (SFS) 4.8 • SFS WASP 1.51 SIL International • Speech Analyser 3.0.1 	<ul style="list-style-type: none"> Dr. Speech, versão 4 • Vocal Assessment • Real Analysis • Speech Trainig • ScopeView • Phonetogram • Speech Therapy 4 FonoTools KayPENTAX • Multi-Speech, Modelo 3700 • Voice Range Profile (VPR), Modelo 4326 • Multi-Dimensional Voice Program (MDVP), Modelo 5105 • Motor Speech Profile (MSP), Modelo 5141 LingWAVES Voice Clinic Suite Pro Seegnal • MasterPitch Pro • VoiceStudio • SingingStudio Estill Voice International • VoicePrint • Estill Voiceprint Plus Time Frequency Analysis Software - TF32 Video Voice Speech Training System 3.0 Vox Metria Vocalgrama 	<ul style="list-style-type: none"> KayPENTAX • Computer Speech Lab (CSL), Modelo 4150B • Visi-Pitch IV, Modelo 3950B

Tabela 3 – Funcionalidades dos softwares e hardwares de análise acústica da voz e da fala

Funcionalidades	
Análise da voz e da fala	Obtenção de parâmetros acústicos Obtenção da forma de onda Análise espectral e de formantes Obtenção do LPC Filtragem do som Obtenção de F1 e F2 das vogais Transcrição fonética e ortográfica Feedback visual para o utente
Síntese da fala	
Intervenção terapêutica	Feedback visual/auditivo para o utente Modulação da intensidade, F_0 , duração dos segmentos da fala <i>Delayed auditory feedback</i> DAF Treino do VOT e início da fonação: ataques bruscos ou fonação suave Treino da articulação Obtenção de F_0 em ST
Especificidades técnicas	Capacidade de gravação áudio/vídeo Requisitos do sistema (Windows, Mac Os e/ ou Linux) Exportação de dados para WAV e MP3





As especificidades técnicas referem-se a: 1) gravação áudio da produção vocal do utente; 2) gravação vídeo da laringe, com auxílio de tecnologias como a endoscopia, videocassetes, sistemas digitais de captação e registo de imagem, câmaras de TV, entre outros; 3) requisitos de sistema dos *softwares* e *hardwares* (i.e., quais os sistemas operativos solicitados: Windows, Mac Os X, Linux); e 4) exportação de dados para os formatos mais conhecidos (i.e., WAV, MP3) (ver Tabela 3).

Na análise acústica e no que se refere aos parâmetros espectrais, todos os *softwares* e *hardwares* fornecem dados relativos à F_0 . O Praat 5.3.04¹⁴ e o MDVP¹⁵ mensuram as quebras de frequência. Ambos, em conjunto com o Multi-Speech¹⁶, MSP¹⁷, CSL¹⁸, Vocal Assessment¹⁹ e Vox Metria²⁰ medem a variabilidade de F_0 . O Vocalgrama²¹, LingWAVES Voice Clinic Suite Pro²², Phonetogram²³, SingingStudio²⁴, VRP²⁵, MDVP¹⁵ permitem a obtenção da EMFV.

Quanto aos parâmetros de intensidade, a intensidade objetiva vocal falada é fornecida por todos os *softwares* e *hardwares*. A variabilidade da amplitude é medida pelo Praat 5.3.04¹⁴, MDVP¹⁵, MSP¹⁷, CSL¹⁸ e LingWAVES Voice Clinic Suite Pro²⁶. A extensão dinâmica adquire-se através do CSL¹⁸, VRP²⁵, Phonetogram²³ e LingWAVES Voice Clinic Suite Pro²². O SPI é medido pelo MDVP¹⁵.

Os parâmetros acústicos de perturbação, *jitter* e *shimmer*, são obtidos pelo Praat 5.3.04¹⁴, MDVP¹⁵, CSL¹⁸, Vocal Assessment¹⁹, Vox Metria²⁰, SFS 4.8²⁷, VoiceStudio²⁸, TF32²⁹ e LingWAVES Voice Clinic Suite Pro²². O parâmetro de ruído espectral S/R é medido pelo TF32²⁹, Vocal Assessment¹⁹ e LingWAVES Voice Clinic Suite Pro³⁰. O índice H/R é fornecido pelo Praat 5.3.04¹⁴, MDVP¹⁵, CSL¹⁸, Vocal Assessment¹⁹ e VoiceStudio²⁸.

O parâmetro temporal débito verbal, fornecido pelo rácio silábico (i.e., número de sílabas por minuto), é medido pelo MSP¹⁷. O Video Voice Speech Training System 3.0³¹ também permite a obtenção do débito verbal. O VOT é medido pelo Praat 5.3.04¹⁴, Speech Therapy 4³² e LingWAVES Voice Clinic Suite Pro³³. A diadococinésia pode ser obtida utilizando-se o MSP¹⁷, o Visi-Pitch IV³⁴ e o LingWAVES Voice Clinic Suite Pro³⁰. Os *softwares* Speech Therapy 4³², Vocalgrama²¹, Vocal Assessment¹⁹, LingWAVES Voice Clinic Suite Pro²⁶ e o *hardware* CSL¹⁸ medem o TMF, por meio do qual se pode calcular o índice S/Z.

A forma de onda é facultada por todos os *softwares* e *hardwares*, com exceção do Phonetogram²³, VRP³⁵ e Speech Therapy 4³². A análise de espectros, espectrogramas e formantes são fornecidas pelo *hardware* e pela maioria dos *softwares*, excetuando o Phonetogram²³, SingingStudio²⁴, VRP³⁵, Vocalgrama²¹ e o MasterPitch Pro³⁶, que não proporciona o espectrograma. O LPC não é obtido pelo Phonetogram²³, VRP³⁵, Vocalgrama²¹, *softwares* da Seegnal Research^{24,28,36}, da Estill Voice International^{37,38}, Video Voice Speech Training System 3.0³¹, Visi-Pitch IV³⁴, Speech Therapy 4³², Audacity 2.0.0³⁹, SFS|WASP 1.51⁴⁰ e ScopeView⁴¹. A filtragem do som é realizada pelo Praat 5.3.04¹⁴, Multi-Speech¹⁶, CSL¹⁸, SFS 4.8²⁷, TF32²⁹, Audacity 2.0.0³⁹, LingWAVES Voice Clinic Suite Pro³³ e EMU Speech Database System 2.3.0⁴².

O F1 e F2 das vogais são fornecidos pelo Multi-Speech¹⁶, CSL¹⁸, Vox Metria²⁰, Speech Analyser 3.0.1⁴³, Real Analysis⁴⁴ e Speech Training⁴⁵. O Praat 5.3.04¹⁴, Multi-Speech¹⁶, CSL¹⁸, EMU Speech Database System 2.3.0⁴² e Speech Analyser 3.0.1⁴³ fazem transcrição fonética. A transcrição ortográfica é realizada com recurso ao Speech Analyser 3.0.1⁴³ e Wavesurfer 1.8.5⁴⁶.

O feedback visual do rácio silábico, da avaliação da diadococinésia, F_0 , variabilidade da amplitude e de F_0 é disponibilizado pelo MSP¹⁷. O feedback visual de F_0 também é fornecido pelo MDVP¹⁵. O Vocal Assessment¹⁹ dá feedback visual da variabilidade de F_0 , *jitter* e *shimmer*, o Vox Metria²⁰ dos parâmetros de perturbação e o VoiceStudio²⁸ destes e do H/R. O LingWaves Voice Clinic Suite Pro²⁶ permite obter feedback visual dos parâmetros *jitter*, *shimmer*, índice S/Z, TMF, F_0 , variabilidade de F_0 , intensidade e variabilidade da amplitude, índice S/R, avaliação da diadococinésia e dos pontos de articulação de sílabas. Estes *softwares* apresentam gráficos que demonstram se o utente apresenta alguma característica vocal ou da fala alterada, por intermédio de cores. Para tal, comparam os parâmetros da voz e da fala do utente com valores da base de dados do próprio sistema. A síntese da fala pode realizar-se através do Praat 5.3.04¹⁴, Multi-Speech¹⁶, CSL¹⁸, SFS 4.8²⁷ e Wavesurfer 1.8.5⁴⁶.

Quanto às funcionalidades para a intervenção terapêutica, o feedback auditivo é fornecido pelo Praat 5.3.04¹⁴, Multi-Speech¹⁶, SFS 4.8²⁷, MasterPitch Pro³⁶, Audacity 2.0.0³⁹, Wavesurfer 1.8.5⁴⁶, CSL¹⁸ e FonoTools⁴⁷. Estes *softwares*





permitem a modificação da onda sonora produzida pelo utente, proporcionando feedback auditivo da voz alterada e auxiliando na consciencialização da variação de determinados parâmetros acústicos. O Visi-Pitch IV³⁴ e o LingWaves Voice Clinic Suite Pro³³ comparam produções alvo (i.e., produzidas pelo profissional) com as produções do utente, através da reprodução áudio. Este último *software* compara ainda as produções do utente nos vários momentos da terapia. O feedback visual que os *softwares* MSP¹⁷, MDVP¹⁵, Vocal Assessment¹⁹, Vox Metria²⁰, VoiceStudio²⁸ e LingWaves Voice Clinic Suite Pro³³ fornecem também se torna importante na fase de intervenção, por mostrar ao utente a sua evolução através dos gráficos. O Visi-Pitch IV³⁴, Video Voice Speech Training System 3.0³¹, Speech Therapy 4³² e LingWaves Voice Clinic Suite Pro³³ fornecem atividades que ilustram as tentativas do utente para a produção correta de determinados parâmetros da voz e da fala, ajudando o utente a modificar estes parâmetros da sua voz.

O Multi-Speech¹⁶, CSL¹⁸, Vox Metria²⁰, Phonetogram²³, SingingStudio²⁴, VRP²⁵ e LingWaves Voice Clinic Suite Pro⁴⁸ permitem a obtenção de F_0 em ST e o Vocalgrama²¹ obtém a variabilidade de F_0 também em ST, pelo que estes *softwares* e *hardware* fornecem ao cantor feedback visual de F_0 . Os últimos cinco *softwares* fornecem feedback auditivo, na medida em que permitem a audição das notas musicais equivalentes à frequência de produção de determinado segmento da fala.

A modulação da intensidade objetiva vocal falada e de F_0 obtêm-se através do Praat 5.3.04¹⁴, SFS 4.8²⁷ (apenas manipula F_0), LingWaves Voice Clinic Suite Pro³³, Video Voice Speech Training System 3.0³¹, Speech Therapy 4³², Visi-Pitch IV³⁴, MasterPitch Pro³⁶ (que também fornece o FAF), Audacity 2.0.0³⁹, Wavesurfer 1.8.5⁴⁶ e FonoTools⁴⁷. A modulação da duração dos segmentos da fala obtêm-se através do Praat 5.3.04¹⁴, Video Voice Speech Training System 3.0³¹, Speech Therapy 4³² e Audacity 2.0.0³⁹. A modulação do TMF é obtida pelo Video Voice Speech Training System 3.0³¹ e Speech Therapy 4³². O *hardware* CSL¹⁸, os *softwares* Multi-Speech¹⁶, MasterPitch Pro³⁶, Audacity 2.0.0³⁹ e FonoTools⁴⁷ incorporam a possibilidade de obter o DAF.

O Speech Therapy 4³² e o LingWaves Voice Clinic Suite Pro³³ possibilitam a modulação do VOT. O primeiro *software*, o Video Voice Speech Training System 3.0³¹ e o Visi-Pitch IV³⁴ fornecem

atividades que permitem treinar a fonação vozeada, não vozeada, assim como o início da fonação confortável e ataques vocais bruscos. Apresentam também atividades para a produção articulatória de fonemas. O LingWaves Voice Clinic Suite Pro³³ possibilita o treino da fonação vozeada/ não vozeada e da articulação de sons, sílabas, palavras e frases.

Grande parte dos *softwares* pode operar nos sistemas operativos Windows, Mac OS e Linux. Os *softwares* e *hardwares* do Dr. Speech^{19,23,32,41,44,45}, KayPENTAX^{15-18,25,35}, Vox Metria²⁰, SFS 4.8²⁷, TF32²⁹, SFS|WASP 1.51⁴⁰, Speech Analyser 3.0.1⁴³ e FonoTools⁴⁷ exigem o sistema operativo *Windows*. O Video Voice Speech Training System 3.0³¹ e os *softwares* da Estill Voice International^{37,38} requerem o *Windows* ou *Mac OS*. O LingWaves Voice Clinic Suite Pro³³ requer o *Windows XP* ou *Vista*. O Vocalgrama²¹ apresenta como requisitos a utilização dos sistemas *Windows XP/Vista/Vista x64/7*.

A maioria dos *softwares* e *hardwares* possibilitam a captura e gravação áudio. O ScopeView⁴¹ possibilita a gravação de vídeo. A exportação de dados pode ser feita para MP3 e WAV, os formatos mais utilizados, através dos *softwares* Audacity 2.0.0³⁹ e Wavesurfer 1.8.5⁴⁶. O Praat 5.3.04¹⁴, SFS 4.8²⁷, *softwares* da Seegnal Research^{24,28,36}, TF32²⁹, *softwares* da Estill Voice International^{37,38}, SFS|WASP 1.51⁴⁰, EMU Speech Database System 2.3.0⁴² e Speech Analyser 3.0.1⁴³ permitem armazenamento de ficheiros áudio em formato WAV.

Contextos clínico, pedagógico e científico

Os *softwares* e *hardwares* de análise acústica da voz e da fala apresentam inúmeras funcionalidades, sendo que cada um foi concebido com um propósito específico: alguns são mais indicados para o contexto clínico, pedagógico, outros estão associados à investigação.

No contexto clínico distinguem-se as fases de avaliação e intervenção terapêutica. Todos os *softwares* e *hardwares* podem ser utilizados para recolher dados para a avaliação da voz e da fala, considerando-se alguns mais indicados para o efeito. Destaca-se o Praat 5.3.04, dado ser um *software* gratuito, que permite a obtenção de vários parâmetros acústicos e os principais gráficos dos mesmos. Os *softwares* MDVP, Multi-Speech,





Vox Metria, Real Analysis, Vocal Assessment e LingWAVES Voice Clinic Suite Pro também fornecem variadas informações sobre as características vocais do utente, embora sejam pagos. O CSL é um *hardware* que pode ser complementado com *softwares* da KayPENTAX, tais como o MDVP ou o VRP, pelo que fornece uma excelente variedade de parâmetros acústicos e de representações gráficas referentes à análise da voz do utente, embora tenha valores de aquisição elevados. Destaca-se ainda o LingWAVES Voice Clinic Suite Pro pelo feedback visual que fornece ao utente aquando da avaliação.

Para a intervenção terapêutica selecionam-se os *softwares* Video Voice Speech Training System 3.0, Speech Therapy 4, LingWAVES Voice Clinic Suite Pro, Visi-Pitch IV e MasterPitch Pro. Os três primeiros *softwares* apresentam atividades que auxiliam na otimização, habilitação e reabilitação da voz e da articulação. O MasterPitch Pro possibilita a obtenção do FAF e DAF, contribuindo para o trabalho ao nível da fluência do discurso. Um aspeto negativo na utilização destes *softwares* e *hardware* é o seu custo elevado. Quando não existe possibilidade financeira para adquirir um destes, considera-se que o Praat 5.3.04, Audacity 2.0.0 e Wavesurfer 1.8.5 são os *softwares* gratuitos mais adequados.

Os *softwares* Video Voice Speech Training System 3.0, MasterPitch Pro, Speech Terapy 4, MDVP, MSP, Vocal Assessment, Phonetogram, Vox Metria, VoiceStudio, SingingStudio, VRP, LingWAVES Voice Clinic Suite Pro, Vocalgrama e o *hardware* Visi-Pitch IV são aqueles que fornecem feedback ao utente aquando da intervenção terapêutica e pedagógica. Assim, apresentam um papel determinante, por permitirem a consciencialização das características da voz do utente e a modificação de parâmetros acústicos. O SingingStudio, Phonetogram, VRP, LingWAVES Voice Clinic Suite Pro e Vocalgrama têm a particularidade de converter Hertz em ST e fornecer representações com instrumentos musicais. Estes fazem a comparação direta da frequência vocal do utente com as notas musicais, auxiliando-o na deteção mais precisa dos limites da função vocal.

Na pedagogia e para o primeiro contacto dos profissionais e futuros profissionais de saúde das áreas da voz e da fala, considera-se que podem ser utilizados *softwares* e *hardwares* de análise acústica distintos nos níveis de dificuldade de utilização de: iniciação, exploração e aprofundamento.

Selecionam-se para a fase de iniciação os *softwares* Audacity 2.0.0, WaveSurfer 1.8.5, SFS 4.8 e SFS|WASP 1.51, por serem gratuitos e simples de manipular, por fornecerem os principais parâmetros acústicos e as mais básicas formas de análise da voz, a análise espectral e de formantes. Para as fases de exploração e aprofundamento considera-se mais adequado o *software* Praat 5.3.04. Este fornece um vasto número de funcionalidades, desde a deteção de parâmetros acústicos e fornecimento de representações gráficas dos mesmos a funcionalidades que podem ser utilizadas aquando da intervenção terapêutica. Salienta-se que estes *softwares* são gratuitos, facilitando aos estudantes a aquisição dos mesmos.

Para a investigação na área da voz e da fala considera-se mais adequado o uso da KayPENTAX e dos *softwares* da Seegnal Research, dado possuírem um maior número de funcionalidades. O Praat 5.3.04 continua a ser uma opção a considerar, dados os baixos custos inerentes à sua obtenção.

Conclusões

O presente artigo pretendeu dar uma contribuição na síntese, organização e comparação das características dos mais utilizados *softwares* e *hardwares* de análise acústica da voz e da fala. O objetivo principal foi auxiliar o profissional de saúde destas áreas na determinação daqueles que são mais adequados, de acordo com os objetivos do seu trabalho (i.e., clínico, pedagógico, científico).

Com o levantamento e caracterização destes *softwares* e *hardwares* foi possível compreender as vantagens inerentes à sua utilização: 1) análise objetiva da voz e da fala; 2) auxílio na intervenção terapêutica, nomeadamente, na otimização, habilitação e reabilitação da voz e da fala; 3) divulgação dos *softwares* e *hardwares* mais direcionados para a pedagogia e 4) divulgação dos *softwares* e *hardwares* mais direcionados para a investigação na área da voz e da fala.

Pretende-se, numa fase posterior, proceder à validação dos *softwares* e *hardwares* de análise acústica e ao estudo de fiabilidade dos mesmos.

Agradecimentos

As autoras deixam os seus agradecimentos especiais a Margarida Navarra e Nádía Simões, pela contribuição na coleta de informações sobre





os softwares e hardwares de análise acústica do sinal da voz e da fala.

Referências Bibliográficas

1. Silva AHP. O estatuto da análise acústica nos estudos fônicos. Cadernos de Letras da UFF – Dossiê: Letras e cognição [periódico na internet]. 2010 [acesso em janeiro 15, 2012]; 41: 213-29: [16 p.]. Disponível em: <http://www.uff.br/cadernosdeletrasuff/41/cotidiano1.pdf>.
2. Behlau M, Feijó D, Madazio G, Brasil OD, Pontes P, Azevedo R. Voz: o livro do especialista. Vol 1. Rio de Janeiro: Revinter; 2001.
3. Kent RD, Read C. Acoustic analysis of speech. 2nd ed. Canada: Thomson Learning; 2002.
4. Colton RH, Casper JK. Compreendendo os problemas de voz: uma perspectiva fisiológica ao diagnóstico e ao tratamento. Porto Alegre: Artes Médicas; 1996.
5. Mendes A, Castro E. Análise acústica da avaliação vocal I: Tarefas fonatórias e medidas acústicas. Rev Port Otorrinolaringol Cirurgia Cérvico-Facial. 2005; 43(2):127-36.
6. Guimarães ICRP. An electrolaryngographic study of dysphonic Portuguese speakers. [thesis submitted for the degree of doctor of Philosophy]. London: University of London; 2002.
7. Borden JG, Harris KS, Raphael LJ. Speech science primer: physiology, acoustics, and perception of speech. 4th ed. Baltimore: Lippincott Williams & Wilkins; 2003. p. 114-5.
8. Ferreira FV, Cielo CA, Trevisan ME. Medidas vocais acústicas na doença de parkinson: Estudo de Casos. CEFAC. [periódico na internet]. 2010 [acesso em janeiro 10, 2012]; 12 (5): 889-98 [9 p.].Disponível em: <http://www.scielo.br/pdf/rcefac/v12n5/10-09.pdf>.
9. Shipley GK, McAfee JG. Assessment in speech-language pathology: a resource manual. 3rd ed. San Diego: Singular Publishing Group; 2009.
10. Andrade, CR. Perfil familiar da fluência da fala - estudo linguístico, acústico e eletromiográfico. Pro Fono. [periódico na internet]. 2010 [acesso em janeiro 10, 2012]; 22 (3): 169-74 [5 p.]. Disponível em: <http://www.scielo.br/pdf/pfono/v22n3/a03v22n3.pdf>.
11. Jindřich IM. Acoustic speech synthesis [website]. University of West Bohemia in Pilsen, Czech Republic: Project MUSSLAP, Artificial intelligence section, Department of cybernetics. [acesso em janeiro 15, 2012]. Disponível em: <http://musslap.zcu.cz/en/acoustic-speech-synthesis/>.
12. Boersma P, Weenink D. Praat: doing phonetics by computer [homepage na internet]. Amsterdam: Institute of Phonetic Sciences, University of Amsterdam; c2011 [acesso em outubro 10, 2011]. Disponível em: <http://www.fon.hum.uva.nl/praat/>.
13. SEEGNAL Research - Tecnologias de Processamento de Sinal, Lda. Produtos [website]. S. Mamede de Infesta: SEEGNAL Research - Tecnologias de Processamento de Sinal, Lda. [acesso em janeiro 15, 2012]. Disponível em: <http://www.seegnal.pt/products/presentation.php>.
14. Boersma P, Weenink D. Praat: doing phonetics by computer versão 5.3.04 [programa de computador]. Amsterdam: Institute of Phonetic Sciences, University of Amsterdam; c 2011. [acesso em outubro 10, 2011]. Disponível em: <http://www.praat.org/>.
15. KayPENTAX. Multi-Dimensional Voice Program, Modelo 5105. [programa de computador]. Montvale, New Jersey: PENTAX Medical Company; c1996-2011.
16. KayPENTAX. Multi-Speech, Modelo 3700. [programa de computador]. Montvale, New Jersey: PENTAX Medical Company; c1996-2011.
17. KayPENTAX. Motor Speech Profile, Modelo 5141. [programa de computador]. Montvale, New Jersey: PENTAX Medical Company; c1996-2011.
18. KayPENTAX. Computerized Speech Lab, Modelo 4150B. [sistema de software e hardware]. Montvale, New Jersey: PENTAX Medical Company; c1996-2011.
19. Tiger DRS, Inc. Vocal Assessment. [programa de computador]. Seattle, WA: Tiger DRS, Inc.; c1998.
20. Behlau M. Vox Metria - software para análise de voz e qualidade vocal, versão 2.0. [programa de computador].Paraná, Brasil: CTS Informática, 2003.
21. Behlau M., Moraes, M. Vocalgrama. [versão de demonstração do programa de computador]. Paraná, Brasil: CTS Informática; [s.d.] [acesso em agosto 21, 2012]. Disponível em: http://www.profala.com/lj/product.php?id_product=29.
22. WEVOS – WEVOSYS. Lingwaves Voice Protocol. [website]. Forchheim, Germany: WEVOS – WEVOSYS; c2010. [acesso em agosto 21, 2012]. Disponível em: http://www.wevosys.com/products/lingwaves/lingwaves_voice_protocol.html.
23. Tiger DRS, Inc. Phonetogram. [website]. Seattle, WA: Tiger DRS, Inc.; 1998. [acesso em janeiro 15, 2012]. Disponível em: <http://www.drspeech.com/PhonetogramV4.html>.
24. Seegnal Research - Tecnologias de Processamento de Sinal, Lda. SingingStudio [versão de demonstração do programa de computador]. São Mamede de Infesta: Seegnal Research - Tecnologias de Processamento de Sinal, Lda; 2008 [acesso em janeiro 15, 2012]. Disponível em: <http://www.seegnal.pt/products/singingstudio.php>.
25. Camargo TF, Barbosa DA, Teles LCS. Características da fonetografia em coristas de diferentes classificações vocais. Rev Soc Bras Fonoaudiol [periódico na internet]. 2007 [acesso em março 14, 2012];12(1):10-7: [7 p.]. Disponível em: <http://www.scielo.br/pdf/rsbf/v12n1/02.pdf>.
26. WEVOS – WEVOSYS. Voice Protocol. [website]. Forchheim, Germany: WEVOS – WEVOSYS; c2010. [acesso em agosto 21, 2012]. Disponível em: http://wevosys.com/products/lingwaves/images/lingWAVES_Voice_Protocol_printout.pdf.
27. Huckvale M. Speech Filling System 4.8. [programa de computador]. London: University College London, Division of Psychology & Language Sciences; c2009 [atualizado em outubro, 2010; acesso em janeiro 25, 2012.]. Disponível em: <http://www.phon.ucl.ac.uk/resource/sfs/>.
28. Seegnal Research - Tecnologias de Processamento de Sinal, Lda. VoiceStudio [versão de demonstração do programa de computador]. São Mamede de Infesta: Seegnal Research - Tecnologias de Processamento de Sinal, Lda; 2008 [acesso em janeiro 15, 2012]. Disponível em: <http://www.seegnal.pt/products/voicestudio.php>.
29. Milenkovic PH. Time Frequency Analysis 32-bit for Windows – TF32 [versão de demonstração do programa de computador]. Wisconsin: University of Wisconsin-Madison, Department of Electrica and Computer Engineering; c2000 [atualizado em 2004; acesso em fevereiro 2, 2012]. Disponível em: <http://userpages.chorus.net/cspeech/>.
30. WEVOS – WEVOSYS. lingWAVES Motor Speech Disorder Assessment (MSDA). [website]. Forchheim, Germany: WEVOS – WEVOSYS; c2010. [acesso em agosto 21, 2012]. Disponível em: http://www.wevosys.com/products/lingwaves/lingWAVES_msda.html.





31. Micro Video Corporation. Video Voice Speech Training System version 3.0 [website]. Michigan: Micro Video Corporation; c1996-2012 [acesso em janeiro 15, 2012]. Disponível em: <http://www.videovoice.com/default.htm>.
32. Tiger DRS, Inc. Speech Therapy 4. [programa de computador]. Seattle, WA: Tiger DRS, Inc.; c1998.
33. WEVOS – WEVOSYS. lingWAVES – a modular system for voice and speech. [website]. Forchheim, Germany: WEVOS – WEVOSYS; c2010. [acesso em agosto 21, 2012]. Disponível em: http://www.orl.biofonia.com/rcs_gene/Folleto_LingWaves.pdf.
34. KayPENTAX. Visi-Pitch IV, Modelo 3950B. [Website]. Montvale, New Jersey: PENTAX Medical Company; c1996-2011 [acesso em fevereiro 2, 2012]. Disponível em: [http://www.kayelemetrics.com/index.php?option=com_product&Itemid=3&controller=product&task=learn_more&cid\[\]=85](http://www.kayelemetrics.com/index.php?option=com_product&Itemid=3&controller=product&task=learn_more&cid[]=85).
35. KayPENTAX. Voice Range Profile, Modelo 4326. [programa de computador]. Montvale, New Jersey: PENTAX Medical Company; c1996-2011.
36. Seegnal Research - Tecnologias de Processamento de Sinal, Lda. MasterPitch Pro [versão de demonstração do programa de computador]. São Mamede de Infesta: Seegnal Research - Tecnologias de Processamento de Sinal, Lda; 2008 [acesso em janeiro 15, 2012]. Disponível em: <http://www.seegnal.pt/products/masterpitch.php>.
37. Estill Voice International, LLC. Voiceprint versão 5.5. [versão de demonstração do programa de computador]. Pittsburgh, PA: Estill Voice International, LLC; c2010 [acesso em fevereiro 3, 2012]. Disponível em: <https://www.estillvoice.com/pages/clinical-software>.
38. Estill Voice International, LLC. Estill Voiceprint Plus versão 6.03. [website]. Pittsburgh, PA: Estill Voice International, LLC; c2010. [acesso em fevereiro 3, 2012]. Disponível em: <https://www.estillvoice.com/products/>.
39. Andrews G, Ash R, Brochec C, Busleiman AB et al. Audacity versão 2.0.0. [programa de computador]. c1999-2012 [acesso em dezembro 20, 2011]. Disponível em: <http://audacity.sourceforge.net/download/>.
40. Huckvale M. SFS|WASP versão 1.51. [programa de computador]. London: University College London, Division of Psychology & Language Sciences; c2009 [atualizado em setembro, 2011; acesso em janeiro 25, 2012]. Disponível em: <http://www.phon.ucl.ac.uk/resource/SFS|WASP.htm>.
41. Tiger DRS, Inc. ScopeView. [versão de demonstração do programa de computador]. Seattle, WA: Tiger DRS, Inc.; c1998 [acesso em janeiro 15, 2012]. Disponível em: <http://www.drsspeech.com/Information.html#Download>.
42. Maximilians L, Albrechts K. The EMU Speech Database System 2.3.0. [programa de computador]. Alemanha: Institute of Phonetics and Speech Processing, LMU Munich; c2010 e Institute of Phonetics and Digital Speech Processing, CAU Kiel; c2006 [acesso em fevereiro 5, 2012]. Disponível em: <http://EMU.sourceforge.net/>.
43. SIL International. Speech Analyser 3.0.1. [programa de computador]. Dallas, Texas: SIL International; c1996-2007 [acesso em dezembro 27, 2011]. Disponível em: <http://www.sil.org/computing/speechtools/>.
44. Tiger DRS, Inc. Real Analysis. [programa de computador]. Seattle, WA: Tiger DRS, Inc.; c1998.
45. Tiger DRS, Inc. Speech Training. [programa de computador]. Seattle, WA: Tiger DRS, Inc.; c1998.
46. Sjölander K, Beskow J. Wavesurfer 1.8.5. [programa de computador]. Estocolmo, Suécia: Centre for Speech Technology (CTT), Kungliga Tekniska högskolan (KTH); c2005 [atualizado em novembro 17, 2011; acesso em janeiro 21, 2012]. Disponível em: <http://www.speech.kth.se/wavesurfer/index.html>.
47. Behlau M. FonoTools: software para distúrbios da comunicação. [website]. Paraná, Brasil: CTS Informática, c2012 [acesso em abril 15, 2012]. Disponível em: <http://www.ctsinformatica.com.br/#fonoTools.html{paginaProduto!6&1>.
48. WEVOS – WEVOSYS. lingWAVES HOW-TO: Professional Voice Range Profile with Voice Diagnostic Center (VDC). [filme]. Forchheim, Germany: WEVOS – WEVOSYS; 2012. [acesso em agosto 22, 2012]. Disponível em: http://www.wevossys.com/support/lingWAVES_how_to/supprrt_how_to.html.

Recebido em junho/12; aprovado em outubro/12.

Endereço para correspondência

Lénia Ferreira

Endereço: Rua Antônio Alexandre Raposo, lote 44
7600-135 Aljustrel

E-mail: lenia.ferreira@gmail.com

