



UNIVERSIDADE D
COIMBRA

João Couceiro e Castro

«MÁQUINA DE OUVER»
REPRESENTAÇÃO DO DISCURSO ORAL PELA TIPOGRAFIA

Dissertação no âmbito do Mestrado em Design e Multimédia
orientada pelo Professor Doutor Fernando Jorge Penousal Martins Machado
e pela Professora Doutora Ana Madalena de Sousa Vasconcelos Matos Boavida
apresentada ao Departamento de Engenharia Informática
da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

setembro de 2019

> {

Máquina de Ulver

}

João Couceiro e Castro
setembro 2019

AGRADECIMENTOS

Ao professor Penousal Machado,
por duvidar sempre de mim, das minhas capacidades e das
minhas ideias descabidas.

À professora Ana Boavida,
por nunca ter deixado de acreditar e por me transmitir o gosto e
o respeito pela tipografia.

Ao professor Pedro Martins,
por todo o apoio e motivação desde o primeiro momento.

Ao professor Ernesto Costa,
por ter sido a pior primeira impressão que tive da universidade e
a melhor relação que tive com um professor.

Ao professor João Bicker,
por, há 3 anos me ter dado o conselho certo no instante certo
e, involuntariamente, ter despoletado a cadeia de eventos que me
trouxe até este momento na minha vida académica.

Aos professores Tiago Cruz e Artur Rebelo,
por me mostrarem que não só podemos, como devemos estar
dispostos a aprender com qualquer pessoa, sendo um exemplo de
integridade, humildade e profissionalismo.

Aos professores Nuno Lourenço e Nuno Coelho,
por terem despendido do seu tempo para me ajudar quando
precisei.

Aos meus grandes amigos, José Maria Simões, Fábio Pereira, Pedro
Matos Chaves, Rui Gaspar, Diogo Ferreira e Matilde Branco,
por me terem recebido de braços abertos e pela relação tão
forte que desenvolvemos ao longo dos anos. Custa-me imaginar

AGRADECIMENTOS

um futuro em que as horas que passemos a trabalhar não serão à mesma mesa. Estejamos onde estivermos a partir de hoje, estou certo que continuaremos sempre assim, juntos.

Aos meus pais, Paula Couceiro e Manuel Castro,
por todo o apoio e amor incondicional. Pelo esforço acrescido que fizeram para poder chegar aqui. Pela liberdade que me deram e pela oportunidade depois de abusar dela. Pelos valores e princípios que tenciono transmitir aos meus filhos. Por tudo.

À minha irmã e ao meu cunhado, Andreia e Rui Pedro,
por serem tão mais do que isso.

Às nossas famílias,
por serem muitos e muito bons.

Aos meus futuros sogros, Carmo e Manuel,
por serem mesmo os meus segundos pais.

Às nossas futuras madrinhas, Ema e Dora,
por tomarem conta da Francisca e por me desencaminharem de vez em quando, também foi preciso.

À Lai Lai,
por ter sido avó, colega de casa, amiga e confidente. Estou certo de que estarás aí para mim, como sempre aqui estiveste.

Para a Francisca,
razão tanto da minha existência quanto desta tese.

*Peço ao tempo que me leve de volta àquele lugar,
Que, com ele, traga tudo o que for capaz de lembrar
E aquilo que hoje gostaria de esquecer,
Isso, que o deixe lá ficar.*

*Não preciso de nada do que fez de mim quem fui,
Se de quem fui não me orgulho ser.
Apenas quero guardar o que sou,
Hoje, depois de te conhecer.*

*Mereces uma homenagem, uma declaração,
Mereces mais que uma viagem, um poema ou uma canção.
Quero dar-te um momento,
Preservá-lo na história,
Fazer do papel um pedaço de memória.*

*Pensei em mil e uma formas de te mostrar,
De te dizer.*

*Esta é a tua carta de amor,
Na minha máquina de ouver.*

RESUMO

A tipografia é considerada por diversos autores a representação visual da linguagem e, através da escrita, o ser humano encontrou uma forma de registar e partilhar informação. Ao longo da história da tipografia, muitos foram os autores que exploraram a ligação entre a palavra e a sua sonoridade, tentando aproximar o que é dito do que se ouve e de como se lê.

Apresentamos a «Máquina de Ouvir», um sistema que analisa gravações de discurso e cria uma representação visual para a sua expressividade, com recurso à composição e variáveis tipográficas.

Este sistema tira proveito das possibilidades de *scripting* do *Praat*, *Adobe InDesign* e *Adobe After Effects* para recolher dados sobre as características sonoras de gravações de discurso e, dinamicamente, criar uma composição tipográfica que resulta num artefacto estático ou dinâmico, como um poster ou vídeo.

Sendo uma das formas de discurso mais dinâmicas e ricas em termos de expressividade, o material de experimentação foca-se, principalmente, em interpretações de poesia.

PALAVRAS CHAVE

Som
Discurso
Tipografia
Expressividade
Composição Tipográfica

ABSTRACT

Typography is considered by many authors to be the visual representation of language and, through writing, the human being has found a way to record and share information. Throughout the history of typography, many authors have explored the link between the word and its sound, trying to approximate what is said with what one hears and how one reads.

We present the «Máquina de Ouvir», a system that analyzes speech recordings and creates a visual representation for their expressiveness, using composition and typographic variables.

This system takes advantage of the scripting capabilities of *Praat*, *Adobe InDesign*, and *Adobe After Effects* to gather data on the sound features of speech recordings and dynamically create a typographic composition that results in a static or dynamic artifact, such as a poster or video.

As one of the most dynamic and richly expressive forms of speech, experimentation material focuses mainly on interpretations of poetry.

KEYWORDS

Sound
Speech
Typography
Expressiveness
Typographic Composition

ÍNDICE

INTRODUÇÃO	13
ESTADO DA ARTE	17
PLANO DE TRABALHOS	37
EXPERIMENTAÇÃO INICIAL	39
«MÁQUINA DE OUVER»	57
AVALIAÇÃO DO SISTEMA	67
DISSEMINAÇÃO DO TRABALHO	73
CONCLUSÃO	75
BIBLIOGRAFIA	77
APÊNDICES	81

INTRODUÇÃO

Desde os primórdios da humanidade, a comunicação oral provou ser uma das capacidades mais importantes para a sobrevivência e evolução da nossa espécie. Com o progresso social e tecnológico, a necessidade de preservar e partilhar informação levou à invenção da escrita. O que começou como um sistema tridimensional recorrendo a fichas de barro como um sistema de contabilidade para acompanhar transações de mercadorias, evoluiu finalmente para uma abstração bidimensional e representação visual do som, através da conjugação de letras do alfabeto para a representação de fonemas (Denise Schmandt-Besserat, 2015).

A transição das técnicas caligráficas para a impressão mecânica, no século xv, proporcionou as condições necessárias para a produção em massa de materiais impressos mas só no início do século xx alguns artistas começaram a olhar para a página impressa enquanto um objeto de exploração criativa (Meggs & Purvis, 2012).

Nos trabalhos de poetas e artistas da vanguarda do século xx como Stéphane Mallarmé (1842-1898), Filippo Marinetti (1876-1944), Guillaume Apollinaire (1880-1918), Hugo Ball (1886-1927), Ilia Zdanevich (1894-1975), Tristan Tzara (1896-1963) ou Kurt Schwitters (1887-1948), é possível encontrar alguns dos primeiros exercícios que exploram a ligação entre o aspeto visual da linguagem e o seu som mas é no trabalho de interpretação gráfica executado por Robert Massin de *La Cantatrice Chauve* (1964), a peça de teatro do absurdo de Eugène Ionesco (1909-1994) que podemos encontrar um dos exemplos mais interessantes e relevantes do uso da tipografia para a transmissão da expressividade sonora da voz, através da manipulação em termos da composição, do espaço em branco e da forma, peso, tamanho ou inclinação da letra.

Com a ascensão da era informática e conseqüente consolidação das ferramentas digitais contemporâneas no apoio ao processo de design, a manipulação da tipografia e a exploração da sua ligação com o som tornou-se mais acessível e, por isso, presente na cena da arte e do design, analógica ou digitalmente, através do trabalho de

INTRODUÇÃO

artistas e designers contemporâneos como Alan Kitching, Paula Scher e Niklaus Troxler.

Apesar do foco principal dos tipógrafos ser, geralmente, a palavra escrita, a noção de que «a forma como usamos as inflexões, cadência e volume vocais, para realçar, exprimir ou clarificar, pode ser interpretada visualmente com o propósito de dar expressão gráfica à palavra impressa» (Jury, 2006).

Partindo da exploração das capacidades expressivas da letra no sentido de representar as variações na entoação da voz humana, esta dissertação posiciona-se entre as áreas da tipografia e do som.

OBJETIVOS

O principal objetivo desta dissertação é o desenvolvimento de um sistema capaz de analisar e representar visualmente as variações na expressividade da voz humana através da tipografia, com recurso a técnicas e ferramentas de análise de som e de composição tipográfica. Para tal, é necessária a definição de um conjunto de regras que permita o mapeamento de características sonoras para variáveis tipográficas.

Assim, com o desenvolvimento prático deste projeto espera-se a consolidação de um contributo científico para área do design, explorando as possibilidades e limitações da manipulação tipográfica por vias da computação, na representação visual do discurso humano.

ESTRUTURA DO DOCUMENTO

Depois de uma breve introdução, enquadramento e apresentação dos objetivos desta dissertação, é apresentado um breve resumo da investigação efectuada relativamente ao **Estado da Arte** das áreas abordadas - tipografia e som. Neste capítulo é apresentada uma breve história acerca das origens da tipografia e são apresentados casos relevantes da utilização da letra como meio expressivo do som na prática artística das vanguardas do século xx, assim como no design contemporâneo. Serão também apresentados conceitos essenciais da área do som, assim como trabalhos relacionados com

ESTRUTURA DO DOCUMENTO

esta tese, na medida em que utilizam técnicas computacionais para a representação do discurso recorrendo à manipulação sistemática da tipografia.

Seguidamente é apresentado o **Plano de Trabalhos** no qual se encontram descritas, de uma forma detalhada, todas as tarefas realizadas e a metodologia adotada no processo de desenvolvimento do projeto.

No capítulo da **Experimentação Inicial** são apresentadas as três principais fases de experimentação, assim como os artefactos resultantes de cada uma delas, as suas análises e a linha de pensamento da qual resultou o conjunto de regras para o mapeamento das características sonoras para as variáveis tipográficas adotado.

O sistema resultante deste trabalho de investigação encontra-se detalhadamente descrito no capítulo «**Máquina de Ouvir**», no qual são apresentadas as quatro etapas do sistema e o conjunto de regras usadas na geração de artefactos estáticos e dinâmicos.

Seguidamente, é apresentada a **Análise do Sistema** realizada através da análise de resultados de mais de 100 respostas a um questionário desenvolvido com o objetivo de avaliar as soluções propostas nesta dissertação.

No capítulo **Disseminação do Trabalho** são enumeradas as várias oportunidades aproveitadas para a divulgação do projecto e dos resultados obtidos.

Finalmente é apresentada a **Conclusão**, na qual se encontram algumas considerações finais, assim como perspectivas futuras para o projeto.

ESTADO DA ARTE

Neste capítulo é apresentada, de uma forma sintetizada, a pesquisa efetuada no sentido de recolher a informação necessária para a conceptualização do projeto proposto para esta investigação.

TIPOGRAFIA

A definição de Tipografia está longe de ser unânime e, ainda mais, de se resumir à que nos é apresentada nos dicionários – arte de compor e imprimir (*tipografia* in Dicionário Infopédia da Língua Portuguesa [online]. Porto: Porto Editora). Mesmo não existindo um consenso sobre ela, é possível identificar uma máxima comum a diversos autores (Bringhurst, 2015; Cullen, 2012; Kane, 2011) afirmando que se baseia na representação visual da linguagem humana. Esta representação da linguagem é alcançada através de técnicas de composição, edição, impressão e visualização dando uso ao conjunto de símbolos a que chamamos caracteres.

Apesar das suas fortes raízes na escrita, criação e reprodução de materiais literários, a tipografia é uma área de conhecimento indispensável na comunicação visual seja ela de cariz informativo, comercial ou artístico.

O processo de escolha, combinação e manipulação de tipos de letra implica uma responsabilidade acrescida na principal tarefa do designer – a transmissão da mensagem. Seja na paginação de um romance ou na composição de um cartaz para um concerto de uma banda *Punk*, é importante que exista uma relação de respeito entre a tipografia e o conteúdo, uma vez que o papel principal da tipografia é honrar o texto e nunca, de forma alguma, sobrepor-se a ele. No entanto, esta tem uma «vida e dignidade própria» que não deve ser descurada (Bringhurst, 2015).

Uma das mais importantes e influentes figuras da tipografia do século xx, Jan Tschichold, no seu ensaio *Clay in a Potter's Hand* de 1948, afirma que «quanto mais importante o conteúdo de um livro, mais tempo deverá ser preservado, e quanto mais equilibrado, mais

CARACTERE

Cada uma das letras, números e sinais que compõem uma fonte tipográfica.

ESTADO DA ARTE

perfeita deverá ser a tipografia» (Tschichold, 1991).

A perfeição a que Tschichold se refere é, com certeza, discutível, mas, de uma forma geral, não depende apenas da escolha do tipo e do tamanho da letra, mas também de aspetos como o alinhamento, espaçamento, entrelinha ou as margens e as proporções e relações entre todos os elementos presentes na mancha e na página. Assim, será tão importante o espaço em branco quanto o elemento impresso.

ESPAÇAMENTO

Ajuste do espaço global entre as letras ou palavras.

ENTRELINHA

Espaço entre as linhas de base de duas linhas consecutivas de texto.

MARGEM

Espaço em branco que rodeia a mancha escrita de uma página.

ORIGENS DA TIPOGRAFIA

Não obstante a pertinência da origem da escrita e das necessidades que despoletaram os processos evolutivos que resultaram na capacidade comunicativa com recurso a artefactos visuais do ser humano para esta investigação, do ponto de vista histórico da evolução da tipografia, segundo autores como Kristin Cullen e John Kane (Cullen, 2012; Kane, 2011), faz sentido que o ponto de partida desta resumida revisão histórica seja marcado pela origem do alfabeto como hoje o conhecemos.

Segundo Meggs (Meggs & Purvis, 2012), um alfabeto é «um conjunto de símbolos visuais ou caracteres usados para representar sons elementares de uma linguagem falada». A palavra "alfabeto" resulta da junção de alfa e beta, as duas primeiras letras do alfabeto grego, uma adaptação do desenvolvido pelos fenícios um século (Kane, 2011) antes e que, por sua vez, foi adotado e revisto pelos romanos, disseminando pelo Império Romano aquele que é o antepassado mais próximo do alfabeto atual (Meggs & Purvis, 2012).

Na Figura 2.1. podemos observar um exemplo de uma inscrição de capitais romanas, na base da coluna de Trajano, em Roma, que, de acordo com as palavras de Allen Haley «tiveram, e continuam a ter, a maior influência no desenho e utilização de letras maiúsculas, mantendo-se como a norma de proporção e dignidade por quase dois mil anos».

Por sua vez, as letras minúsculas são o resultado de centenas de anos de evolução (Cullen, 2012) partindo da escrita cursiva romana, usada na documentação temporária como, por exemplo, transações quotidianas, de forma a facilitar o processo de escrita (Kane, 2011).



FIG 2.1.
Inscrição na coluna Trajana.

CLAREZA

Refere-se à percepção e mede-se através da velocidade com que cada carácter consegue ser compreendido (Tracy, 2003).

CAIXA-BAIXA

Alusivo às caixas usadas na tipografia tradicional, onde os caracteres minúsculos se encontram na zona inferior.

No final do século VIII, Carlos Magno determina que todos os textos eclesiásticos sejam padronizados (Kane, 2011) no que toca ao esquema de página, estilo de escrita e decoração, levando, também a uma reforma no alfabeto (Meggs & Purvis, 2012). Assim, no sentido de alcançar uma maior clareza no texto aliada à facilidade de escrita, nasce a minúscula Carolíngia, o principal antecessor da letra de caixa-baixa atual (Cullen, 2012).



FIG 2.2.
Páginas 144 e 145 da Bíblia de Gutenberg.

A transição das técnicas caligráficas para a impressão mecânica, na civilização Ocidental, deu-se em meados do século XV. A figura principal à qual se atribui maior relevância é o alemão Johannes Gutenberg (Cullen, 2012) que, apesar de não ser o responsável pela invenção dos tipos móveis, as suas capacidades nas áreas da engenharia, fundição e química permitiram-lhe criar moldes para

ESTADO DA ARTE

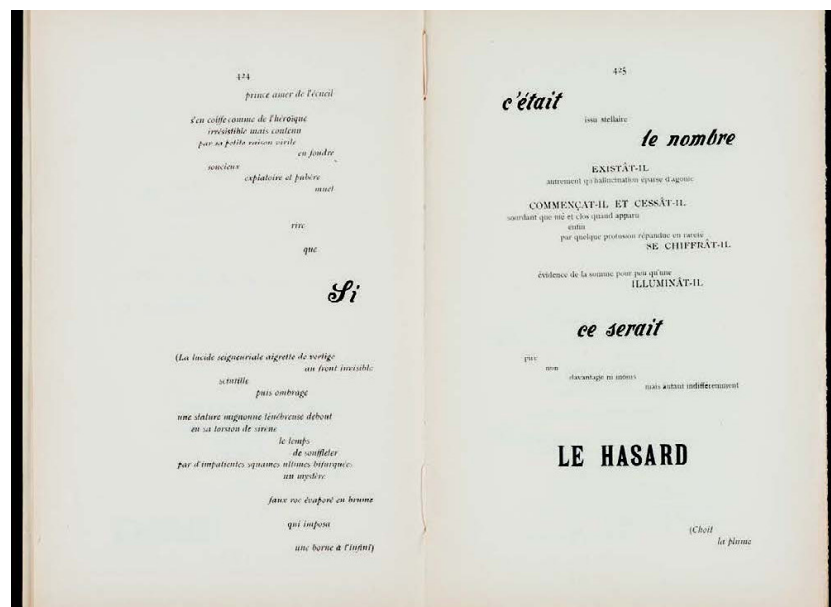
fundir tipos móveis e reutilizáveis feitos de chumbo. Para além disso, idealizou uma prensa baseada nas usadas na produção de vinho e até uma tinta com propriedades favoráveis à impressão com tipos de metal (Kane, 2011). Esta inovação não só lhe permitiu a produção da tão aclamada *Bíblia de Gutenberg*, apresentada na Figura 2.2., como proporcionou condições para a produção massiva de inúmeros materiais impressos e uma grande facilidade na reutilização dos tipos (Lupton, 2004).

Desde então a história fica marcada por inúmeras personalidades, desde tipógrafos, designers de tipos, designers gráficos e artistas, que exploram o uso da tipografia como meio para criar, comunicar e questionar o mundo que os rodeia.

PRÁTICAS TIPOGRÁFICAS NAS VANGUARDAS

Uma vez analisada a história do design e, mais especificamente da tipografia, optámos por dar especial atenção ao trabalho desenvolvido por alguns dos protagonistas dos movimentos de vanguarda que marcaram as primeiras duas décadas do século XX, não só pela sua forte expressividade tipográfica, mas também pelo carácter experimental dos mesmos.

FIG 2.3.
Un coup de dés
Stéphane Mallarmé
1897

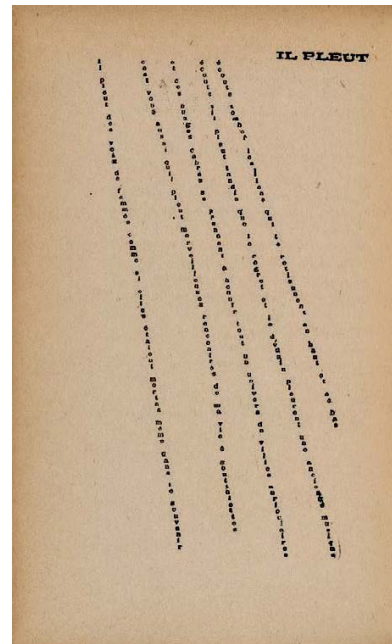
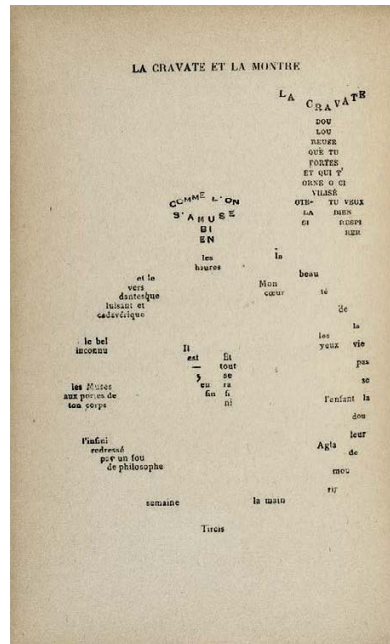


ESTADO DA ARTE

de forma a serem interpretadas, não só como texto, mas também como figuras resultando numa série de poemas visuais.

Apesar da sua relevância histórica, o trabalho desenvolvido por Apollinaire nesta área centra-se principalmente na dimensão figurativa do texto, pelo que o interesse para esta dissertação é limitado pela sua ausência de expressividade.

FIG 2.5.
Poemas da obra
Calligrammes
Guillaume
Apollinaire
1918



A contribuição mais significativa do Futurismo Russo parte de Ilia Zdanevich (1894-1975), artista e escritor russo, para o qual «a tipografia era mais do que um meio, era uma substância da sua própria expressão» (Drucker, 1994), explorando o potencial gráfico da linguagem nos seus poemas e peças *zaum* – linguagem poética experimental isenta de significado.

A sua experiência enquanto impressor e compositor, leva-o a nunca amontoar elementos tipográficos, quebrar palavras ou destruir letras, ao contrário de Marinetti.

Na sua obra *Ledentu as a Beacon*, apresentada na Figura 2.6. Ilia Zdanevich investiga de forma extensiva a estrutura fonética da língua russa e, posteriormente, procura uma forma de a representar em termos tipográficos (Drucker, 1994). Uma vez que não

Para cada personagem escolheu uma fonte tipográfica, variando o seu tamanho e posição consoante a expressividade de cada fala. O texto está para a página, como os atores para o palco.

Na sua obra *La lettre et l'image*, Massin refere-se ao seu próprio trabalho como sendo «um outro exemplo de texto "vísivel" que oferece um novo modo de se ler uma peça. O seu design combinou as técnicas do cinema às da banda desenhada, além de utilizar as verdadeiras faces dos atores, que adquiriram a aparência de um ideograma pelas mãos do fotógrafo Henry Cohen. Atuando como uma espécie de diretor de cena, o designer procura transmitir a atmosfera, o movimento, as falas e os silêncios da peça, tentando ao mesmo tempo transmitir uma ideia de duração temporal e espacialidade do palco através do simples jogo entre imagem e texto» (Massin, 1970).

Este não é o único trabalho deste género produzido por Massin. Em 1966 explora um outro texto de Ionesco em *Délire à Deux* (Figura 2.9.) na qual, segundo o autor Gustavo Piqueira, «brinca com diversos pesos, inclinações e sobreposições para construir volumes, entoações e ação» (Piqueira, 2018).



FIG 2.9.
Délire à Deux
Robert Massin
1966

ESTADO DA ARTE

EXPRESSÃO TIPOGRÁFICA NA ERA DIGITAL

Após o processo de digitalização da tipografia nos anos 1960 e 1970, a distinção entre estúdios de design, tipografias e gráficas começa a ser cada vez mais ténue e, aliando este fenómeno à democratização do acesso à tecnologia por via do computador pessoal, a tensão entre as técnicas tradicionais de composição tipográfica e os processos digitais levam a uma alteração drástica de paradigma na prática de design (Baines & Haslam, 2015) que se estende até aos dias de hoje.

Uma das figuras contemporâneas que melhor representa esta tensão intemporal entre a tipografia tradicional e moderna é Alan Kitching. Na Figura 2.10, do lado esquerdo, pode observar-se um dos posters de uma série que desenvolveu para uma produtora musical londrina no qual se pode identificar uma clara influência futurista. Neste trabalho o tamanho, ângulo de posicionamento e sobreposição das letras transmite uma sensação de movimento e tensão (Hillner, 2009).

Do lado direito da Figura 2.10, pode observar-se um poster produzido para o *London Design Festival* no qual é possível identificar uma clara intenção de transmitir a vocalização da palavra *táxi*, com uma forte e crescente intensidade.

FIG 2.10.
(à esquerda)
*Glory to the
New Born King*
Alan Kitching
1993

(à direita)
Taxi
Alan Kitching
2009



TIPOGRAFIA

A manipulação tipográfica, tanto ao nível do desenho da letra quanto à composição da página, na expressão visual de dinamismo e som, pode ser constatado no trabalho de outras grandes figuras do design contemporâneo como Paula Scher ou Niklaus Troxler.

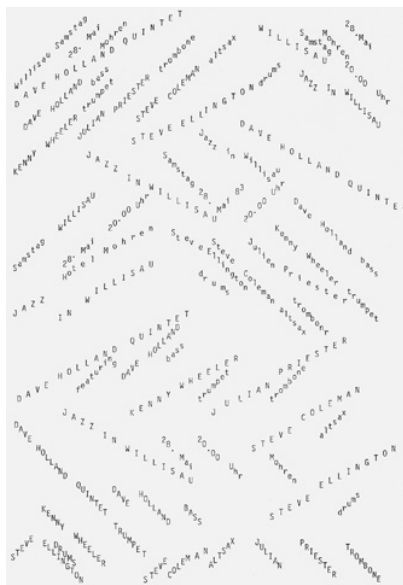
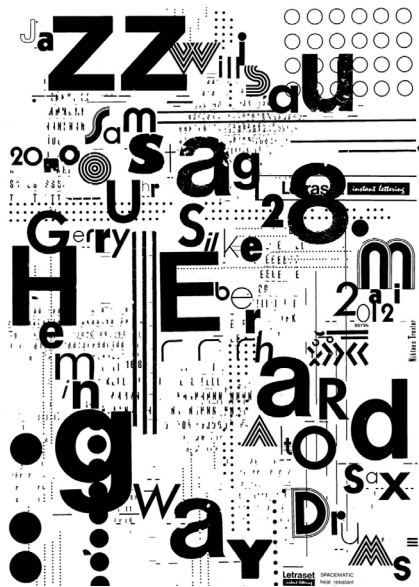


FIG 2.11.
(à esquerda)
Silke Eberhard-Gerry
Hemingway
Niklaus Troxler
2012

(à direita)
Dave Holland Quintet
Niklaus Troxler
1983

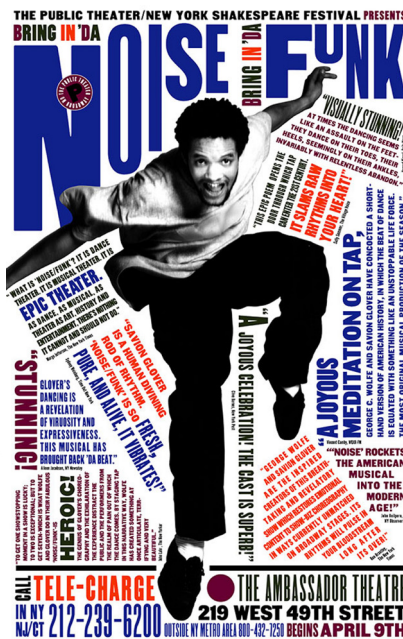
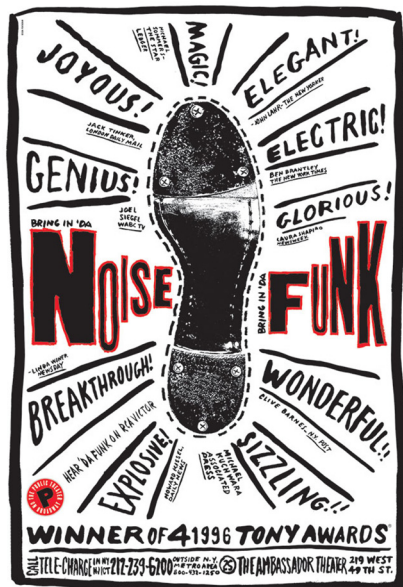


FIG 2.12.
Posters para
Public Theater
Paula Scher
2008

ESTADO DA ARTE

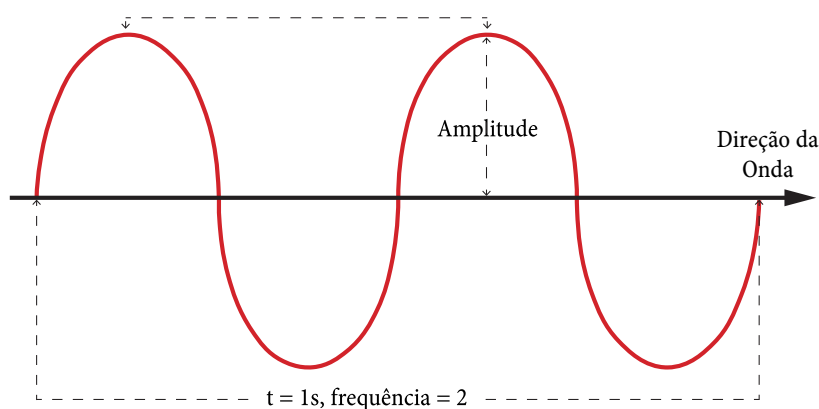
SOM

Segundo F. Alton Everest (1909-2005), engenheiro acústico americano, «o som tem uma natureza dupla: pode ser uma perturbação física num meio físico como o ar, ou pode ser considerado uma percepção psicofísica resultante de impulsos nervosos que estimulam o córtex auditivo do cérebro» (Ballou, 1987).

CARACTERÍSTICAS SONORAS

Partindo do trabalho de vários autores, o compositor canadiano especialista em música electroacústica e comunicação acústica Barry Truax, afirma existir uma distinção entre os parâmetros acústicos «objectivos», como a amplitude, frequência e a forma de onda, e os seus homólogos «subjectivos» psicoacústicos nomeadamente intensidade, altura e timbre, respectivamente, que constituem a resposta cerebral a esses parâmetros (Truax, 2001).

FIG 2.13.
Características de
uma onda.



INTENSIDADE Representa a intensidade do sinal áudio, que está directamente relacionada com a amplitude dos sinais, por vezes também referida como energia ou potência do sinal (Jang, n.d.).

A intensidade é medida através de incrementos de energia denominados decibéis, numa escala logarítmica de energia sonora em que cada dez pontos representam dez vezes mais intensidade. Assim, zero dB (decibéis) pode considerar-se o limiar inferior de audição (1 unidade de energia). Apesar de, de forma generaliza-

da, o ouvido humano ter capacidade para captar sons até aos 130 db (10 triliões de unidades de energia) antes de sentir dor, a voz humana é 1 milhão de vezes mais potente que o som emitido por uma lâmpada (Sonnenschein, 2001).

ALTURA Representa a frequência de vibração do sinal áudio, pode ser representada pela frequência fundamental ou, de forma equivalente, a inversa do período fundamental de sinais de áudio vocais (Jang, n.d.).

A altura do som refere-se ao espectro de frequências, altas (agudas) e baixas (graves), que podem ou não ser captadas pelo ouvido humano. A audição humana, normalmente, tem a capacidade de distinguir frequências entre os 20 e os 20.000Hz (Hertz, ou ciclos por segundo), podendo esta ser afetada e diminuída com o avançar da idade.

Sons mais graves são chamados de infra-sónicos e as suas frequências são tão baixas que não constituem som e as suas vibrações são percecionadas como um conjunto de batidas individuais e são, muitas vezes, passíveis de serem confundidas com ritmo.

Sons mais agudos denominam-se de ultra-sónicos e apesar de muitas vezes não serem possíveis de ouvir, se forem emitidos com muita intensidade podem revelar-se incómodos (Sonnenschein, 2001).

TIMBRE Quando as ondas sonoras pulsam em intervalos regulares (também referidas como periódicas), criam um som puro ou tonal, por oposição ao som ruidoso resultante da sobreposição e mistura de frequências que constituem ondas sonoras altamente complexas (intervalos irregulares, não-periódicos). Algures entre estes extremos podem ser encontrados os instrumentos musicais que emitem várias frequências conhecidas como harmónicos. Para cada instrumento, os harmónicos têm formas de onda regulares que constituem o seu timbre.

Como exemplo de um timbre caracterizado por harmónicos mais simples temos a flauta e, por oposição, podemos considerar o violino como um instrumento com um timbre constituído por um padrão mais complexo de harmónicos. (Sonnenschein, 2001)

ESTADO DA ARTE

SINAIS DE ÁUDIO E VOZ HUMANA

Sinais de áudio referem-se, geralmente, a sinais que são audíveis para os humanos e, normalmente, provenientes de uma fonte sonora que vibra no intervalo de frequências audíveis. Considerando a fonte desses sinais, podemos classificá-los em duas categorias: sons animais, como a voz humana, ou sons não-animais, como o som produzido por um motor de um carro (Jang, n.d.).

Segundo Isabel Guimarães, na perspectiva biológica, a voz é definida como «som audível resultante da inter-relação complexa entre a pressão e velocidade do fluxo de ar expiratório (que influenciam directamente o volume, ou seja, a intensidade da vocalização), os diferentes padrões de mobilidade das pregas vocais (que influenciam directamente a sonoridade ou registo, também designado de timbre) e as propriedades de reflexão e configuração das estruturas do tracto vocal (que influenciam directamente a ressonância)» (Guimarães, 2007) sendo que, «o instrumento vocal compõe-se essencialmente de três partes: a fonte de energia – os pulmões; a fonte de vibração – a laringe e as pregas vocais; e as cavidades de ressonância – o tracto vocal» (Guimarães, 2011).

SOM QUASI-PERIÓDICO As ondas sonoras consistem em padrões repetidos e semelhantes de tal forma que somos capazes de perceber a altura. Exemplos são a reprodução monofónica da maioria dos instrumentos musicais (pianos, violinos, guitarras, etc) e seres humanos a falar/cantar (Jang, n.d.).

SOM APERIÓDICO As ondas não consistem de padrões repetidos óbvios e, por isso, não somos capazes de perceber uma altura estável. Exemplos são bater palmas, trovão, partes da dicção não vocal (Jang, n.d.).

De acordo com Jang, cada pequeno segmento da voz humana pode ser dividido em dois tipos: som vocal, produzido pela vibração das cordas vocais e nos quais somos capazes de perceber uma altura, e o som não vocal, produzido pelo fluxo rápido de ar através da boca, nariz ou dentes (Jang, n.d.).

Geralmente, uma vibração regular na produção de voz gera sons vocais quase-periódicos. Por outro lado, se a produção de som for a partir de um fluxo de ar irregular, então teremos sons não vocais (Jang, n.d.). Na Figura 2.14 podemos observar claramente que “s” e “x” são sons não-vocais, enquanto que “i” é um som vocal.

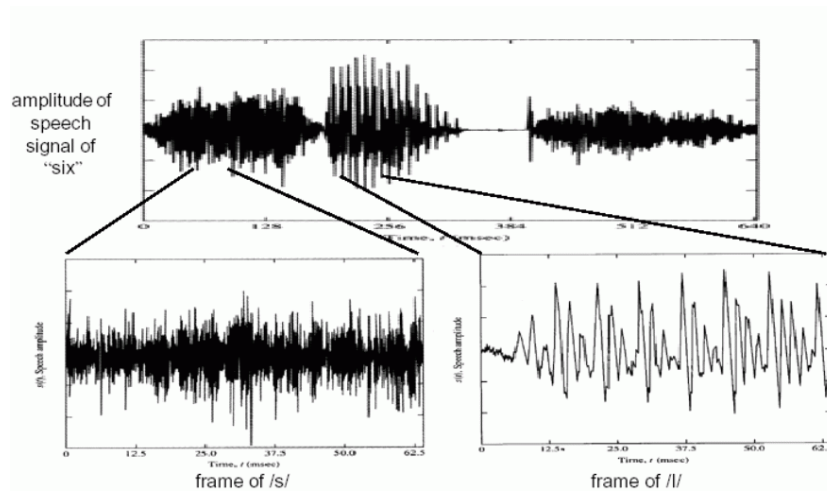


FIG 2.14.
Sons vocais e
não vocais.

PERFORMANCE MUSICAL EXPRESSIVA

Tirando partido das palavras de Juslin, «música é provavelmente a forma de arte mais praticada e apreciada. Uma possível explicação poderá ser o efeito emocional profundo que a música tem nas pessoas» (Juslin, 1997) e nas últimas décadas têm sido desenvolvidas variadas investigações no sentido de decifrar essa ligação entre o ouvinte e o performer, com base na hipótese de que a música pode transmitir realmente emoções.

Na opinião de Gabrielsson, a experiência que o ouvinte retira da música «depende muito de propriedades específicas dos sons produzidos pelo performer» (Gabrielsson, 1988).

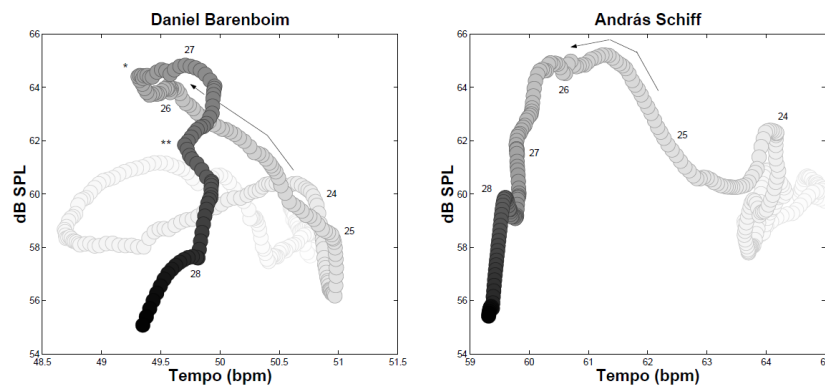
Deste modo, revela-se altamente pertinente a análise dos estudos efectuados nesta área, uma vez que as suas conclusões poderão ser relevantes para a nossa investigação no que diz respeito às características acústicas mais pertinentes de obter.

No trabalho desenvolvido por Simon Dixon, Werner Goebel

ESTADO DA ARTE

e Gerhard Widmer do Instituto de Investigação Austríaco para Inteligência Artificial, concentram a sua atenção «em duas das mais importantes dimensões expressivas: flutuações no tempo e na intensidade (dinâmica)» através de uma ferramenta desenvolvida por eles, com base numa representação visual proposta anteriormente por Jorg Langner, cuja representação gráfica é apresentada na Figura 2.15, à qual chamam de *Performance Worm* (Dixon, Goebel & Widmer, 2002).

FIG 2.15.
Performance Worm



TRABALHOS RELACIONADOS

Não são muitos os trabalhos de investigação que procuram respostas no tópico das alternativas computacionais para a representação visual da expressividade do discurso. Assim foi feita uma análise dos que despertam maior interesse para esta dissertação.

KINETIC TYPOGRAPHY: ISSUES IN TIME-BASED PRESENTATION OF TEXT (FORD, FORLIZZE & ISHIZAKI, 1997)

Neste trabalho, Ford, Forlizzi e Ishizaki, exploram as possibilidades temporais dos meios digitais, na tentativa de entender o potencial comunicativo da tipografia cinética. Assim, existem alguns aspectos que são pertinentes para a nossa investigação como, por exemplo, a sugestão de representar a variação da intensidade das vocalizações através da variação de tamanho e do peso da fonte. Na Figura 2.16 é apresentado um exemplo.

TRABALHOS RELACIONADOS

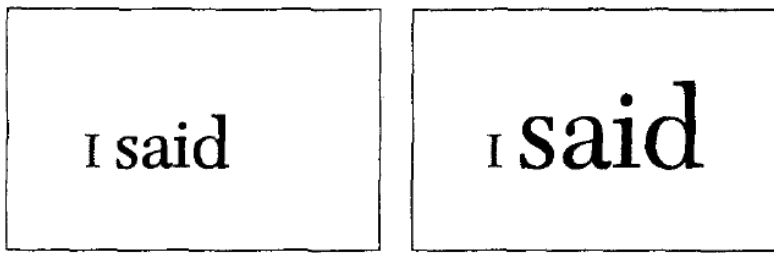


FIG 2.16.
Alterações no tamanho e no peso da fonte de forma a sugerir um aumento no volume vocal.

PROSODIC FONT: TRANSLATING SPEECH INTO GRAPHICS (ROSENBERGER & MACNEIL, 1999)

A investigação levada a cabo por Rosenberger e MacNeil, consiste na criação de uma fonte com base na prosódia do discurso. No decorrer do seu trabalho, formularam certas questões e executaram algumas experiências que nos podem ser úteis no desenvolvimento da nossa tese como, por exemplo, o facto de aplicar alterações ao nível da palavra ou da sílaba sendo que, a divisão por sílabas, torna a leitura mais próxima da forma como ouvimos discurso oral.

Outro ponto pertinente a retirar do trabalho de Rosenberger e MacNeil, centra-se nas alterações visuais feitas à forma da letra, sendo que um discurso suave produz fontes de peso leve e tamanho reduzido, enquanto que o discurso mais intenso produz fontes maiores e mais pesadas. No que toca à altura do discurso, quanto mais agudo, mais estreitas e esticadas se tornam as letras e quanto mais grave, mais achatadas e alongadas ficam.



FIG 2.17.
Exemplo de fonte prosódica.

ESTADO DA ARTE

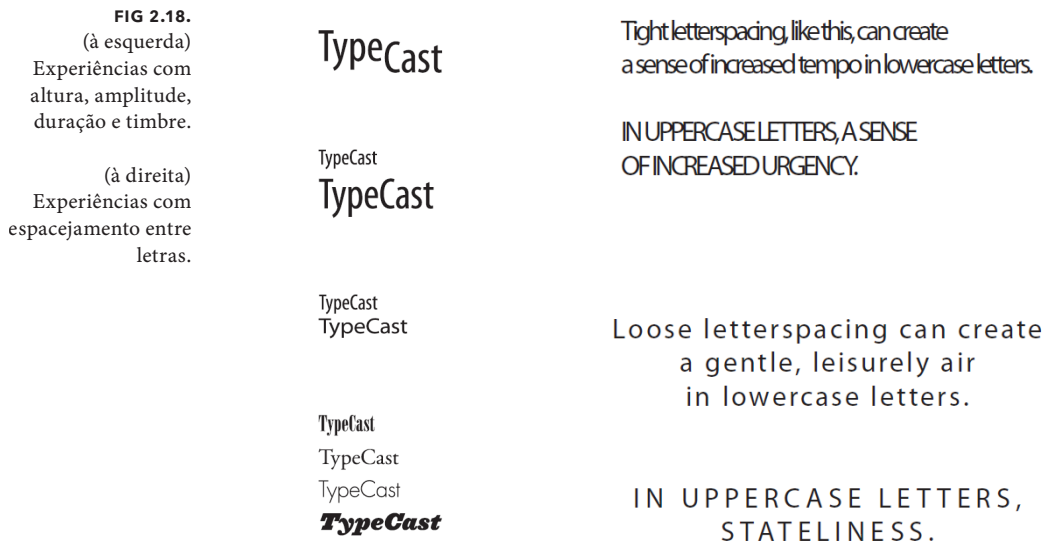
De forma a testar a sua teoria, fizeram um pequeno ensaio com quatorze indivíduos em que doze deles foram capazes de associar as imagens geradas com a fonte prosódica ao discurso oral no qual se baseavam.

TYPECAST: THE VOICE OF TYPOGRAPHY (FELLOWS, 2009)

O trabalho desenvolvido por Fellows, procura encontrar o tom de voz de várias personagens de um filme e atribuir-lhes tipos de letra que os representem.

Como podemos ver no lado esquerdo da Figura 2.18, as suas experiências iniciais traçam a ponte entre aquelas que considera as principais qualidades do som – altura, amplitude, duração e timbre – e as variáveis tipográficas que melhor as representam, entrelinha, peso e tamanho, espaçamento entre letras e largura, e o estilo da fonte. Outra experiência que nos é útil analisar consiste no uso do espaçamento entre letras.

No lado direito da Figura 2.18 podemos observar o resultado da diminuição do espaçamento entre letras que, conclui Fellows, no caso dos caracteres minúsculos cria uma sensação de aumento do tempo e, no caso das maiúsculas, causa uma sensação de desconforto.



TRABALHOS RELACIONADOS

VOICE-DRIVEN TYPE DESIGN (WOFEL, SCHILIPPE & STITZ, 2015)

A investigação de Wolfel, Schlippe e Stitz tem como principal foco a fonética. Assim, concentra-se em trabalhar ao nível do fonema e do grafema ou grafemas correspondentes e, a partir da amplitude, altura e velocidade de cada um deles, variar a forma de cada caractere individualmente.

Tendo como base o tipo de letra *Futura* de Paul Renner (1927), as características do desenho dos caracteres a serem influenciadas pela amplitude, altura e velocidade são, respectivamente, o peso do traço vertical, o peso do traço horizontal e a largura do caractere.

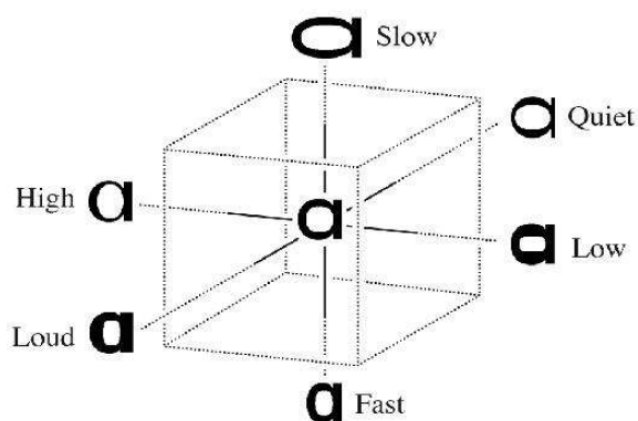


FIG 2.19.
Esquema da variação dos parâmetros.



FIG 2.20.
Aplicação a legendas.

PLANO DE TRABALHOS

METODOLOGIA

A metodologia adotada para o desenvolvimento deste projeto aproxima-se à de *Design Science Research* no sentido em que se baseia em três pilares: a obtenção do máximo de conhecimento possível nas áreas a abordar, a experimentação orientada à criação de artefactos e a análise do resultados obtidos a cada fase de experimentação na constante procura da solução mais adequada à resolução do problema em mãos.

TAREFAS

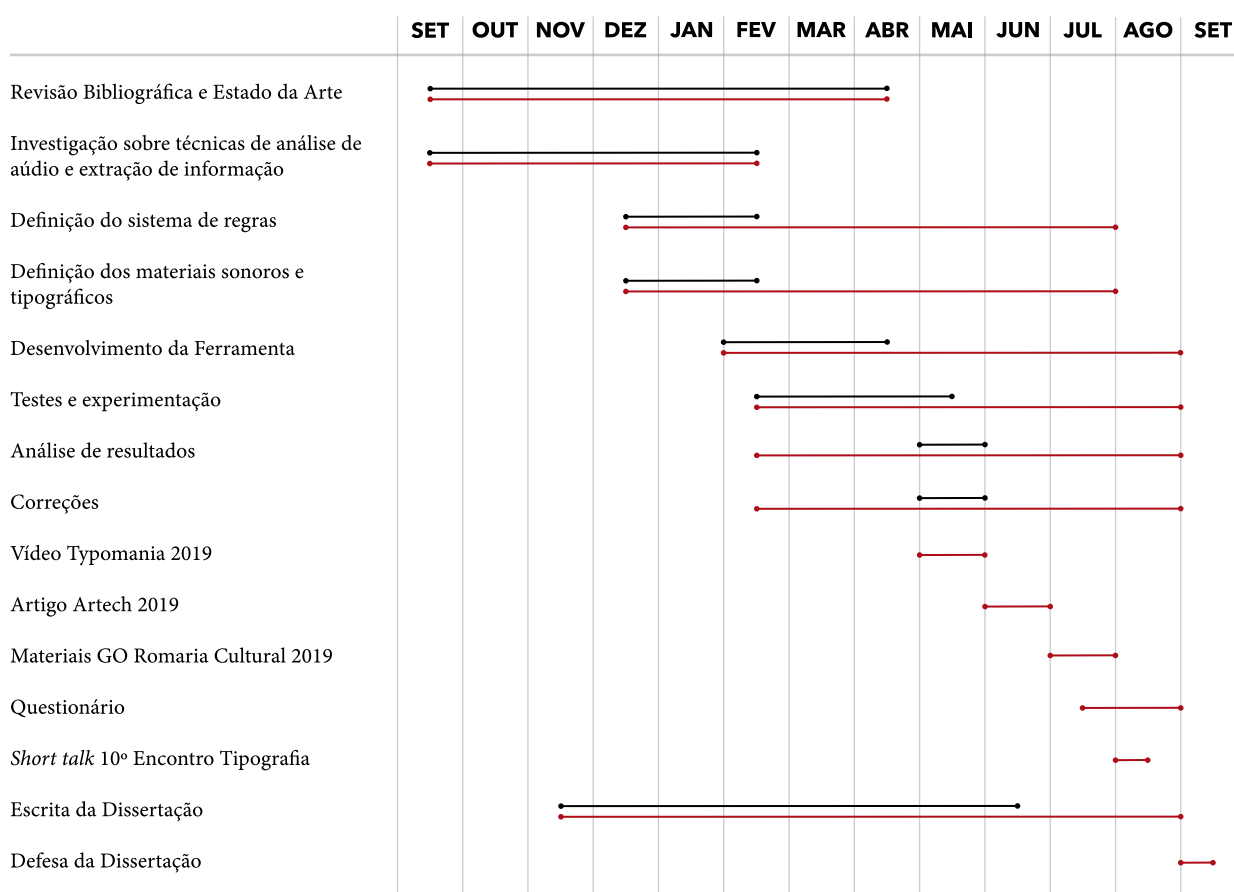


FIG 3.1.
Plano de Trabalhos proposto.

PLANO DE TRABALHOS

O processo de desenvolvimento deste projeto iniciou-se com uma **Revisão Bibliográfica e Estado da Arte**, de forma a reunir o máximo de informação necessária não só nas áreas da tipografia e do som, no geral, mas também acerca de trabalhos relacionados com o cruzamento das duas na tentativa de explorar as possibilidades expressivas da tipografia enquanto representação do som.

A passo com este trabalho de pesquisa, foi essencial fazer uma **Investigação sobre técnicas de análise de áudio e extração de informação** uma vez que foi durante esta tarefa que se teve conhecimento da existência de uma das ferramentas essenciais no sistema desenvolvido.

Inicialmente terá sido proposto que a **Definição do Sistema de Regras** fosse executada antes do início do segundo semestre, mas dado o grau de complexidade do problema proposto a resolver e a metodologia adoptada, o sistema de regras foi sofrendo alterações ao longo de todo o processo de **Testes e Experimentação e Desenvolvimento da Ferramenta**, assim como a **Definição dos materiais sonoros e tipográficos** a utilizar.

Estaria previsto inicialmente que seria efectuada uma **Análise de Resultados** de um processo único de experimentação e feitas as devidas **Correções**. Mais uma vez, estas são tarefas que acompanharam o processo de desenvolvimento e que foram executadas a cada ciclo de experimentação, na procura de uma solução cada vez mais eficaz.

Tarefas como a preparação de um vídeo para o festival Typomania 2019, a escrita e submissão de um artigo científico para a *Artech 2019*, a preparação de uma *short talk* para o 10º Encontro de Tipografia e os materiais para a *GO Romaria Cultural* não foram previamente planeadas mas acrescentadas ao plano de trabalhos de forma espontânea uma vez que se trata da disseminação do projeto.

Finalmente constatou-se a importância de validar os resultados obtidos e decidiu-se levar a cabo a criação de um **Questionário** a partir do qual se fez uma análise estatística das respostas recolhidas.

A acompanhar todo este processo, até ao momento da sua conclusão, foi feita a **Escrita da Dissertação**, assim como a preparação da apresentação pública e **Defesa da Tese**.

EXPERIMENTAÇÃO INICIAL

Feita a devida recolha e análise do trabalho desenvolvido nesta área até à data e das ferramentas necessárias à execução deste projeto, iniciou-se o trabalho de experimentação com o objetivo de testar a possibilidade teórica subjacente a esta tese de representar a expressividade do discurso humano recorrendo à análise de gravações sonoras e posterior manipulação tipográfica.

MAPEAMENTO MANUAL INTUITIVO

Não existindo quaisquer evidências da exequibilidade desta tese, partiu-se do princípio de que seria precipitado avançar com o desenvolvimento de um sistema computacional sem esboçar um mapeamento manual que o justificasse. Assim, a primeira fase de experimentação foi realizada sem recorrer a qualquer técnica computacional, baseando-se inteiramente na intuição e perceção individual do autor.

SETUP EXPERIMENTAL

O ficheiro de som usado para este exercício experimental foi a interpretação do *diseur* e ator português João Villaret (1913-1961) do poema *Cântico Negro*, originalmente escrito pelo poeta português José Régio (1901-1969). Esta gravação foi retirada do álbum *João Villaret no São Luís*, editado em 1961 pela Valentim de Carvalho. Sendo dotada de elevado grau de dinamismo, esta interpretação facilita a identificação das variações na expressividade do discurso do interlocutor.

Recorreu-se ao *software Adobe InDesign* para criar o artefacto tipográfico resultante desta experiência.

A fonte usada para compor este artefacto foi a *TheSans*, parte da superfamília *Thesis*, desenhada nos anos 90 pelo designer de tipos holandês Luc(as) de Groot. A escolha desta fonte baseia-se no vasto número de estilos de que dispõe e do seu desenho humanista. Esta

EXPERIMENTAÇÃO INICIAL

opção possibilita a composição com uma grande variedade de tamanhos e pesos de letra, sem comprometer totalmente a clareza do texto apresentado.

RESULTADOS EXPERIMENTAIS

Parte do artefacto resultante desta experiência é apresentado na Figura 4.1 na qual podem ser observadas e identificadas as regras criadas durante o processo da sua criação. O artefacto completo encontra-se no Apêndice A.

FIG 4.1.
Excerto do artefacto
resultante do
Mapeamento
Intuitivo.

“**Veeem** por aqui!!!”

dizem-me alguns com olhos **doce**s estendendo-me os **braço**s
e seguros de que seria bom que eu os ouvisse quando me **dizee**m.

“**VEEEM** por aqui!!!”

Eu **oo**lho-os com olhos **la**ssooos,
há nos meus olhos **ironi**as e **can**ni: **Sa**ços e cruço os **bra**çoos, e não vou por ali!!..

a minha **GLÓÓRIA É** ESSAAA!

Criar **de**sumanidade, não acompanhar ninguém
- Que eu vivo com o **mm**esmo sem-vontade com que **rrrr**rasguei o **vee**ntre a minha mãe

Nãããã, não vou por aiíí
Só vou por onde me levam **mm**meus próprios **pp**passos...

Se às coisas que eu pergunto **ee**m vão ninguém responde

Porque me dizeis **vóóóss**s: “**VEM** por aqui!!!”?

Prefiro **esccrrrr**regar nos **bb**ecos lamacentos, **rrrr**redemoinhar aos **vv**entos,
como **farrrr**apos **arrrrr**astar os pés **sannng**rentos, a **iiii**r por aiííi...

Se eu vim ao mundo, foi só para desbravar florestas virgens,
e desenhar meus **pró**prios **pass**os na areia inexplorada!
O mais que faço **nn**ão vale nada.

Como pois sereis vós que me dareis **ferrrr**amentas, machados e coragem
P_ra eu **derrrr**ubar os **mm**meus **obsss**táculos?...

COORRE nas vossas veias sangue **vel**ho dos avós,
E **mós** amais o que é **fff**ácil!

Eu amo o Longe e a Miragem,

Amo os abiiiiismos, as torrreentes, os deseertos...

MAPEAMENTO MANUAL INTUITIVO

A pronúncia prolongada de uma letra é dada pela repetição do caractere que a representa. O tamanho e o peso da fonte são usados para representar a ênfase dada a uma determinada enunciação, quanto mais intensa ou aguda a voz do intérprete, maior e mais pesada a fonte. As palavras ditas com um nível de intensidade que possa ser identificado como um berro são escritas em caixa alta.

CAIXA ALTA

O mesmo que maiúsculas.

A cor - escala de cinzentos - é usada em alguns caracteres para representar letras que fazem parte de uma palavra, mas que não são pronunciadas ou, quando usada juntamente com a repetição de caracteres, para criar um rasto visual transmitindo a extensão da letra aliada à sua crescente intensidade.

Quanto à relação entre o texto e a própria página, não é possível identificar uma regra para o seu posicionamento. No que toca ao posicionamento das palavras em relação às palavras que as antecedem, é possível identificar uma tentativa de que o espaço visual horizontal entre elas seja proporcional à duração da pausa no discurso. O mesmo se aplica ao espaço vertical entre os versos na forma de entrelinha.

ENTRELINHA

Espaço entre linhas.

ANÁLISE DE RESULTADOS

O tamanho e o peso da fonte são tidos como as duas variáveis tipográficas mais fortes. Uma vez que não estão a ser consideradas características específicas e os seus respetivos valores exatos não estão a ser analisados, as divergências nestas variáveis não representam verdadeiramente as diferenças percecionadas no discurso.

O texto é claro enquanto o tamanho da fonte não é demasiado pequeno e, por isso, deverá ser aplicado um valor mínimo para este parâmetro. Sendo uma fonte não serifada, compromete a composição em termos de legibilidade.

O largo espectro de pesos provou ser uma mais valia da fonte escolhida.

O uso da cor pode ser uma escolha válida, mas não é essencial, uma vez que está a ser usada da mesma forma que poderíamos usar o peso ou o tamanho da letra.

Constatou-se também que a solução para representar as pausas entre palavras e entre versos é um uso natural do espaço em branco.

EXPERIMENTAÇÃO INICIAL

A possibilidade de representar a duração de uma pronúncia através da repetição de caracteres parece direta e intuitiva, mas levanta a questão de como tornar visível a duração de uma palavra como um todo e se a solução se aproximará à utilizada para representar a duração das pausas.

Apesar do seu grau elevado de subjetividade e imprecisão, a criação deste artefacto provou ser uma tarefa valiosa na exploração de um conjunto de regras preliminar e dela resultou um artefacto que motivou a execução de uma nova fase de experimentação, recorrendo a técnicas computacionais.

MAPEAMENTO BASEADO NA PERFORMANCE MUSICAL EXPRESSIVA

Esta segunda fase de experimentação, tem como objetivo a exploração de técnicas computacionais para a obtenção de dados provenientes das gravações sonoras e a criação de artefactos tipográficos que as representem.

No planeamento destes testes, foi tomado como foco o campo da Performance Musical Expressiva, mais especificamente o trabalho de autores como Gerhard Widmer, Werner Goebel, Elias Pampalk, Simon Dixon e Jorg Langner, na tentativa de mimetizar a sua abordagem para a medição da expressão de uma performance musical baseada na dinâmica e no tempo (Dixon, Goebel, & Widmer, 2002; Goebel, Pampalk, & Widmer, 2004; Langner & Goebel, 2003).

SETUP EXPERIMENTAL

De forma a sistematizar o mapeamento das características sonoras às suas respetivas variáveis tipográficas, foi necessária uma simplificação quer dos materiais sonoros a utilizar, quer da quantidade de regras a aplicar.

Como material sonoro a ser analisado pelo sistema, foram usadas gravações de várias versões de uma mesma frase – «Isto nunca acaba.» – ditas pelo autor. Estas gravações foram realizadas em ambiente controlado numa cabine insonorizada com tratamento acústico situada nos estúdios da Casa das Caldeiras, explo-

MAPEAMENTO PERFORMANCE MUSICAL EXPRESSIVA

rados pelo *Projecto Imagem, Media e Comunicação* da Divisão de Comunicação da Universidade de Coimbra. Os níveis de ganho do microfone foram sempre mantidos no nível 6 e a distância entre o locutor e o microfone foi sempre de, aproximadamente, 10 centímetros, com recurso a um gravador *ZOOM H1N*. Estas condições foram mantidas em todas as gravações. Todas as gravações passaram por um processo de edição com o objetivo de isolar os excertos pretendidos e de remover o ruído de fundo.

Como é possível constatar na , manipulando certas características da oralidade como a duração das palavras, pausas, altura e intensidade é possível obter graus de expressividade diferentes. Com isto, esperam-se constatar variações na tipografia reagindo às alterações sonoras.

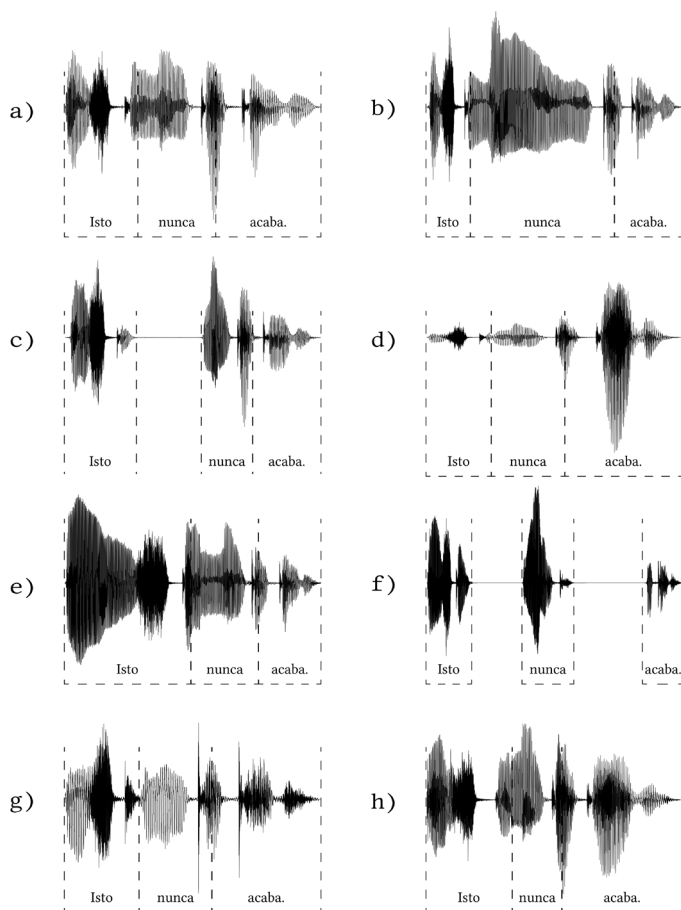


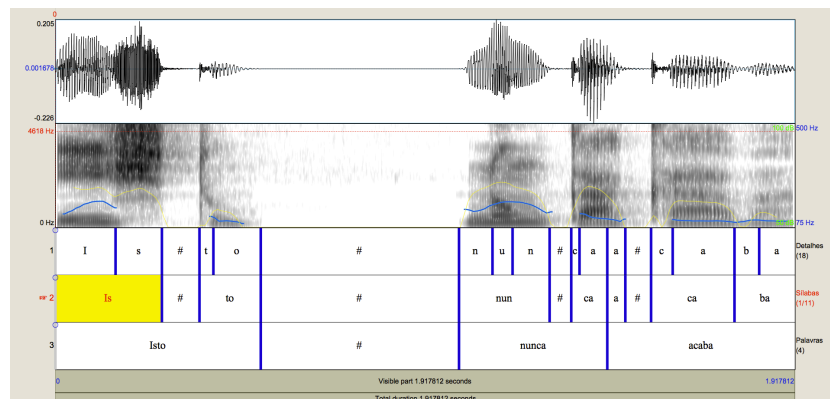
FIG 4.2.
Representação visual das gravações sonoras através das respetivas ondas.

EXPERIMENTAÇÃO INICIAL

Para a análise das gravações sonoras optou-se por recorrer ao software *Praat* (Boersma & Weenik, 2009), uma alternativa livre desenvolvida por Paul Boersma e David Weenik do Instituto de Ciências Fonéticas da Universidade de Amesterdão, com uma interface de utilizador simples, documentação detalhada e credibilidade científica. Para além disso, oferece a possibilidade de criação e execução de *scripts*, o que se prova útil para a extração automática de informação.

De forma a executar a extração de dados das gravações sonoras, é necessária uma fase prévia de anotação. Na Figura 4.3 é apresentada a interface gráfica deste *software*, exemplificando a estrutura da anotação aplicada aos excertos descritos anteriormente. A anotação é constituída por três camadas: Palavras, Sílabas e Letras. Com o objetivo de assinalar as pausas no discurso dá-se uso ao caractere “#”, uma vez que este não faz parte do conjunto de glifos do alfabeto da língua portuguesa para a formação de palavras. Assim, torna-se possível identificar programaticamente um momento de pausa, procurando apenas por esse caractere.

FIG 4.3.
Exemplo de anotação da gravação sonora c).



Os valores correspondentes a cada uma das características sonoras são extraídos programaticamente com recurso a um *script* desenvolvido pelo autor que recebe um ficheiro áudio e um ficheiro do tipo *TextGrid* – ficheiro resultante do processo de anotação. Este *script* cria um ficheiro do tipo *Comma Separated Values* (csv) com os valores temporais de início e fim, assim como a duração e os valores numéricos da média da intensidade para cada intervalo

pertencente à camada selecionada. Na Figura 4.4 encontra-se representado um dos ficheiros csv criados pelo *script*.

Palavra	Start	End	Duration	Mean Intensity
Isto	0	0.57	0.57	56.79
#	0.57	1.1	0.53	10.85
nunca	1.1	1.46	0.36	63.27
acaba	1.46	2.05	0.59	53.15

FIG 4.4.
Exemplo do ficheiro csv relativo à camada *Palavras* da gravação sonora c).

O software escolhido para o desenvolvimento dos artefactos nesta fase experimental foi o *Adobe InDesign*. Com recurso à linguagem proprietária *Adobe ExtendScript*, um dialeto variante de *ECMAScript* desenvolvido pela *Adobe Systems* e da ferramenta *Adobe ExtendScript Toolkit*, permite ao utilizador controlar todas as funcionalidades da aplicação recorrendo a *scripts*. Uma vez que a programação nesta linguagem pode ser confusa e de aprendizagem difícil para designers, uma equipa do Instituto de Comunicação Visual em Basel, Suíça, decide em 2013, criar uma biblioteca *JavaScript* chamada *BasilJS*. Esta biblioteca torna a programação em *ExtendScript* mais próxima da programação em *Processing* - linguagem de programação baseada em *Java* - uma vez que esta é já uma linguagem mais simples, direcionada para as artes visuais e de aprendizagem rápida.

Para o primeiro conjunto de testes foi desenvolvido um *script* com recurso à biblioteca *BasilJS* que importa o ficheiro csv resultante da etapa da análise e extração de dados e guarda essa informação no formato de *Array* de objetos. Assim, a cada linha da tabela corresponde um objeto e cada propriedade desse objeto corresponde a uma coluna da tabela. Esta solução permite o acesso a todas as letras, sílabas ou palavras do texto e aos respetivos valores de cada parâmetro a elas associadas.

Os artefactos gerados são apresentados em formato A4 contendo o resultado das oito gravações em simultâneo. Neste caso, a posição de cada frase na página não tem relevância expressiva e serve apenas para efeitos de comparação.

A escolha das características sonoras a ter em conta para o primeiro teste desta fase experimental, centrou-se na intensidade e na

EXPERIMENTAÇÃO INICIAL

duração das pausas e das palavras.

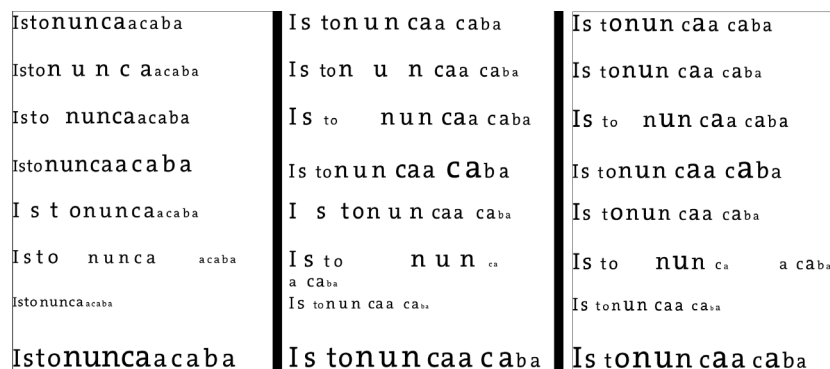
Quanto às variáveis tipográficas a utilizar, optou-se pelo tamanho e peso da fonte e o espaçamento entre caracteres.

Uma vez que um dos objetivos a longo prazo é a representação visual de excertos de discurso mais extensos, a escolha do tipo de letra para a criação dos artefactos nesta fase experimental foi ligeiramente alterada, passando para a versão serifada da superfamília *Thesis*, i.e., *TheSerif*.

RESULTADOS EXPERIMENTAIS

Na Figura 4.5 são apresentados os artefactos gerados pelo *script BEMP_1* (*Baseado na Expressive Music Performance – Versão 1*) aplicado ao nível das palavras, sílabas e letras, respetivamente.

FIG 4.5.
Resultados do *script BEMP_1* ao nível das palavras, sílabas e letras.



TRACKING

Ajuste de espaçamento uniforme entre um conjunto de caracteres.

KERNING

Ajuste de espaçamento entre caracteres individuais.

Como pode ser consultado na Tabela 4.1, os valores da intensidade da voz são mapeados para o tamanho da fonte e a duração de cada palavra, sílaba ou letra é representada na forma de *tracking* e o espaço em branco que representa a duração das pausas é aplicado na forma de *kerning* por questões técnicas. Uma vez que as pausas são marcadas com o caractere “#”, ele é adicionado à caixa de texto assim como todos os outros. Posteriormente, esse caractere é substituído por um caractere vazio e a solução encontrada para aumentar e diminuir esse espaço vazio foi a de variar o valor do *kerning* desse caractere.

De forma a testar a viabilidade do uso do peso da fonte enquanto variável tipográfica, no *script BEMP_2* (*Baseado na Expressive*

MAPEAMENTO PERFORMANCE MUSICAL EXPRESSIVA

Music Performance – Versão 2), os oito pesos disponíveis na família *TheSerif – ExtraLight, Light, SemiLight, Regular, SemiBold, Bold, ExtraBold, Black* – foram usados para mapear, a par do tamanho, a intensidade da voz. Na Tabela 4.2 encontra-se descrito o mapeamento usado pelo *script BEMP_2* na criação dos artefactos apresentados na Figura 4.6.

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração	0.26 - 1.39	<i>Tracking</i>	0 - 500
Duração Pausas	0.26 - 1.39	<i>Kerning</i>	0 - 5000
Intensidade	43.93 - 73.17	Tamanho	12 - 72

TABELA 4.1.
Mapeamento utilizado no *script BEMP_1*.

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração	0.26 - 1.39	<i>Tracking</i>	0 - 500
Duração Pausas	0.26 - 1.39	<i>Kerning</i>	0 - 5000
		Tamanho	12 - 72
Intensidade	43.93 - 73.17	Peso	ExtraLight - Black (8 pesos)

TABELA 4.2.
Mapeamento utilizado no *script BEMP_2*.

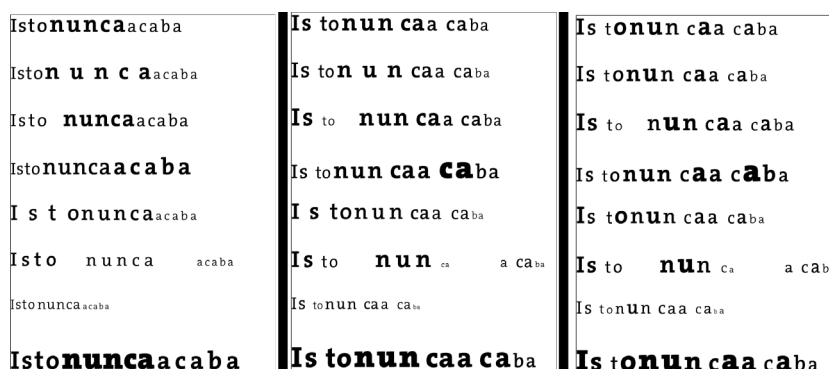


FIG 4.6.
Resultados do *script BEMP_2* ao nível das palavras, sílabas e letras.

Após observação dos artefactos criados através dos *scripts BEMP_1* e *BEMP_2*, constata-se que a solução para representar a duração de uma pronúncia só produz resultados minimamente satisfatórios quando aplicada ao nível das palavras. O uso do *tracking* ao nível das sílabas e das letras gera resultados extremamente ambíguos, uma vez que, visualmente, o que é apresentado pode resultar do ato de soletrar, pois o espaço em branco entre caracteres está a

EXPERIMENTAÇÃO INICIAL

ser usado para representar pausas no discurso. Assim, é tomada a decisão de procurar uma solução diferente para a representação da duração de cada pronúnciação.

De forma a não dar início a uma nova ronda de experimentação sem garantir que a solução usada para a representação da duração das pausas estaria refinada ao ponto de poder ser incluída no conjunto de regras que se procura conceber, foi desenvolvido o *script P_1 (Pausas – Versão 1)*. Na Figura 4.7 são apresentados os artefactos gerados por este script usando o mapeamento presente na Tabela 4.3.

FIG 4.7.
Resultados do *script P_1*
ao nível das palavras,
sílabas e letras.

Isto nunca acaba	Is tonun caa caba	Is tonun caa caba
Isto nunca acaba	Is tonun caa caba	Is tonun caa caba
Isto nunca acaba	Is to nun caa caba	Is to nun caa caba
Isto nunca acaba	Is tonun caa caba	Is tonun caa caba
Isto nunca acaba	Is tonun caa caba	Is tonun caa caba
Isto nunca acaba	Is to nun ca a caba	Is to nun ca a caba
Isto nunca acaba	Is tonun caa caba	Is tonun caa caba
Isto nunca acaba	Is tonun caa caba	Is tonun caa caba

TABELA 4.3.
Mapeamento utilizado
no *script P_1*.

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração Pausas	0.03 – 1.39	<i>Kerning</i>	0 – 5000

Apesar deste *script* ter sido erradamente desenvolvido, como pode confirmar-se pela discrepância entre o tamanho das mesmas pausas nos diferentes artefactos, ao analisar estes resultados foi possível tomar duas decisões fulcrais e definitivas para o desenvolvimento do sistema.

Observando os resultados ao nível das sílabas e das letras, podemos constatar que existe uma pausa, entre a sílaba “Is” e a sílaba “to” da primeira palavra “Isto”, que é significativamente representada, mas, no entanto, impercetível quando ouvidas. Apesar dessa pausa existir realmente e de ser possível detetá-la no momento da anotação, tem uma duração de 0.03 segundos e, por isso, não parece fazer sentido representá-la. Posto isto, após recolha e análise das dura-

MAPEAMENTO PERFORMANCE MUSICAL EXPRESSIVA

ções de todas as pausas presentes nestas gravações, decide-se que não serão consideradas pausas com durações abaixo de 0.1 segundos. Esta decisão veio mais tarde a ser validada aquando da consulta do trabalho desenvolvido por MacArthur, Zellou e Miller no qual os autores afirmam que “pausas com menos de 100 milissegundos não são consideradas, uma vez que fazem naturalmente parte do discurso contínuo” (MacArthur, Zellou, & Miller, 2018).

Uma vez que é possível fazer o mapeamento dos valores obtidos na análise de uma gravação ao nível da letra, não existe razão alguma para que o mapeamento seja feito em níveis de detalhe inferior. Ou seja, ao aplicar as alterações ao nível da letra, estão a ser representadas, automaticamente, alterações ao nível da sílaba e da palavra. Deste modo, fica definido no conjunto de regras do sistema que o mapeamento é sempre feito na camada das letras.

Na Figura 4.8 encontram-se os artefactos gerados pelo *script PC_1* (*Pausas Controladas – Versão 1*) que aplica as regras de mapeamento presentes na Tabela 4.4, ignorando todas as pausas com duração menor que 0.1 segundos.

SOM	VALORES (min. - máx)	TIPOGRAFIA	VALORES (min. - máx)
Duração Pausas	0.1 – 1.39	<i>Kerning</i>	0 – 5000

TABELA 4.4.
Mapeamento utilizado no *script PC_1*.

Istonuncaacaba
Istonuncaacaba
Isto nuncaacaba
Istonuncaacaba
Is tonunca caba
Isto nun ca acaba
Istonuncaacaba
Istonuncaacaba

FIG 4.8.
Artefacto gerado pelo *script PC_1*.

EXPERIMENTAÇÃO INICIAL

Analisando este artefacto, é possível atestar que a aplicação do limiar de 0.1 segundos para que uma pausa seja mapeada produz resultados visuais mais próximos ao que é percebido auditivamente.

Depois de solucionado o problema da representação da duração das pausas, seria imperativo avançar no sentido de encontrar uma resolução para a forma de representar a duração da pronúncia de cada letra. A solução proposta pelo *script PC_REP_1 (Pausas Controladas com Repetição – Versão 1)*, baseia-se numa alternativa usada diariamente de forma intuitiva na escrita em computador, nomeadamente através de serviços de *Instant Messaging*, e-mail e redes sociais – a repetição de caracteres.

Aplicada a solução para a representação da duração de cada letra e analisados os resultados gerados por este *script*, é identificada a estranheza causada pela falta de existência de separação entre cada palavra. Dado que a observação destes artefactos desencadeia, quase imediatamente, um processo de leitura, para que esse não seja desconfortável para o observador, opta-se por acrescentar uma nova regra ao sistema que garanta que existe sempre, no mínimo, um carácter vazio entre duas palavras. Na Figura 4.9 são apresentados os resultados do *script PC_REP_2 (Pausas Controladas com Repetição – Versão 2)* em comparação com o *script PC_REP_1*, ambos baseados no mapeamento descrito na Tabela 4.5.

FIG 4.9.
Artefactos gerados
pelos *scripts PC_REP_1* e
PC_REP_2.

Istonuncaacaba	Isto nunca acaba
Istonuuuuuunnaacaba	Isto nuuuuuunna acaba
Isto nuncaacaba	Isto nunca acaba
Istonuncaacaba	Isto nunca acaba
IIIIIIss tonnuncaa caba	IIIIIIss to nnunca a caba
IISTOO nunn ca acaba	IISTOO nunn ca acaba
Istonuncaacaba	Isto nunca acaba
Istonuncaacaaba	Isto nunca acaaba

MAPEAMENTO PERFORMANCE MUSICAL EXPRESSIVA

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração Pausas	0.1 – 1.39	<i>Kerning</i>	0 – 5000
Duração Letras	0.1 – 1.39	Repetição Caracteres	1 a cada 0.1s

TABELA 4.5.
Mapeamento utilizado nos scripts PC_REP_1 e PC_REP_2.

Concluído o aprimoramento da representação das durações das pausas e da pronúncia de cada letra, é feita a inclusão da representação da intensidade com as duas variáveis tipográficas testadas anteriormente – o tamanho e o peso da fonte.

Na Figura 4.10 são apresentados os artefactos resultantes da aplicação dos scripts *IT_1* (*Intensidade: Tamanho – Versão 1*) e *IP_1* (*Intensidade: Peso – Versão 1*) com base nos mapeamentos descritos na Tabela 4.6 e na Tabela 4.7, respetivamente.

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração Pausas	0.1 – 1.39	<i>Kerning</i>	0 – 5000
Duração Letras	0.1 – 1.39	Repetição Caracteres	1 a cada 0.1s
Intensidade	43.93 – 73.17	Tamanho	12 – 72

TABELA 4.6.
Mapeamento utilizado no script *IT_1*.

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração Pausas	0.1 – 1.39	<i>Kerning</i>	0 – 5000
Duração Letras	0.1 – 1.39	Repetição Caracteres	1 a cada 0.1s
Intensidade	43.93 – 73.17	Peso	ExtraLight – Black (8 pesos)

TABELA 4.7.
Mapeamento utilizado no script *IP_1*.

ANÁLISE DE RESULTADOS

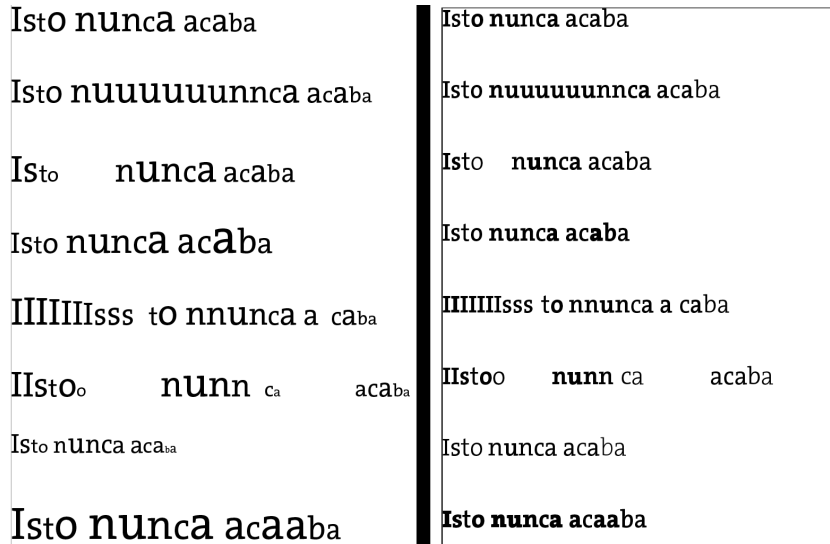
Partindo dos resultados obtidos nesta fase de experimentação é possível confirmar que o tamanho e o peso da fonte poderão ser usados para representar visualmente a intensidade do discurso e que esta é, sem dúvida, uma característica sonora a ter em conta na criação do sistema de regras.

A decisão de aplicar o mapeamento ao nível da letra e de representar a duração da pronúncia através da repetição de caracteres provou ser fundamental para a relação entre o que se ouve nas

EXPERIMENTAÇÃO INICIAL

gravações e o que se vê nos artefactos, assim como a opção de garantir que existe sempre, no mínimo, um espaço em branco entre duas palavras.

FIG 4.10.
Artefactos gerados pelos
scripts IT_1 e IP_1.



MAPEAMENTO PARA «OUVER»

Apesar dos resultados satisfatórios da fase de experimentação anterior, a consideração da intensidade como única característica sonora a representar limitada a potencialidade representativa do sistema. A variação da altura na voz é uma característica essencial do discurso humano possível de ser extraída e é o que permite, por exemplo, distinguir uma afirmação de uma questão.

Nesta fase de experimentação pretende-se tirar proveito do conjunto de regras estabelecidas anteriormente e expandi-lo, explorando alternativas tipográficas para a representação da altura da VOZ.

SETUP EXPERIMENTAL

Para esta fase de experimentação, foi usado um excerto do poema *Amar ou Odiar*, escrito pelo autor português Fausto Guedes Teixeira e interpretado por João Villaret, retirado do álbum *João Villaret*

MAPEAMENTO PARA «OUVER»

no *São Luís* editado em 1961 pela Valentim de Carvalho. A escolha desta gravação deve-se ao facto de conter uma interpretação com grande variação na altura da voz do locutor.

As variáveis tipográficas escolhidas para esta ronda de testes foram a posição, a cor e o tamanho e peso da fonte.

A opção de testar o mapeamento da altura da voz à posição Y da letra, tem que ver com a representação implementada nas pautas musicais nas quais as notas musicais são posicionadas de acordo com a sua altura. Notas mais agudas são posicionadas mais acima e notas mais graves são posicionadas mais abaixo.

RESULTADOS EXPERIMENTAIS

Os primeiros testes na procura da representação visual da altura da voz são realizados recorrendo à posição e à cor.

De forma a ser possível alterar a posição Y de cada carácter, recorreu-se à manipulação dos valores do parâmetro *Baseline Shift* presente no *Adobe InDesign*. Este parâmetro permite-nos aplicar um valor negativo ou positivo para alterar a distância da base da letra à linha de base.

Quanto ao uso da cor para representar a altura, a opção passa por dar uso apenas a uma escala de cinzentos que, programaticamente, varia entre os valores 0 e 255, sendo 0 (preto) usado para representar a voz mais aguda e 255 (branco) para a voz mais grave.

Na Figura 4.11 são apresentados os artefactos resultantes da aplicação dos *scripts IPAYT_1* (*Intensidade: Peso / Altura: Posição Y + Tamanho – Versão 1*) e *IPAC_1* (*Intensidade: Peso / Altura: Cor – Versão 1*) com base nos mapeamentos descritos na Tabela 4.8 e na Tabela 4.9, respetivamente.

LINHA DE BASE

Recta imaginária sobre a qual é feito o alinhamento horizontal das letras.

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração Pausas	0.1 – 0.5	<i>Kerning</i>	0 – 2000
Duração Letras	0.1 – 0.5	Repetição Caracteres	1 a cada 0.1s
Intensidade	39 – 78	Peso	ExtraLight – Black (8 pesos)
Altura	100 – 346	Posição Y	-20 – +20

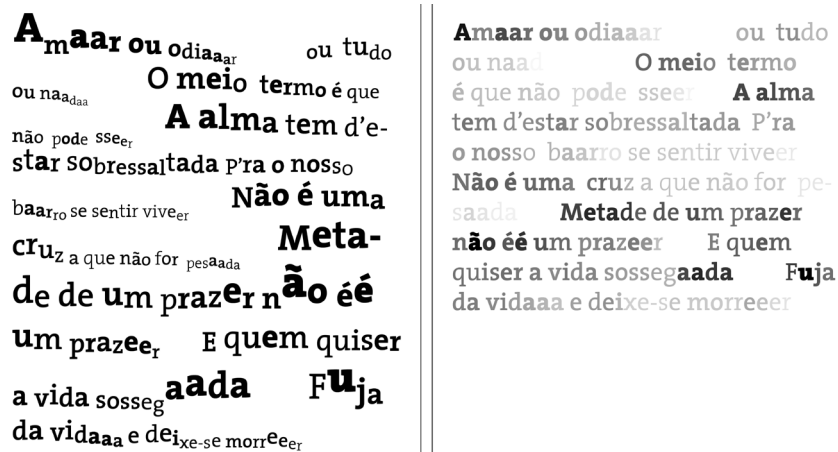
TABELA 4.8.
Mapeamento utilizado no *script IPAYT_1*.

EXPERIMENTAÇÃO INICIAL

TABELA 4.9.
Mapeamento utilizado
no script IPAC_1.

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração Pausas	0.1 – 0.5	<i>Kerning</i>	0 – 2000
Duração Letras	0.1 – 0.5	Repetição Caracteres	1 a cada 0.1s
Intensidade	39 – 78	Peso	ExtraLight – Black (8 pesos)
Altura	100 – 346	Cor	255 – 0 (Escala de Cinzento)

FIG 4.11.
Artefactos gerados pelos
scripts IPAYT_1 e IPAC_1.



No conjunto seguinte de testes, a intensidade e a altura da voz são representadas, alternadamente, pelo tamanho e peso da fonte. Desta forma, procuram-se testar as variáveis tipográficas e as características sonoras mais promissoras e espera-se obter a solução com a ligação mais próxima entre o discurso e o texto.

Na Figura 4.12 são apresentados os artefactos resultantes da aplicação dos scripts *ITAP_1* (*Intensidade: Tamanho / Altura: Peso – Versão 1*) e *IPAT_1* (*Intensidade: Peso / Altura: Tamanho – Versão 1*) com base nos mapeamentos descritos na Tabela 4.10 e na Tabela 4.11.

TABELA 4.10.
Mapeamento utilizado
no script ITAP_1.

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração Pausas	0.1 – 0.5	<i>Kerning</i>	0 – 2000
Duração Letras	0.1 – 0.5	Repetição Caracteres	1 a cada 0.1s
Intensidade	39 – 78	Tamanho	12 – 72
Altura	100 – 346	Peso	ExtraLight – Black (8 pesos)

MAPEAMENTO PARA «OUVER»

SOM	VALORES (mín. - máx)	TIPOGRAFIA	VALORES (mín. - máx)
Duração Pausas	0.1 – 0.5	<i>Kerning</i>	0 – 2000
Duração Letras	0.1 – 0.5	Repetição Caracteres	1 a cada 0.1s
Intensidade	39 – 78	Peso	ExtraLight – Black (8 pesos)
Altura	100 – 346	Tamanho	12 – 72

TABELA 4.11.
Mapeamento utilizado
no script *IPAT_1*.

Amaar ou odiaaa. ou
tudo ou naaaa. O meio ter-
mo é que não pode ssee. A
alma tem d'estar sobressalta-
da pra o noSso baarro se seni.i.ee.
Não é uma cruz a.e não. Or pesaaada
Metade de um prazer não
é um prazeer. E quem quis-
er a vida sossegadaa. Fuja
da vidaaa e deixe-se morreeer.

Amaar ou odiaaar ou tudo
ou naadaa. O meio termo é que
não pode sseer. A alma tem d'es-
tar sobressaltada P'ra o nosso
baarro se sentir viveer. Não é uma
cruz a que não for pesaada. Meta-
de de um prazer não é
um prazeer. E quem quis-
er a vida sossegadaa. Fuja da
vidaaa e deixe-se morreeer.

FIG 4.12.
Artefactos gerados pelos
scripts *ITAP_1* e *IPAT_1*.

Analisando os artefactos obtidos nestes últimos testes, acredita-se ter-se encontrado um conjunto de regras possivelmente válidas para o cumprimento do objetivo desta tese. No entanto, no que toca à composição do texto como um todo, os resultados não são satisfatórios. Pelo facto de se tratar de textos poéticos, existe uma estruturação baseada em estrofes e versos que não está a ser tida em conta e que poderá ter influência na perceção dos artefactos.

Na Figura 4.13 é apresentado o artefacto resultante da aplicação do script *IPAT_O_1* (*Intensidade: Peso / Altura: Tamanho / Estrutura Original - Versão 1*) que respeita o mesmo mapeamento do script *IPAT_1* (Tabela 4.11) com a adição da opção do texto ser composto de acordo com a sua estruturação original.

ANÁLISE DE RESULTADOS

Apesar do extenso leque de possibilidades para o mapeamento da altura, o tamanho e o peso da fonte continuam a ser as variá-

EXPERIMENTAÇÃO INICIAL

veis tipográficas que melhor a representam. Ouvindo a gravação e acompanhando-a com a leitura dos artefactos produzidos, a opção com resultados mais promissores parece ser o mapeamento da intensidade ao peso e da altura ao tamanho da letra.

FIG 4.13.
Artefacto gerado pelo
script IPAT_O_1.

Amaar ou odiaaar ou tudo ou naadaa

O meio termo é que não pode sseer

A alma tem destar sobressaltada

Pra o nosso baarro se sentir viveer

Não é uma cruz a que não for pesaada

Metade de um prazer não é um prazeer

E quem quiser a vida sossegadaa

Fuja da vidaaa e deixese morreeer

Em relação aos scripts desenvolvidos para o *Adobe InDesign*, apesar de cumprirem o seu objetivo, fazem-no de uma forma demorada, uma vez que o seu tempo de execução é de, em média, 66 segundos, dez vezes mais que os scripts desenvolvidos na fase de experimentação anterior. Isto deve-se à complexidade acrescida dos scripts, mas também ao facto de toda a informação de entrada – os ficheiros csv – estar a ser carregada e tratada pelo *Adobe InDesign*. Assim, quanto mais extenso for o texto, mais morosa será a execução do *script*.

Finda esta fase de experimentação, dá-se como concluída a base essencial do conjunto de regras para o mapeamento tipográfico da expressividade do discurso.

«MÁQUINA DE OUVER»

Do conjunto de regras para a representação visual da expressividade do discurso humano pela tipografia consolidado nas fases de experimentação deste projeto, nasce a «Máquina de Ouvir», um sistema que, de forma programática, analisa gravações e converte as suas características sonoras em artefactos tipográficos.

O sistema é constituído por uma fase inicial de *Anotação* onde, de forma manual, é feita uma transcrição e marcação do texto presente na gravação sonora. Seguidamente, essa anotação é usada na fase de *Análise de Som* onde são extraídos e guardados os valores para cada uma das características sonoras correspondentes a cada elemento textual. Após esta fase, é feito o *Tratamento de Dados* que garante que todos os elementos textuais têm um valor válido para cada uma das características sonoras. Desta fase resulta um ficheiro que contém toda a informação necessária para a *Geração de Artefactos* que podem ser em formato de imagem estática ou vídeo.

ANOTAÇÃO

De forma a ser possível produzir dinamicamente os ficheiros com os dados necessários à geração dos artefactos, é necessária uma fase prévia de anotação que pressupõe uma transcrição da gravação sonora.

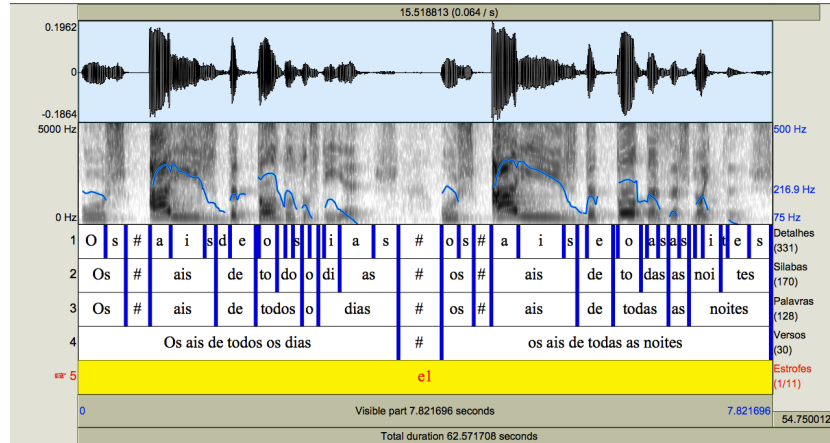
A anotação é estruturada em cinco camadas textuais: *Estrofes*, *Versos*, *Palavras*, *Sílabas* e *Letras*. Na Figura 5.1 é apresentada uma captura de ecrã da interface gráfica do *software Praat* onde pode ser observada a aplicação desta estrutura e da respetiva anotação de um excerto do programa de televisão português dos anos 80 chamado *A Dificuldade está na Escolha – Poesia Portuguesa I* onde o ator e *diseur* português Mário Viegas interpreta o poema *Cantiga dos Ais*, originalmente escrito pelo poeta português Armindo Mendes de Carvalho.

Desta fase resulta um ficheiro do formato TextGrid que guarda a informação temporal de todos os elementos textuais presentes em

«MÁQUINA DE OUVER»

cada um dos intervalos marcados em cada camada da anotação.

FIG 5.1.
Screenshot da interface gráfica do Praat.



ANÁLISE DE SOM

A fim de possibilitar a extração dos valores da intensidade e da altura dos elementos presentes em cada uma das camadas anotadas na fase de *Anotação*, foi desenvolvido um *script* para o *Praat* que recebe como inputs o ficheiro de som a analisar e o ficheiro *TextGrid* resultante da fase de *Anotação*.

Para além de extrair os valores da intensidade e da altura e guardar os dados temporais de cada um dos intervalos, este *script* contabiliza o número de repetições que cada caractere deve sofrer baseando-se na sua duração, como exemplificado na Figura 5.2.

Desta fase resultam cinco ficheiros do formato csv com toda a informação acerca do ficheiro sonoro analisado, um por cada camada de anotação – *Estrofes*, *Versos*, *Palavras*, *Sílabas* e *Letras*.

FIG 5.2.
Amostra do ficheiro csv.

Detalhe	Y	Start	End	Duration	Intensity	Pitch
o		16.17	16.32	0.15	61.99	199.95
s		16.32	16.52	0.19	47.8	158.9
#		16.52	16.8	0.28	31.8	--undefined--
a		16.8	17.18	0.13	67.19	250.64
_a		16.8	17.18	0.13	70.75	245.23
_a		16.8	17.18	0.13	69.06	171.2
i		17.18	17.69	0.1	65.73	289.08
_i		17.18	17.69	0.1	67.05	275.63

TRATAMENTO DE DADOS

Depois da análise efetuada na fase anterior, é necessário um tratamento dos dados recolhidos, uma vez que nem sempre é possível ao *Praat* efetuar uma medição realista dos valores da altura de certos elementos textuais o que resulta na atribuição de valores erróneos ou da atribuição do valor *undefined*. Para tal, foi desenvolvido um script em *JavaScript* que, numa primeira instância, combina os cinco ficheiros *csv* resultantes da etapa de *Análise de Som* num único ficheiro do tipo *JavaScript Object Notation (JSON)* que possibilita o acesso a qualquer elemento textual da gravação: *Estrofes, Versos, Palavras, Sílabas* ou *Letras*.

Depois de uma análise manual prévia, são fornecidos ao *script*, os valores mínimos e máximos para a duração das pausas, intensidade e altura.

Com base nesta informação, caso o valor da altura de um certo elemento seja mais alto que o máximo indicado ou mais baixo que o mínimo, são convertidos para *undefined*.

Seguidamente, o script percorre todos os elementos do objeto procurando identificar os que guardam o valor *undefined* e, uma a uma, verifica se existe uma letra anterior com um valor válido para a altura, caso exista, iguala ao anterior.

Se esta condição não for verdadeira, iguala à letra na posição seguinte, caso exista.

Uma vez não se verificando esta condição, procura pela palavra anterior e iguala à média da altura dessa palavra.

Quando nenhuma das condições anteriores for respeitada, o valor para a altura dessa letra é igualado à altura da palavra seguinte.

Finalmente é gerado um ficheiro *JSON* com toda a estrutura textual, os seus valores correspondentes e a sua *metadata*, como, por exemplo, o nome do poema, o seu autor, o ano em que foi escrito e o intérprete.

Neste objeto *JSON* encontram-se também guardados os valores mínimos e máximos para a duração das pausas, intensidade e altura a fim de serem usados futuramente no mapeamento para as variáveis tipográficas.

GERAÇÃO DE ARTEFACTOS

De forma a materializar o mapeamento proposto, foram identificadas duas opções: artefactos estáticos, como uma imagem ou poster, ou artefactos dinâmicos em forma de vídeo. Assim, foram desenvolvidas duas versões de um mesmo *script* para gerar os dois tipos de artefactos no *software Adobe InDesign*, sendo que no segundo caso, é necessária uma etapa acrescida de criação do vídeo que é responsabilidade de um *script* desenvolvido para o *software Adobe After Effects*.

ARTEFACTOS ESTÁTICOS

Inicialmente, são definidos os valores mínimos e máximos para o mapeamento da intensidade para tamanho da fonte, a família tipográfica a ser utilizada e os respetivos pesos a serem mapeados para os valores da altura, o caminho para o ficheiro *JSON* resultante da etapa anterior e as dimensões da página que resultará no artefacto final.

Depois de fornecida a informação necessária, o *script* percorre todos os elementos textuais por camadas, da mais geral – *Estrofes* – à mais pormenorizada – *Letras*. Sempre que é encontrado um elemento textual com o valor de “#”, o *script* reconhece como sendo uma pausa e, caso seja na camada das *Estrofes* ou dos *Versos*, essa pausa é representada na forma de entrelinha. Os valores mínimo e máximo para a duração das pausas são mapeados para um valor mínimo e máximo definido pelo utilizador e o cálculo dessa entrelinha é feito com base nesse valor multiplicado pelo valor da entrelinha inicial – 20% do valor médio entre o tamanho mínimo e máximo da fonte. Para cada verso com informação textual diferente de “#”, é criada uma nova caixa de texto cuja posição em relação à caixa de texto anterior é dada pelo resultado do cálculo da entrelinha despoletado pela pausa que a antecede.

Assim que o *script* encontra um *Verso* válido, percorre os seus elementos textuais na camada das *Letras* e sempre que encontra uma pausa marcada pelo caractere “#” aplica o mapeamento da duração dessa pausa a um valor entre um mínimo e máximo de

GERAÇÃO DE ARTEFACTOS

caracteres vazios – “ ” – atribuído previamente pelo utilizador. Para cada caractere diferente de “#”, é aplicado o mapeamento aos valores da intensidade e da altura para os do peso da fonte e o seu tamanho, respetivamente.

Finalmente, depois de percorrido todo o ficheiro *JSON*, o artefacto final é exportado no formato pretendido.

Na Figura 5.3, é apresentado o artefacto final, com dimensões de impressão A4, da aplicação do *script* com as regras de mapeamento presentes na Tabela 5.1, a um excerto do programa de televisão português dos anos 80 chamado *A Dificuldade está na Escolha – Poesia Portuguesa I* onde o ator e *diseur* português Mário Viegas interpreta o poema *Cantiga dos Ais*, originalmente escrito pelo poeta português Armindo Mendes de Carvalho.

OOOss **aa**iiis deee **to**odos os diaaass
os **aa**aiiiiis deee **to**odas as noiiteess
AAiiis do ffaa doo e do folclooree
o **aa**i de **óó** **aa**ii **ó**llindaa
os **aa**aiiiiiss que vvêem do pei too
os **aaa**aiiiiis que vêem da aalmaa
Os **aa**aiiiiiss que vêem do ssexoo
os **aaaa**aiis do prazee na camaa
AAAi **po**obre da quele velhiinho
AAAii que saudaaaaades meniinna
aaaii a velhiice é tão triiiSStee
AAiiis do pei to e da poesiaa
e os **ai**s de **ou**tras **co**isas **ma**aiiiiis
os **ai**iiiis da vidaaaa e da mortee
Aii ooss **ai** ssss deste paí sss

FIG 5.3.
Artefacto gerado pelo
script final.

«MÁQUINA DE OUVER»

TABELA 5.1.
Mapeamento utilizado
no *script* final.

SOM	VALORES (min. - máx)	TIPOGRAFIA	VALORES (min. - máx)
Duração Pausas (entre Palavras, Sílabas ou Letras)	0.1 – 1.7	Repetição Caracteres Vazios	0 – 30
Duração Pausas (entre Estrofes ou Versos)	0.1 – 1.7	Entrelinha	0 – 30*
Duração Letras	0.1 – 1.7	Repetição Caracteres	1 a cada 0.1s
Intensidade	29 – 72	Peso	ExtraLight – Black (8 pesos)
Altura	75 – 350	Tamanho	9 – 36

*multiplicado por 20% do valor médio entre o mínimo e o máximo estabelecido para o tamanho da fonte

ARTEFACTOS DINÂMICOS

A solução adotada para a automatização da criação de artefactos em formato vídeo passa pela modificação do *script* desenvolvido para a criação de artefactos estáticos, no sentido de exportar uma imagem do estado atual da página a cada letra ou alteração acrescentada. Deste modo, é calculada a quantidade de imagens necessária a exportar de cada alteração ao texto com base na sua duração. Assim, assumindo que o vídeo final será produzido com um frame rate de 30 frames por segundo, se uma determinada letra tem a duração de 1 segundo, significa que o *script* exporta trinta vezes a imagem correspondente a essa alteração o que corresponde a 1 segundo de vídeo.

Para ser possível apresentar o texto composto no seu estado original, de forma neutra com o seu valor inicial para o tamanho e peso da fonte e entrelinha e, no decorrer do vídeo, apresentar as alterações que são feitas pelo discurso ao texto de forma sincronizada com a gravação sonora, a solução proposta nesta versão do *script* é a de percorrer o texto já alterado, no seu estado final, a partir da última letra para a primeira, revertendo todo o processo e exportando a quantidade de imagens necessárias pelo método explicado anteriormente.

Nas páginas 66 e 67 é apresentado, do lado esquerdo, o primeiro *frame* do vídeo – último a ser exportado pelo *script* – e, do lado direito, o último *frame* do vídeo – primeiro a ser exportado pelo *script*.

Seguidamente, foi desenvolvido um *script* para o *Adobe After*

GERAÇÃO DE ARTEFACTOS

Effects que seleciona todas as imagens geradas na fase anterior, importando-as como uma sequência de imagens e a gravação sonora correspondente. Finalmente, o *script* cria uma nova composição com a sequência de imagens e a gravação sonora, exportando-o no formato de vídeo mp4 com um *frame rate* de 30 *frames* por segundo.

O artefacto dinâmico final encontra-se disponível em https://student.dei.uc.pt/~jccastro/cantiga_dos_ais.mp4

Os ais de todos os dias
os ais de todas as noites
Ais do fado e do folclore
o ai de ó ai ó linda
os ais que vêm do peito
os ais que vêm da alma
Os ais que vêm do sexo
os ais do prazer na cama
Ai pobre daquele velhinho
Ai que saudades menina
ai a velhice é tão triste
Ais do peito e da poesia
e os ais de outras coisas mais
Os ais da vida e da morte
Ai os ais deste país

OOOss **aaiiis** de **todo**s os diaaass

os **aaaiiiiis** de **tood**as as noit^{te}ss

AAiis do ffaad^oo e do folclo^oree

o **aai** de **óó** **aai** **ó llin**daa

os **aaaiiiiiss** que vvêem do pei too

os **aaaaaiiis** que vêem da aalmaa

Os **aaaiiiiiss** que vêem do ssexoo

os **aaaaaiis** do prazer na camaa

AAAi **poobre** da quele velhiinho

AAAi que sauda^{aa}ades meni^{inna}

aaai a velhiice é tão triiis^{Stee}

AAiiiss do pei to e da poes^{iaa}

e os **ais** de **ou**tras **coi**sas ma^{aiiis}

os **aiiis** da vid^{aaa} e da mort^{ee}

Aii ooss **ai** ssss deste paí sss

AValiação DO SISTEMA

Com o objetivo de testar o sucesso e eficácia dos artefactos produzidos pelo sistema apresentado, optou-se pela realização de um curto inquérito *online* – *acessível em* <https://student.dei.uc.pt/~-jccastro/ouver> – de resposta rápida e descomplicada. Para tal, o inquérito foi estruturado em três partes: apresentação do projeto e objetivos do inquérito, questões de correspondência entre artefactos estáticos e gravações sonoras e, por último, questões de classificação de artefactos dinâmicos. Depois de apresentado o projeto e os objetivos do questionário numa página inicial, é pedido ao utilizador que forneça alguma informação pessoal como a idade e o sexo e é pedido que, numa escala de 1 a 5, classifique o seu grau de conhecimento na área de Design ou Tipografia e de Música ou Som. Com esta informação pretende-se efetuar uma análise no sentido de averiguar se o grau de sucesso do sistema depende do grau de conhecimento nas áreas abordadas.

O grupo de questões de correspondência é constituído por dois subgrupos, cada um dos subgrupos apresenta duas questões. No primeiro subgrupo é apresentada uma imagem e pede-se ao utilizador que escolha um de três sons possíveis que corresponda a ela. No segundo subgrupo a questão é feita de forma inversa, sendo apresentada uma gravação sonora e pedido ao utilizador que escolha uma de três imagens possíveis. Em qualquer uma das questões deste grupo existe sempre uma e apenas uma resposta correta. De forma a diminuir o enviesamento das respostas, as duas questões de cada subgrupo são retiradas de um total possível de oito. Com este grupo de questões espera-se obter a quantidade de respostas corretas e erradas a fim de averiguar o grau de sucesso do sistema.

No segundo grupo são apresentadas duas soluções possíveis na geração de artefactos dinâmicos na forma de vídeo e pede-se ao utilizador que classifique, de 1 a 5, qual, no seu entendimento, é o grau de ligação entre a imagem e o som apresentado em cada um dos vídeos. O *Vídeo 1* trata-se de um artefacto dinâmico no qual o texto é apresentado com as devidas alterações visuais à medida que

AVALIAÇÃO DO SISTEMA

o som avança, de forma sincronizada. O *Vídeo 2* consiste num artefacto dinâmico no qual é apresentado, desde o primeiro momento, o texto original com formatação e composição neutra, à medida que o som avança, são apresentadas as alterações correspondentes de forma sincronizada. Com esta questão pretende-se averiguar qual a alternativa mais eficaz na transmissão da ligação entre a imagem e o som, com base na preferência dos inquiridos.

Ao longo de todo o questionário foi disponibilizado um campo de texto no qual o utilizador poderia deixar comentários ou justificar as suas respostas. Todos os comentários e sugestões podem ser consultados no Apêndice D.

Uma vez que a aleatoriedade das perguntas e respostas não foi possível de implementar em serviços de questionários como o *TypeForm* ou *Google Forms*, o questionário foi materializado numa página web desenvolvida em *HTML*, *CSS* e *JavaScript*, com recurso à *framework VUEJS*, a partir da qual toda a informação foi enviada para uma base de dados alojada nos serviços *Google Firebase*.

No Apêndice B encontram-se capturas de ecrã de cada uma das páginas do questionário online.

RECOLHA E ANÁLISE DE RESULTADOS

O questionário esteve acessível online entre os dias 15 e 31 do mês de agosto de 2019, período durante o qual foram recolhidas 128 respostas (Apêndice C), sendo que 77 são de indivíduos do sexo masculino e 51 do sexo feminino, com uma média de idades de 30 anos. Na Figura 6.1, à esquerda, é apresentado o gráfico circular para a distribuição do grau de conhecimento em Design ou Tipografia e, à direita, o mesmo tipo de gráfico para a distribuição do grau de conhecimento em Música ou Som dos inquiridos. Em ambos os casos, o nível 1 corresponde a não ter qualquer tipo de conhecimento a nível teórico ou prático nas áreas, sendo que o nível 5 pressupõe ser profissional nas áreas ou ser detentor de um curso ligado a elas. Da análise destes gráficos podemos constatar que no conjunto total de respostas existe um maior conhecimento especializado na área de Design ou Tipografia do que na área de Música ou Som. No entanto, a percentagem de pessoas com grau

RECOLHA E ANÁLISE DE RESULTADOS

de conhecimento inferior 3 na área de Design ou Tipografia é de 44,1% e na área de Música ou Som é de 49,2% pelo que podemos sugerir que temos uma amostra equilibrada em ambas as áreas.

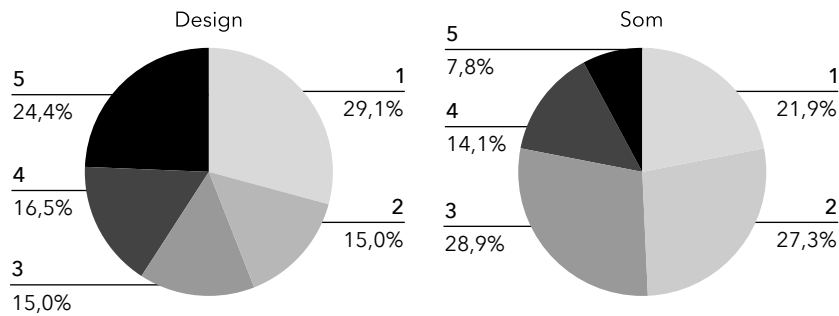


FIG 6.1. Distribuição dos graus de conhecimento nas áreas de Design ou Tipografia e Som ou Música

No primeiro grupo de respostas foram registadas 400 respostas corretas, de um total de 512, sendo que 189 foram no subgrupo *Imagem – Sons* e 211 no subgrupo *Som – Imagens*, como é possível observar na Tabela 6.1. Uma vez que o conteúdo possível a apresentar nas perguntas e respostas é o mesmo nos dois subgrupos, o maior número de respostas corretas no subgrupo em que é apresentado um ficheiro de som e o utilizador tem que escolher uma de três imagens possíveis, sugere que a possibilidade de contextualizar uma imagem e de a comparar com outras duas, pode facilitar na identificação da resposta correta. Esta diferença pode sugerir também ser mais fácil identificar alterações visuais do que alterações sonoras, quando acompanhadas por mais do que uma opção.

	CORRETAS (%)	ERRADAS (%)
Subgrupo Imagem – Sons	73,8	26,2
Subgrupo Som – Imagens	82,4	17,6
TOTAL	78,1	21,9

TABELA 6.1. Percentagens de respostas corretas e erradas no primeiro grupo de questões.

Aplicando o teste *chi square*, definindo a hipótese nula como *As respostas foram dadas de forma aleatória*, verifica-se que para um grau de confiança de 0,05% o valor crítico seria de 3,841 – segundo tabela consultada em <https://www.itl.nist.gov/div898/handbook/eda/section3/eda3674.htm> – e, para estes dados, o valor x^2 obtido

AVALIAÇÃO DO SISTEMA

foi de 25,45. Uma vez que este valor é superior ao tabelado, encontram-se reunidas as condições para rejeitar a hipótese nula. Assim, sendo que o número de respostas corretas é de, aproximadamente, 80% e que, como provado pelo teste anterior, as respostas não foram dadas de forma aleatória, pode sugerir-se que o conjunto de regras para o mapeamento do som através da tipografia e a sua aplicação através do sistema desenvolvido nesta tese produz resultados eficazes e com sucesso.

De forma a analisar os dados com maior detalhe, considerou-se que indivíduos com graus de conhecimento de 3, 4 ou 5 são categorizados como *Com Conhecimentos Mínimos* e os indivíduos com graus de conhecimento de 1 ou 2 como *Sem Conhecimentos Mínimos*. Deste modo, foram criados dois grupos: *Com Conhecimentos Mínimos em Design e Som (CCMDS)* e *Sem Conhecimentos Mínimos em Design e Som (SCMDS)*.

TABELA 6.2.
Percentagens de respostas corretas e erradas no primeiro grupo de questões – CCMDS vs. SCMDS

	CORRETAS (%)	ERRADAS (%)
CCMDS	81,3	18,7
SCMDS	72,6	27,4

Observando a Tabela 6.2 é possível verificar que existe uma percentagem superior de respostas corretas no grupo com conhecimentos mínimos tanto em design como em som, o que poderá sugerir que indivíduos com maior conhecimento nas áreas abordadas por este projeto tenham mais facilidade em identificar a ligação entre as variações nas características sonoras e as alterações visuais na tipografia. No entanto, a percentagem de respostas corretas no grupo sem conhecimentos mínimos em design ou som é de 72,6% o que, mesmo assim, é considerado um resultado satisfatório.

Na Figura 6.2 pode ser observada a distribuição das classificações atribuídas aos dois artefactos dinâmicos apresentados no segundo grupo, sendo que a classificação 1 corresponde à não existência de ligação entre a imagem e o som e a classificação 5 corresponde a uma existência clara de ligação entre o som e a imagem. Das 128 respostas, o *Vídeo 1* obteve uma classificação média de 4,0 enquanto que o *Vídeo 2* obteve uma classificação média de 4,7.

RECOLHA E ANÁLISE DE RESULTADOS

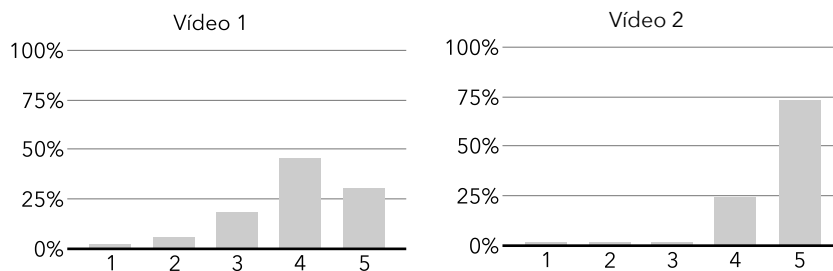


FIG 6.2. Gráficos de Barras – Distribuição do grau de ligação som / imagem atribuído no segundo grupo de questões.

Apesar de uma primeira análise superficial aos resultados obtidos no segundo grupo sugerir que existe uma tendência a preferir a segunda opção, de forma a confirmar se os dados obtidos nos permitem corroborar esta hipótese, foi executado um teste estatístico de hipótese não paramétrico – *Wilcoxon signed-rank test*. Para tal, recorreu-se a uma calculadora online disponibilizada pela plataforma *Social Science Statistics* em www.socscistatistics.com. Os resultados obtidos neste teste fundamentam a afirmação de que existe uma preferência, por parte dos inquiridos, pela segunda alternativa apresentada.

DISSEMINAÇÃO DO TRABALHO

Foram várias as oportunidades aproveitadas para a divulgação dos resultados obtidos no seguimento do desenvolvimento do trabalho prático apresentado nesta dissertação, quer no âmbito científico e académico, quer num âmbito mais abrangente direccionado ao público em geral.

Numa fase final do processo de experimentação inicial, recorreu-se à «Máquina de Ouver» para gerar um artefacto dinâmico para concorrer ao concurso de vídeo do *Festival Internacional Typomania* em Moscovo, acabando por ser o primeiro trabalho português na história do festival a arrecadar um dos prémios.

Seguidamente foi redigido e submetido o artigo científico «Máquina de Ouver» — *From Sound to Type: Finding the Visual Representation of Speech by Mapping Sound Features to Typographic Variables* para a 9ª Conferência Internacional em Artes Digitais e Interativas - *Artech 2019* – acabando por ser aceite, será apresentado entre os dias 23 a 25 de outubro de 2019 em Braga.

Tendo sido seleccionado pelos órgãos de coordenação dos cursos de Licenciatura e Mestrado em Design e Multimédia da Faculdade de Ciências e Tecnologia da Universidade de Coimbra para ser um dos trabalhos a representá-los na 1ª Bienal de Design do Porto, de 22 de novembro de 2019 a 20 de abril de 2020, terá a oportunidade de integrar a exposição *Y, Desenhar Portugal — Projetos de Escolas de Design Nacionais*, a ter lugar na Galeria Municipal de Matosinhos, com curadoria de Francisco Providência, Joana Quental e Rui Costa.

Foi ainda submetida a proposta para *short talk* no 10º Encontro de Tipografia, a acontecer entre os dias 21 e 23 de novembro de 2019, na Escola Superior de Artes e Design de Matosinhos. Esta publicação aguarda decisão de aceitação.

Entre os dias 26, 27 e 28 de julho de 2019, em representação do CDVLab, a «Máquina de Ouver» marcou presença em Gouveia, no evento *GO Romaria Cultural*.

CONCLUSÃO

Depois de um longo e exaustivo processo de experimentação e da procura incessante pela possibilidade de «ouvir», partindo da análise das características sonoras da voz humana através de gravações e recorrendo a técnicas computacionais, os resultados obtidos nesta dissertação são muito satisfatórios e promissores.

Embora num estado de desenvolvimento que pode ainda ser considerado embrionário, os artefactos estáticos e dinâmicos gerados por esta «Máquina de Ouvir» corroboram a hipótese inicial de que é possível transmitir a expressividade do discurso humano partindo da análise de características sonoras como a intensidade e a altura, quando aliadas ao fator da cadência temporal da fala na forma da duração de cada pronúnciação.

O aproveitamento de soluções como a repetição de caracteres, o uso do espaço em branco e a exploração da forma da letra sem a sua alteração anatómica, com foco em variáveis tipográficas como o tamanho e o peso da letra apresentam-se como fatores chave para a simplicidade interpretativa dos artefactos gerados.

Sendo que uma das aplicações de um sistema automático deste género passa pela visualização da expressividade do discurso em tempo real, a transcrição e anotação manual são uma limitação significativa. Trabalho futuro neste sentido, poderá tirar benefícios da exploração e implementação de métodos e ferramentas alternativas de *speech-to-text* e *forced alignment*.

Futuramente, o conjunto de regras estabelecido nesta dissertação poderá evoluir na forma de uma fonte variável reativa, tirando proveito da sua implementação em ferramentas online.

BIBLIOGRAFIA

- Baines, P., & Haslam, A. (2005). *Type & Typography*. Retrieved from <https://books.google.pt/books?id=rs6gyb2hPF4C>
- Ballou, G. (1987). *Handbook for Sound Engineers: The New Audio Cyclopedia*. Indiana: Howard W. Sams & Company.
- Boersma, P., & Weenink, D. (2009). *Praat: doing phonetics by computer (Version 5.1.13)*. Retrieved from <http://www.praat.org>
- Bringhurst, R. (2015). *The Elements of Typographic Style* (Fourth Edi). Hartley & Marks Publishers.
- Cullen, K. (2012). *Design Elements Typography Fundamentals : A Graphic Style Manual for Understanding How Typography Affects Design*. Beverly: Rockport Publishers.
- Denise Schmandt-Besserat. (2015). Evolution of Writing. *International Encyclopedia of the Social & Behavioral Sciences, 2nd Editio*, 761–766. <https://doi.org/https://doi.org/10.1016/B978-0-08-097086-8.81062-4>
- Dixon, S., Goebel, W., & Widmer, G. (2002). *Real Time Tracking and Visualisation of Musical Expression*. 2445. https://doi.org/10.1007/3-540-45722-4_7
- Drucker, J. (1994). *The Visible Word: Experimental Typography and Modern Art, 1909-1923*. London: The University of Chicago Press.
- Fellows, K. S. (2009). *Typecast: the voice of typography*. University of Iowa.
- Ford, S., Forlizzi, J., & Ishizaki, S. (1997). Kinetic Typography: Issues in time-based presentations of text. *CHI '97*.
- Gabrielsson, A. (1988). Timing in music performance and its relations to music experience. In *Generative processes in music: The psychology of performance, improvisation, and composition*. (pp. 27–51). New York, NY, US: Clarendon Press/Oxford University Press.
- Goebel, W., Pampalk, E., & Widmer, G. (2004). *EXPLORING EXPRESSIVE PERFORMANCE TRAJECTORIES: SIX FAMOUS PIANISTS PLAY SIX CHOPIN PIECES*.

- Guimarães, I. (2007). *A Ciência e a Arte da Voz Humana* (ESSA, Ed.). Alcabideche.
- Guimarães, I. (2011). A Diversidade das Manifestações da Voz Falada. *Lumen Veritas*, xvii(1), 3–4.
- Hillner, M. (2009). *Basics Typography 01: Virtual Typography*. Retrieved from <https://books.google.pt/books?id=xGISlV2AwloC>
- Janacek, G. (1984). *The Look of Russian Literature*. Princeton, NJ: Princeton University Press.
- Jang, J. (n.d.). *Audio Signal Processing and Recognition*.
- Jury, D. (2006). *What is Typography?* RotoVision SA.
- Juslin, P. (1997). Emotional Communication in Music Performance: A Functionalist Perspective and Some Data. *Music Perception: An Interdisciplinary Journal*, 14(4), 383–418.
- Kane, J. (2011). *A Type Primer* (Second Edi). London: Laurence King Publishing Ltd.
- Langner, J., & Goebel, W. (2003). Visualizing Expressive Performance in TempoLoudness Space. *Computer Music Journal - COMPUT MUSIC J*, 27, 69–83. <https://doi.org/10.1162/014892603322730514>
- Lupton, E. (2004). *Thinking with Type: A Critical Guide for Designers, Writers, Editors & Students* (2nd Editio). Princeton Architectural Press.
- MacArthur, M., Zellou, G., & Miller, L. (2018). Beyond Poet Voice: Sampling the (Non-) Performance Styles of 100 American Poets. *Journal of Cultural Analytics*. <https://doi.org/10.22148/16.022>
- Massin, R. (1970). *La lettre et l'image*. Retrieved from <https://books.google.pt/books?id=89ZVwwEACAAJ>
- Meggs, P. B., & Purvis, A. W. (2012). *Meggs' History of Graphic Design* (Fifth Edit). New Jersey: John Wiley & Sons, Inc.
- Piqueira, G. (2018). *A Cantora Careca de Massin*. São Paulo: Lote 42.
- Rosenberger, T., & MacNeil, R. L. (1999). Prosodic Font: Translating speech into graphics. *CHI '99*.
- Sonnenschein, D. (2001). *Sound design : the expressive power of music, voice, and sound effects in cinema / by David Sonnenschein*.

BIBLIOGRAFIA

Retrieved from <http://www.loc.gov/catdir/enhancements/fy0715/2001035804-d.html>

- Truax, B. (2001). *Acoustic Communication*. Westport: Ablex.
- Tschichold, J. (1991). *The Form of The Book: Essays on the morality of good design*. London: Lund Humphries.
- Wölfel, M., Schlippe, T., & Stitz, A. (2015). Voice Driven Type Design. *SpeD 2015 - The 8th International Conference on Speech Technology and Human-Computer Dialogue*.

APÊNDICES

A - ARTEFACTO COMPLETO RESULTANTE DO MAPEAMENTO MANUAL INTUITIVO

“Veeem por aqui!!!”

dizem-me alguns com olhos **dóces** estendendo-me os **braços**
e seguros de que seria bom que eu os ouvisse quando me dizem...

“**VEEM** por aqui!!!”

Eu **oolho**-os com olhos **laassooos**,
há nos meus olhos **ironiiias** e **cannis** **saços** e cruzo os **braços**, e não vou por ali...

a minha **GLÓÓRIA É ESSAAA!**

- Criar **desumanidade**, não acompanhar **ninguém**
Que eu vivo com o **mesmo** sem-vontade com que **rrrrrasguei** o **veentre** a minha mãe

Nãããã, não vou por **aíii**
Só vou por onde me levam **mmmeus** **próprios** **pppassos**...

Se às coisas que eu pergunto **eeem** vão ninguém responde

Porque me dizeis **vóóóssss**: “**VEM** por aqui!!!”?

Prefiro **escorrreg**ar nos **bbecos** lamacentos, **rrrrredemoinhar** aos **vventos**,
como **farrrrap**os **arrrrrr**astar os pés **sannngrentos**, a **iiiiir** por **aíiii**...

Se eu vim ao mundo, foi só para desbravar florestas virgens,
e desenhar meus **próprios** **passos** na areia inexplorada!
O mais que faço **nnão** vale nada.

Como pois sereis vós que me dareis **ferrrr**amentas, machados e coragem
Para eu **derrrrubar** os **mmmeus** **obsssstáculos**?...

COORRRE nas vossas veias sangue **velho** dos avós,
E **mós** amais o que é **fffácil**!

Eu amo o Longe e a Miragem,

Amo os abiiismos, as torrrreentes, os deseertos...

IIIDEEE!

Tendes jardins,
tendes canteiros,
tendes estradas,
tendes pátrias,
tendes tectos,

E **tendes** **livros**, e **tratados**, e **filósofos**, e **sábios**...

**EU TENHO A MINHA LOUCURA!
LEVANTO-A COMO UM FACHO A ARRRDERRR NA NOITE ESCURA
E SINTO ESPUUMA E SSSAAANGUEEE,
E CÂÂÂÂNTICOS NOS LÁBIOSSS.
DEEUS E O DIABO É QUEM ME GUIA...**

MAAIS NNINGUÉÉÉMM!

TOODOS TIVERAM PAAAI...

TOODOS TIVERAM MÃEEE...

**MAS EU QUE NUNCA PRINCIPIO NEM ACABO,
NASCI DO AMOOOR QUE HÁ ENTRE DDDEUS E O DIABO.**

AH, QUE NINGUÉM ME DE PIEDOSAS INTENÇÕES
NÃO ME PEÇAM DEFINIÇÕES!
NINGUÉM ME DIGA:

“VEEEM POR AQUI!”

A MINHA VIDA É UM VENDAVAL QUE SE SOLTOU.
É UMA ONDA QUE SE ALEVANTOU.
É UM ÁTOMO A MMAISS QUE SE ANIMOU...

**NÃO SEI POR ONDE VOU,
NÃO SEI PARA ONDE VOOOU**

- SEI QUE NÃO VOU POR AÍ!

APÊNDICES

B - IMAGENS DO QUESTIONÁRIO

APÊNDICE B – IMAGENS DO QUESTIONÁRIO

Máquina de Ouvir — Question: X +

student.dei.uc.pt/~jcastro/ouvir/

>{
Máquina
de Ouvir
}

COMPUTATIONAL
DESIGN &
VISUALIZATION
LAB.

A «Máquina de Ouvir» é um sistema que analisa gravações sonoras de discurso humano e cria uma representação visual para a sua expressividade, dando uso a variáveis tipográficas.

O presente questionário serve para testar a eficácia do sistema.

Este trabalho está a ser desenvolvido por João Couceiro e Castro no âmbito da sua Tese de Mestrado em Design e Multimédia pela Faculdade de Ciências e Tecnologia da Universidade de Coimbra, sob orientação de Penousal Machado, Ana Boavida e Pedro Martins.

Para mais informações sobre o projecto, visite-nos em cdv.dei.uc.pt.

AVANÇAR*

* Ao clicar em "Avançar", está a aceitar participar neste estudo, dando permissão para que os dados fornecidos por si sejam alvo de análise, confiando em que apenas serão utilizados para esta investigação.

Máquina de Ouvir — Question: X +

student.dei.uc.pt/~jcastro/ouvir/

Por favor, preencha os seus dados pessoais. 1/7

Idade Sexo

Qual o seu grau de conhecimento na área do Design ou Tipografia?

Não tenho qualquer tipo de conhecimento Teórico ou Prático na área do Design ou Tipografia. 1 2 3 4 5

Designer Profissional. Professor ou Tipógrafo. Tenho um curso na área do Design ou Tipografia.

Qual o seu grau de conhecimento na área da Música ou Som?

Não tenho qualquer tipo de conhecimento Teórico ou Prático na área da Música ou Som. 1 2 3 4 5

Músico Profissional. Professor, Artista ou Técnico. Tenho um curso na área da Música ou Som.

ANTERIOR

Se desejar, pode justificar aqui a sua resposta.

SEGUINTE

APÊNDICE B – IMAGENS DO QUESTIONÁRIO

Máquina de Ouver — Question: X +

student.dei.uc.pt/~jccastro/ouver/

Escolha o excerto sonoro correspondente à imagem apresentada: 2/7

IIIIIIIss to nunca a ba.

Som 1 Som 2 Som 3

ANTERIOR

Se desejar, pode fazer um comentário ou justificar aqui a sua resposta.

SEGUINTE

Máquina de Ouver — Question: X +

student.dei.uc.pt/~jccastro/ouver/

Escolha o excerto sonoro correspondente à imagem apresentada: 3/7

Isto nunca acaba.

Som 1 Som 2 Som 3

ANTERIOR

Se desejar, pode fazer um comentário ou justificar aqui a sua resposta.

SEGUINTE

APÊNDICE B – IMAGENS DO QUESTIONÁRIO

Máquina de Ouvir — Question: x +

student.dei.uc.pt/~jcastro/ouvir/ ☆ ⓘ

Escolha a imagem correspondente ao excerto sonoro apresentado: 4/7

▶ CLIQUE AQUI PARA OUVIR O EXCERTO SONORO

Isto nunca acaba.

Isto nunca acaba.

Isto nunca acaba.

ANTERIOR Se desejar, pode fazer um comentário ou justificar aqui a sua resposta. SEGUINTE

Máquina de Ouvir — Question: x +

student.dei.uc.pt/~jcastro/ouvir/ ☆ ⓘ

Escolha a imagem correspondente ao excerto sonoro apresentado: 5/7

▶ CLIQUE AQUI PARA OUVIR O EXCERTO SONORO

Isto nunca acaba.

IIIIIIss to nunca acaba.

Isto nunca acaba.

ANTERIOR Se desejar, pode fazer um comentário ou justificar aqui a sua resposta. SEGUINTE

APÊNDICE B – IMAGENS DO QUESTIONÁRIO

Máquina de Ouvir — Question: x +

student.dei.uc.pt/~jccastro/ouvir/

De 1 a 5, classifique o grau de ligação entre o som e a imagem dos vídeos apresentados. 6/7

0:00 / 1:54

1 2 3 4 5

Não existe ligação entre a imagem e o som. Existe uma clara ligação entre o som e a imagem.

0:00 / 1:05

1 2 3 4 5

Não existe ligação entre a imagem e o som. Existe uma clara ligação entre o som e a imagem.

ANTERIOR

Se desejar, pode fazer um comentário ou justificar aqui a sua resposta.

TERMINAR

Máquina de Ouvir — Question: x +

student.dei.uc.pt/~jccastro/ouvir/

Obrigado!

Agradecemos a sua participação.

Se desejar fazer um comentário final ou deixar alguma sugestão, por favor faça-o no campo seguinte.

ENVIAR COMENTÁRIO

INÍCIO

APÊNDICES

C - RESPOSTAS AO QUESTIONÁRIO

APÊNDICE C – RESPOSTAS AO QUESTIONÁRIO

Data	Sexo	Idade	Design	Som	Video1	Video2	iss1		iss2		siii1		siii2	
15-08-19 21:04	M	27	1	2	4	4	2	S	5	S	5	S	7	S
15-08-19 21:25	F	22	5	3	3	4	2	S	5	S	7	S	3	S
15-08-19 18:25	F	24	3	2	5	5	4	S	7	S	3	S	1	S
15-08-19 19:28	F	45	2	2	2	5	2	S	7	S	4	S	1	S
15-08-19 20:32	M	26	2	4	5	5	6 (3)	N	6	S	8	S	6	S
15-08-19 18:50	M	22	5	1	5	5	6	S	4	S	5	S	6	S
15-08-19 18:35	F	26	4	4	4	5	8 (1)	N	1 (7)	N	7	S	1	S
15-08-19 18:32	F	22	5	3	4	5	3	S	4	S	8	S	1	S
15-08-19 19:24	M	30	5	3	3	5	2	S	8	S	3 (7)	N	3	S
15-08-19 19:08	F	27	4	1	4	5	2	S	5	S	6	S	8 (4)	S
15-08-19 18:52	M	23	2	3	4	5	6 (8)	N	4	S	4	S	2	S
15-08-19 19:31	M	28	2	5	3	4	4	S	8	S	3	S	2	S
15-08-19 19:15	F	54	2	2	3	4	8 (1)	N	6	S	7	S	8 (6)	N
15-08-19 20:03	F	28	1	4	2	5	2	S	4	S	3 (7)	N	6	S
15-08-19 19:43	M	28	4	2	4	5	2	S	3 (1)	N	1	S	4	S
15-08-19 19:35	M	28	5	3	3	5	3	S	4	S	1	S	3	S
15-08-19 20:53	F	49	4	1	5	5	2	S	4	S	6	S	8	S
15-08-19 19:10	F	61	1	1	5	5	7	S	1	S	5	S	7	S
15-08-19 18:46	M	23	4	1	3	5	5	S	6	S	7	S	4	S
15-08-19 20:31	M	28	3	1	3	5	1	S	2 (6)	N	3	S	1	S
15-08-19 20:50	F	24	2	3	5	5	3	S	6 (8)	N	4	S	2	S
15-08-19 22:02	F	26	5	4	4	5	6	S	5	S	2	S	4	S
15-08-19 19:18	M	58	2	4	4	5	8	S	4	S	5	S	2	S
15-08-19 19:41	F	22	1	1	4	5	5 (1)	N	3	S	5	S	4	S
15-08-19 18:48	F	24	5	4	5	5	4 (1)	N	2	S	1	S	4	S
15-08-19 20:04	M	20	1	1	5	5	5	S	1	S	5	S	1	S
15-08-19 20:13	M	32	1	2	4	4	3	S	2	S	1 (2)	N	4 (1)	N
15-08-19 19:21	M	23	3	2	4	5	4	S	2	S	1	S	1 (4)	N
15-08-19 18:37	M	23	4	2	3	5	8 (1)	N	7	S	8	S	2	S
15-08-19 18:33	M	23	4	2	3	5	7	S	8	S	2	S	1 (3)	N
15-08-19 18:45	M	34	1	1	3	5	2	S	8 (1)	N	6 (8)	N	4	S
15-08-19 18:46	F	52	2	2	5	5	4	S	5 (3)	N	6	S	7	S
15-08-19 19:29	M	24	3	3	5	5	2	S	5	S	8	S	5	S
15-08-19 22:06	F	28	5	4	3	4	5 (7)	N	4	S	2	S	6	S
15-08-19 20:03	M	21	3	3	4	5	6	S	2 (4)	N	8 (3)	N	8	S
15-08-19 19:04	F	22	1	1	5	5	4 (7)	N	5	S	1 (7)	N	1 (4)	N
15-08-19 18:57	M	25	2	2	4	5	5	S	2	S	6	S	7 (1)	N
15-08-19 18:58	F	24	4	3	4	4	3 (7)	N	8	S	5	S	1 (3)	N
15-08-19 19:38	M	29	3	3	5	5	2	S	5	S	4	S	2	S

APÊNDICE C – RESPOSTAS AO QUESTIONÁRIO

15-08-19 19:30	M	29	1	2	5	5	6 (8)	N	5 (3)	N	8 (3)	N	8	S
15-08-19 22:04	M	28	4	4	4	4	3 (8)	N	3	S	1	S	4	S
16-08-19 00:12	F	30	1	3	4	5	6 (3)	N	8	S	1 (7)	N	4	S
15-08-19 23:19	F	49	5	1	5	4	5	S	2	S	6	S	1 (7)	N
15-08-19 22:50	M	22	1	4	3	5	6	S	2	S	2	S	1 (6)	N
16-08-19 00:36	M	42	2	2	4	5	3 (7)	N	5	S	4	S	5	S
16-08-19 01:28	F	22	1	1	3	5	4 (1)	N	6	S	4	S	7	S
16-08-19 11:29	M	39	1	3	5	5	2	S	4	S	2	S	5 (8)	N
16-08-19 11:53	F	23	5	3	4	4	6	S	8	S	2	S	7	S
16-08-19 13:14	M	29	1	5	5	5	6	S	7	S	4	S	2	S
16-08-19 11:20	M	35	5	3	2	5	6	S	2	S	2 (6)	N	2	S
16-08-19 11:03	F	24	4	3	4	5	1 (8)	N	4	S	6	S	8	S
16-08-19 12:03	F	24	5	3	5	5	6 (3)	N	8	S	1	S	5	S
16-08-19 10:18	F	27	4	1	5	5	8 (4)	N	2	S	5	S	7	S
16-08-19 04:23	M	24	4	3	4	4	1	S	6	S	5	S	8	S
16-08-19 12:00	M	28	3	3	4	5	3	S	5	S	1	S	4 (3)	N
16-08-19 09:01	M	58	5	3	5	5	7	S	8 (5)	N	3 (1)	N	7	S
16-08-19 10:49	M	34	2	2	4	5	7 (1)	N	2	S	8	S	2	S
16-08-19 11:55	M	27	1	3	4	5	7 (8)	N	5	S	1	S	5	S
16-08-19 11:39	M	31	2	3	4	5	5	S	2	S	3	S	2	S
16-08-19 13:04	M	26	2	5	4	4	1	S	8	S	1	S	7	S
16-08-19 10:20	M	32	2	2	5	5	6 (4)	N	5	S	5	S	4	S
16-08-19 02:02	F	25	4	4	5	5	5	S	2	S	6 (3)	N	1	S
16-08-19 02:49	M	29	5	1	2	5	2	S	4 (7)	N	6	S	7	S
16-08-19 12:02	M	29	1	3	3	4	4	S	8	S	1 (7)	N	6	S
16-08-19 13:16	F	28	5	1	4	4	8	S	3	S	2 (6)	N	8 (3)	N
18-08-2019 04:13	F	24	1	1	5	3	1	S	6	S	5	S	2	S
16-08-2019 15:58	M	30	1	2	4	5	7	S	5	S	4	S	2	S
16-08-2019 14:36	M	31	1	1	5	5	8 (7)	N	5	S	5	S	7	S
17-08-2019 04:15	M	28	1	5	5	5	5 (8)	N	5	S	7	S	2	S
17-08-2019 10:11	F	34	3	1	4	4	4	S	6 (3)	N	8	S	7 (1)	N
16-08-2019 16:22	M	27	5	3	3	5	2	S	7	S	2	S	7	S
16-08-2019 15:54	M	61	2	2	5	5	5	S	3	S	2	S	5	S
18-08-2019 18:39	M	28	3	2	5	5	5 (8)	N	7 (1)	N	7	S	5	S
16-08-2019 13:39	M	25	1	5	3	4	3	S	4	S	4	S	5	S
16-08-2019 14:08	M	29	3	3	4	5	2	S	1	S	2	S	5	S
16-08-2019 21:29	M	25	1	1	4	4	6	S	8	S	2	S	4 (7)	N
17-08-2019 12:01	M	29	1	3	3	5	6 (2)	N	3	S	6	S	4	S
16-08-2019 15:27	F	60	2	2	5	5	2 (1)	N	2(3)	N	3	S	1	S
18-08-2019 15:36	M	21	2	3	4	5	6 (1)	N	2	S	2	S	3	S

APÊNDICE C – RESPOSTAS AO QUESTIONÁRIO

16-08-2019 15:59	M	27	1	5	4	5	8	S	5	S	5	S	2	S
18-08-2019 17:32	M	27	1	3	4	5	7	S	6	S	8 (6)	N	4	S
16-08-2019 19:13	F	26	1	1	4	5	4	S	6 (3)	N	6	S	2	S
16-08-2019 19:17	M	29	1	2	5	5	2 (8)	N	4	S	1 (6)	N	5	S
18-08-2019 14:56	M	26	1	1	5	5	4	S	1	S	3	S	1	S
17-08-2019 19:40	M	26	5	2	5	5	6	S	2	S	6	S	7	S
16-08-2019 18:14	M	27	3	4	3	5	6 (3)	N	1	S	5 (3)	N	7	S
16-08-2019 19:01	M	25	3	3	4	4	3 (8)	N	3	S	7	S	4	S
17-08-2019 00:45	F	29	1	5	4	5	6	S	8	S	3	S	4	S
17-08-2019 18:06	F	23	1	1	4	4	8 (7)	N	5 (8)	N	2	S	6 (3)	N
18-08-2019 21:02	F	22	3	2	3	5	2 (1)	N	6 (8)	N	8	S	7	S
18-08-2019 20:30	F	36	1	1	4	4	3	S	7	S	4	S	5	S
18-08-2019 21:07	F	23	5	2	4	5	8 (1)	N	8	S	2	S	3	S
18-08-2019 21:51	M	28	5	3	4	4	4 (7)	N	6	S	1 (7)	N	5	S
18-08-2019 21:31	F	26	5	3	4	4	8	S	4	S	5	S	3	S
18-08-2019 22:56	M	19	3	4	5	4	4	S	5	S	5	S	6	S
19-08-2019 01:27	F	24	1	1	4	5	6 (7)	N	5	S	2 (1)	N	3	S
19-08-2019 09:53	F	45	5	1	5	5	5	S	4	S	2	S	7	S
19-08-2019 00:32	M	34	2	4	1	2	8 (7)	N	2	S	7	S	4 (1)	N
19-08-2019 10:08	M	51	3	2	3	4	3	S	6	S	5	S	8	S
19-08-2019 10:56	F	25	5	2	4	5	2	S	3 (1)	N	2	S	6	S
19-08-2019 10:58	M	26	3	5	4	5	8	S	4	S	7	S	4	S
19-08-2019 11:26	M	34	4	3	4	5	3	S	2	S	2	S	8	S
19-08-2019 01:15	M	21	5	2	4	5	6	S	3	S	8	S	2	S
19-09-2019 11:13	M	24	4	2	4	5	4	S	7	S	3 (8)	N	6	S
19-09-2019 09:00	F	35	5	2	4	5	8	S	7	S	8 (3)	N	5	S
19-08-2019 17:49	F	19	3	1	2	5	6 (1)	S	6	S	3	S	1	S
19-08-2019 20:42	M	22	3	4	4	4	3	S	2	S	4 (7)	N	2	S
19-08-2019 18:25	F	25	5	4	1	1	1 (8)	N	2 (1)	N	2 (6)	N	5	S
19-08-2019 18:28	M	28	2	4	4	5	3 (7)	N	8	S	7	S	2	S
19-08-2019 18:34	F	32	1	1	5	5	2	S	3	S	4	S	7	S
19-08-2019 20:30	F	23	4	1	4	5	5	S	1	S	5	S	3 (4)	N
19-08-2019 15:16	F	23	4	3	5	5	6	S	8	S	8	S	4	S
20-08-2019 12:01	M	66	1	2	2	4	3 (7)	N	6 (1)	N	8	S	2	S
20-08-2019 01:23	M	26	4	3	3	5	1	S	3 (4)	N	2	S	7	S
19-08-2019 23:06	F	27	1	2	5	5	8	S	3	S	7	S	6 (3)	N
22-08-2019 18:06	M	34	1	5	4	4	6	S	4	S	8	S	1	S
20-08-2019 14:56	M	21	5	3	4	5	8	S	1	S	3	S	7	S
21-08-2019 08:09	F	27	4	2	3	5	4	S	2	S	5	S	1 (3)	N
21-08-2019 19:44	M	29	5	3	5	4	5	S	8	S	3 (1)	N	7	S

APÊNDICE C – RESPOSTAS AO QUESTIONÁRIO

21-08-2019 10:15	F	22	4	2	4	5	1	S	6 (3)	N	8	S	4	S
22-08-2019 20:58	F	25	1	3	5	4	7 (1)	N	6 (4)	N	8 (5)	N	4 (3)	N
23-08-2019 22:05	M	25	5	4	5	4	4	S	3	S	3	S	4	S
24-08-2019 10:56	M	38	3	2	3	5	8 (1)	N	5	S	2	S	6	S
25-08-2019 21:57	M	28	5	3	4	5	3 (1)	N	4	S	1	S	4	S
23-08-2019 11:55	M	21	5	4	4	5	5	S	8 (1)	N	6	S	1	S
23-08-2019 11:20	M	28	4	2	4	5	8	S	6 (3)	N	1	S	7	S
23-08-2019 09:44	M	27	5	5	4	5	3	S	4	S	2	S	1	S
27-08-2019 09:39	F	43	5	2	2	4	6	S	2	S	2	S	7	S

APÊNDICES

D - COMENTÁRIOS SUBMETIDOS NO QUESTIONÁRIO

APÊNDICE D – COMENTÁRIOS QUESTIONÁRIO

Comentário
Deixo como sugestão aumentar o contraste dos tamanhos: quando alguém grita, haver um crescimento exponencial e mais “repentino” da letra, para evidenciar essa variação. Além disso, aconselhava a aumentar no geral o tamanho da fonte, pelo simples facto do tamanho mais pequeno que surge nos vídeos e nos exemplos chegar a comprometer a legibilidade. De resto, parabéns pelos resultados!
Está brutal! Parabéns! Sugestões só me ocorre tentar afinar um pouco a normalização do volume ou do tamanho do texto, para garantir que o texto fica sempre entre determinada gama de tamanho. No entanto, estava a ver os vídeos num telemóvel em ecrã vertical, pode bem ser disso e não do algoritmo :) mais uma vez, parabéns. Enquanto leitor parece-me ser uma ideia fantástica.
GOSTEI outra vez.
Muito interessante, uma outra forma de escrever música! Foi um prazer Ouvir!
Bom trabalho artista :)
O texto e poesia são maravilhosos...O projeto é INCRÍVEL...com enorme potencial... e inúmeras aplicabilidades!!!
Muita força para a tese. Gostei muito da ideia do projeto! Um abraço :)
Parabéns!! Com este questionário percebi que para além do texto aparecer enquanto é mencionado, o formato do texto (tamanho, negrito, espaçamento, ...) altera-se com a entoação dada pelo leitor. Poderia então dizer que nestas letras se pode encontrar a interpretação da emoção de cada um enquanto o faz. Brillhante!!
Bom trabalho.
Os meus parabéns :D
As imagens podem criar dissonância cognitiva quando não estão alinhadas com a mensagem sonora. Ou por outro lado, podem amplificar o impacto e clarificar o conteúdo. Acredito que o design tem o poder de democratizar, simplificar, reforçar, clarificar. Uso sempre nos meus projectos ligados à gestão de talento. Não sei exactamente qual o propósito do estudo. Talvez fosse interessante torna-lo mais claro no final para não influenciar as respostas. O que pretendem demonstrar com este exercício. Bom trabalho!!
Adorei os vídeos da última parte do questionário.
Incribél
Boa sorte João! Ótimo trabalho!
Esse sistema, essa máquina de ouvir, é uma das coisas mais poéticas que vi recentemente. Dar imagem à forma como a palavra é dita, é algo belíssimo! Parabéns!
Olá! 2 comentários: Sendo a língua portuguesa um dos focos deste trabalho, é fundamental que não haja gralhas. O penúltimo texto (sobre a emancipação das mulheres) tem alguns erros que deviam ser corrigidos. Compreendo o exercício, como aproximação à poesia concreta, mas acho que devia ser mais explorado do ponto de vista gráfico. Abraço
Well done
Acho que é uma ferramenta que tem muito potencial no contexto de design gráfico assim como para performance ao vivo.
Parabéns pelo projecto, pela aparente simplicidade, pelas potencialidades práticas do mesmo e pelas possibilidades evolutivas que se adivinham.
Parabéns pelo trabalho! Por vezes o mais simples não é o menos complicado, mas sim o melhor pensado.
Um belo desafio para testar a articulação entre ouvir e ver! Nada fácil, meu caro. Bom trabalho!
Fantástico trabalho. Boa sorte. Agradeço a oportunidade de poder ajudar!
bom trabalho!

