

# TRANSCRIÇÃO DE ENTREVISTAS SOCIOLINGÜÍSTICAS COM O ELAN

Livia Oushiro

## INTRODUÇÃO

O ELAN (HELLWIG; GEERTS, 2013) é um programa para anotação de arquivos de áudio e vídeo, desenvolvido pelo Instituto Max Planck de Psicolinguística. Entre suas principais vantagens, encontram-se:

- a sincronização entre o arquivo de mídia e a transcrição/anotação, o que facilita enormemente a análise linguística dos dados (por exemplo, para codificação de variantes de variáveis fonéticas);
- a possibilidade de criação de múltiplas trilhas, que propicia não só a separação da fala de diferentes participantes, mas também a anotação detalhada de outros aspectos linguísticos e contextuais, bem como a representação de ações simultâneas (por exemplo, sobreposição de vozes, ações gestuais concomitantes às verbais);
- ferramentas mais sofisticadas de buscas dentro de um *corpus* (por exemplo, para encontrar todas as ocorrências de variantes de uma variável);
- a ampla flexibilidade de formatos de exportação da transcrição (.txt, .textgrid etc.) e, conseqüentemente, a compatibilidade com outros programas como Word, Excel, R, Rbrul, Praat, etc.;
- o fato de ser gratuito, e que vem sendo utilizado cada vez mais entre estudiosos da língua em uso, entre os quais os sociolinguistas (NAGY;

MEYERHOFF, 2013), o que configura um ponto positivo no compartilhamento de dados e metodologias.

O programa está disponível em <<http://www.lat-mpi.eu/tools/elan/>>. Nessa página, clique em *Download the latest version*. Na página que se abrir, baixe a versão correspondente a seu sistema operacional (Windows, Mac OS ou Linux). Além do programa, nesse site é possível baixar um manual detalhado e inscrever-se na lista de discussão. Após instalá-lo, você pode mudar a língua da interface clicando em *Opções > Língua > Português*. O principal objetivo deste tutorial é apresentar as ferramentas que serão de maior utilidade para transcrições de entrevistas sociolinguísticas. Ao final, apresentam-se algumas breves notas sobre a ferramenta de buscas, a exportação de arquivos e normas gerais de transcrição.

## 1. ARQUIVOS DE ANOTAÇÃO

Ao iniciar o ELAN, aparecerá uma janela vazia. Para começar uma nova anotação, clique em *Arquivo > Novo...*, ou use o atalho [Ctrl] + [N]<sup>1</sup>. Uma nova janela se abrirá (Figura 1). Do lado esquerdo, selecione o arquivo de som que deseja transcrever e clique no botão [>>] para adicioná-lo aos *Arquivos Selecionados* à direita. É preferível utilizar os arquivos de som em formato .wav (não .mp3), pois eles permitem a visualização da onda sonora<sup>2</sup>. Clique em OK. O arquivo será aberto em uma nova janela.

1 No Mac, substitua a tecla [Ctrl] por [Command].

2 É possível que, mesmo abrindo o arquivo de áudio no formato .wav, a onda sonora não esteja tão claramente visível. Para melhorar a visualização, vale a pena ajustar a amplitude da onda. Isso pode ser feito no programa Audacity, disponível gratuitamente em <<http://audacity.sourceforge.net/download/>>. Após instalar o programa faça, primeiramente, uma cópia do arquivo sonoro que você deseja amplificar – é sempre bom manter uma cópia do original, sem alterações! Abra o novo arquivo no programa. No menu superior, clique em *Effect > Amplify...* Na janela que abrir, aumente a amplitude através do slider. Clique em *Allow clipping* e clique em OK. No menu superior, clique em *File > Export...* e salve o arquivo sonoro com onda amplificada. Use esse novo arquivo para trabalhar no ELAN.

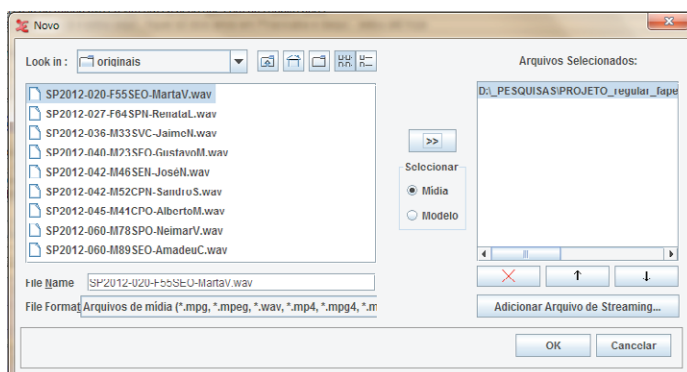


Figura 1 – Seleção de um arquivo de mídia para criação de nova anotação.

O ELAN organiza os arquivos em *projetos*. Cada projeto consiste em pelo menos dois arquivos: um ou mais *arquivos de mídia* (em geral, em estudos sociolinguísticos, um arquivo de som, mas também é possível fazer anotações de arquivos de vídeos), e um *arquivo de anotação* (com a extensão *.eaf*, se criado no ELAN, mas é possível inserir outros formatos de arquivo).

O ELAN associa um arquivo de anotação a seu respectivo arquivo de áudio, e armazena tais informações no arquivo *.eaf*. Isso significa que cada vez que você abre um arquivo do ELAN, o programa automaticamente busca um arquivo de som associado a ele. Se o ELAN não encontrar o arquivo sonoro no local especificado previamente (porque você está abrindo o arquivo de anotação em outro computador ou porque o moveu para um novo local no disco rígido), o programa pedirá novamente a localização desse arquivo (Figura 2).

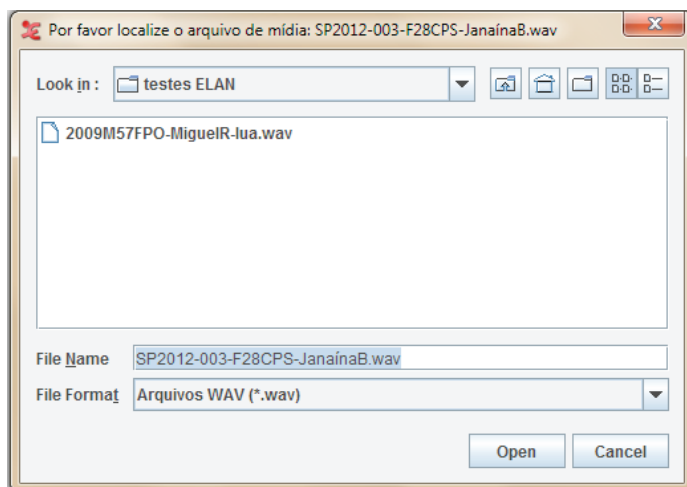


Figura 2 – Janela para localizar o arquivo de mídia.

Para abrir um arquivo de anotação criado previamente, clique em *Arquivo* > *Abrir...*, ou use o atalho [Ctrl] + [O]. Para salvar o seu trabalho, clique em *Arquivo* > *Salvar...*, ou o atalho [Ctrl] + [S].

## 2. AS PRINCIPAIS FUNÇÕES DO ELAN

Ao abrir um arquivo de áudio para transcrição, a janela principal do ELAN aparecerá (Figura 3). Ela contém diversas ferramentas que são descritas detalhadamente a seguir.

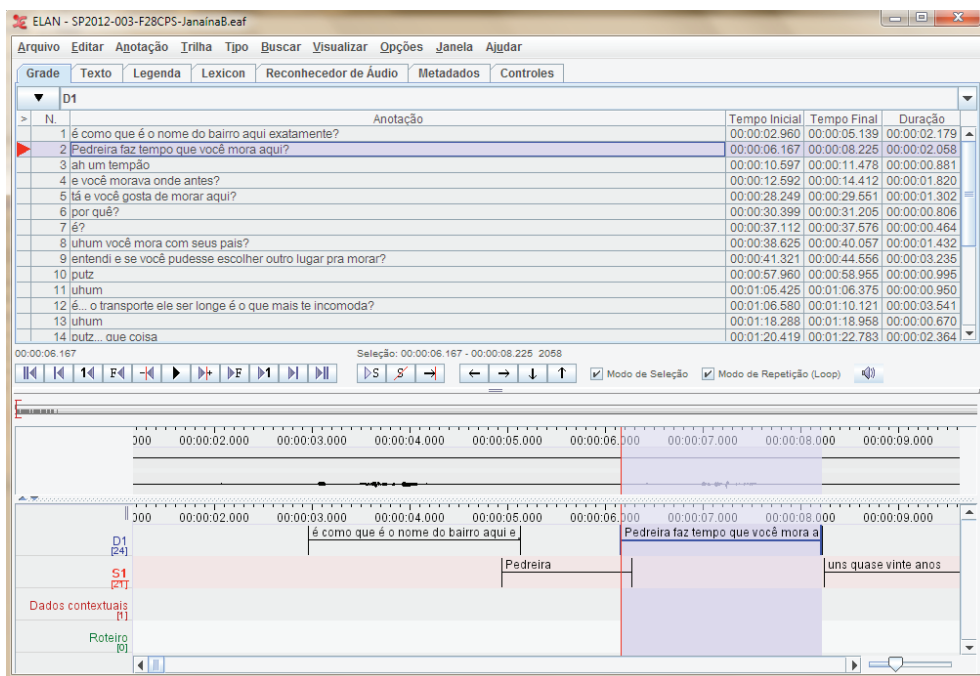


Figura 3 – Janela principal do ELAN.

Para nossos propósitos, as abas *Controles*, *Grade* e *Texto* na metade superior da janela principal do ELAN são as mais relevantes. A aba *Controles* tem dois *sliders*, que permitem ajustar o volume e a velocidade da gravação. Este último pode ser útil para transcrever falas muito rápidas (Figura 4).

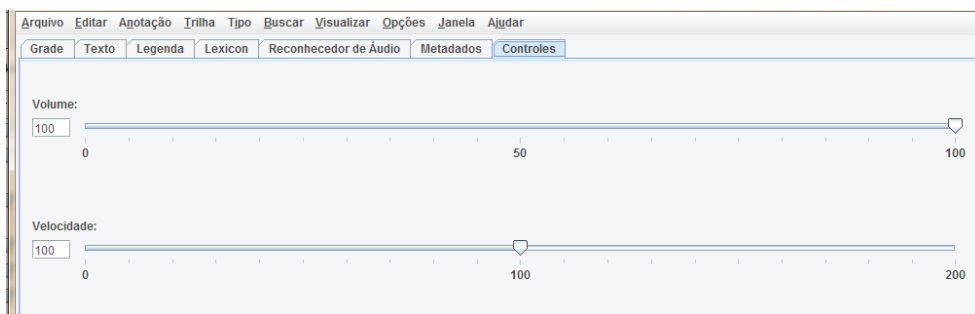


Figura 4 – Janela principal do ELAN – Controles de volume e velocidade.

A aba *Grade* apresenta uma tabela de unidades de anotação de uma determinada trilha, junto com as informações de tempo (inicial, final, duração) para cada anotação (Figura 5). Elas podem ser selecionadas (com um clique) e editadas (com duplo clique). A anotação selecionada é destacada com uma caixa azul escura e indicada por um triângulo vermelho. A *Grade* é sincronizada com a onda sonora, de modo que ao selecionar uma anotação, automaticamente a porção sonora é exibida no *Visualizador da Onda Sonora*. De modo inverso, a seleção de um trecho sonoro no visualizador destaca o intervalo correspondente na *Grade*.

N	Anotação	Tempo Inicial	Tempo Final	Duração
1	é como que é o nome do bairro aqui exatamente?	00:00:02.960	00:00:05.139	00:00:02.179
2	Padreira faz tempo que você mora aqui?	00:00:06.167	00:00:08.225	00:00:02.058
3	ah um tempo	00:00:10.597	00:00:11.478	00:00:00.881
4	e você morava onde antes?	00:00:12.592	00:00:14.412	00:00:01.820
5	tá e você gosta de morar aqui?	00:00:28.249	00:00:29.551	00:00:01.302
6	por quê?	00:00:30.399	00:00:31.205	00:00:00.806
7	é?	00:00:37.112	00:00:37.576	00:00:00.464
8	uhum você mora com seus pais?	00:00:38.625	00:00:40.057	00:00:01.432
9	entendi e se você pudesse escolher outro lugar pra morar?	00:00:41.321	00:00:44.556	00:00:03.235
10	putz uhum você mora com seus pais?	00:00:57.960	00:00:58.955	00:00:00.995
11	uhum	00:01:05.425	00:01:06.375	00:00:00.950
12	é... o transporte ele ser longe é o que mais te incomoda?	00:01:06.580	00:01:10.121	00:00:03.541
13	uhum	00:01:18.288	00:01:18.958	00:00:00.670
14	putz... que coisa	00:01:20.419	00:01:22.763	00:00:02.364

Figura 5 – Janela principal do ELAN – Grade.

A aba *Texto* apresenta um texto corrido de todas as transcrições em uma determinada trilha (Figura 6). As fronteiras entre as unidades de anotação são indicadas por pontos. De modo semelhante à aba *Grade*, uma unidade de anotação pode ser selecionada com um clique e editada com o duplo-clique (através de uma janela de edição na qual a transcrição pode ser modificada), e a onda sonora e o texto são sincronizados.

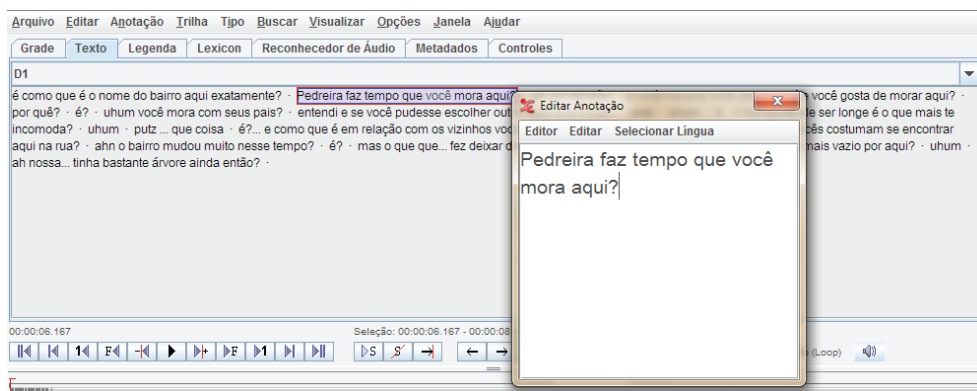


Figura 6 – Janela principal do ELAN – Texto.

Na parte central da janela, os *Botões de Controle* permitem tocar ou pausar o arquivo de áudio, navegar pela gravação, selecionar partes do arquivo de áudio e navegar entre as anotações (Figura 7).



Figura 7 – Janela principal do ELAN – Botões de controle.

O primeiro grupo de botões permite tocar ou pausar a gravação, e navegar pelo arquivo de áudio. Da esquerda para a direita, as respectivas funções são: (i) ir para o começo da mídia; (ii) ir para o enquadramento anterior; (iii) voltar um segundo; (iv) voltar um *frame*; (v) voltar um *pixel*; (vi) tocar/pausar; (vii) ir ao próximo *pixel*; (viii) ir ao próximo *frame*; (ix) adiantar um segundo; (x) ir para o próximo enquadramento e (xi) ir para o final da mídia. O segundo conjunto de botões controla as seleções. Da esquerda para a direita: (i) tocar o intervalo selecionado; (ii) limpar o intervalo selecionado e (iii) mover a linha vermelha para os extremos da seleção. O terceiro conjunto de botões permite navegar para frente e para trás entre as unidades de anotação nas trilhas. Da esquerda para a direita: (i) ir para a anotação anterior; (ii) ir para a próxima anotação; (iii) ir para a trilha de cima; e (iv) ir para a trilha abaixo. Por fim, há duas caixas: (i) se *Modo de Seleção* estiver selecionada, um trecho tocado será automaticamente selecionado; (ii) se *Modo de Repetição (Loop)* estiver selecionada, o intervalo selecionado será tocado repetidas vezes ao clicar sobre o botão *Play*.

O *Visualizador de Densidade de Anotações* se localiza abaixo dos botões de controle (Figura 8) e fornece uma rápida impressão visual de quanto do arquivo sonoro já foi transcrito. As marcações em cinza indicam as regiões no áudio que

contêm unidades de anotação, de modo que você pode visualizar rapidamente quais partes das gravações já foram transcritas. Esse visualizador também é um modo fácil de navegar pelo arquivo, já que a extensão da barra corresponde ao da gravação completa, independente do *zoom* no *Visualizador da Onda Sonora*. A barrinha vertical vermelha indica a posição atual do cursor. É possível navegar facilmente pelo arquivo ao mudar a barrinha para a esquerda ou para a direita.



Figura 8 – Janela principal do ELAN – Visualizador de densidade de anotações.

Na metade inferior da janela principal do ELAN, o *Visualizador da Onda Sonora* (Figura 9) apresenta a amplitude (eixo vertical) ao longo do tempo (eixo horizontal) do arquivo de áudio. Ele também mostra a atual posição do cursor (linha vertical vermelha), informação de tempo e quais partes do arquivo de áudio estão selecionadas (destacadas em azul claro). Você pode aumentar ou diminuir a escala de tempo visualizada segurando o botão [Ctrl] e usando a roda de rolagem de seu *mouse*.

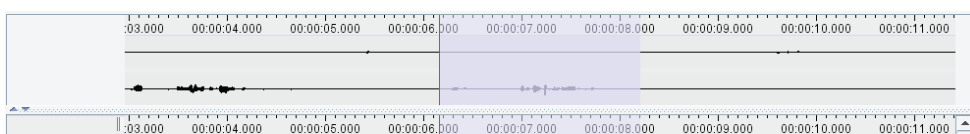


Figura 9 – Janela principal do ELAN – Visualizador da Onda Sonora.

Por fim, abaixo do *Visualizador da Onda Sonora* estão as *Trilhas de Anotação*, sendo uma trilha por falante. Cada trilha contém unidades de anotação, que por sua vez contêm a transcrição. No lado esquerdo do painel estão os nomes das trilhas; aquela selecionada ou ativa está em vermelho claro (por exemplo, na Figura 10, a trilha selecionada é S1). Ao posicionar o cursor do *mouse* sobre os nomes (sem clicar), aparecerá uma janela com informações mais específicas sobre a trilha, com o seu nome, nome do participante, nome do anotador, etc.

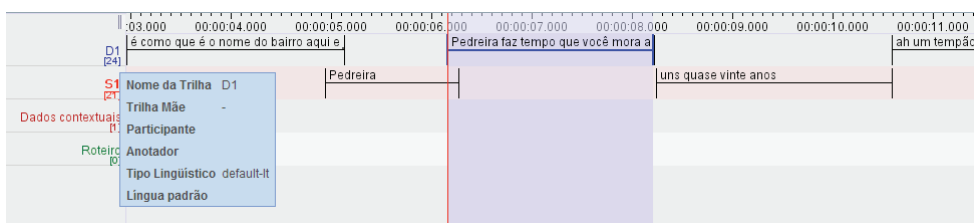


Figura 10 – Janela principal do ELAN – Trilhas de anotação.

### 3. TRILHAS

De modo geral, uma trilha de anotação no ELAN contém um texto transcrito junto com marcações de tempo. Essas marcações são usadas pelo programa para sincronizar a exibição da onda sonora e da anotação.

As trilhas no ELAN podem ter diferentes funções. Por exemplo, para um mesmo falante, você pode criar uma trilha para a transcrição ortográfica das sentenças, outra para as palavras individuais, outra para as unidades morfológicas, outra para unidades fonéticas, etc. Para estudos sociolinguísticos, é interessante criar pelo menos as seguintes trilhas:

- uma trilha para cada participante da gravação (documentador, informante, falantes adicionais, etc.);
- uma trilha para dados contextuais, que se referem à ações simultâneas dos participantes ou de outros eventos no ambiente de gravação, mas não aos enunciados da entrevista [risos], [tosse], [barulho de moto], etc.;
- uma trilha para identificar partes do roteiro (bairro, infância, lista de palavras etc.)

Para criar uma nova trilha, clique em *Trilha > Adicionar nova trilha...* ou [Ctrl] + [T]. Uma nova janela se abrirá, com uma lista na metade superior e diferentes abas e campos na metade inferior (Figura 11). A aba *Adicionar* estará destacada. Nessa janela, insira as seguintes informações para cada trilha que você deseja criar:

- nome da trilha: inserir a identificação do falante (por exemplo, “D1” para documentador, “S1” para o informante, “S2, S3” para falantes adicionais) ou de outros tipos de anotação (por exemplo, “Dados Contextuais”, “Roteiro”);
- participante: inserir o pseudônimo dos informantes e o nome dos documentadores;
- anotador: inserir o seu nome completo.



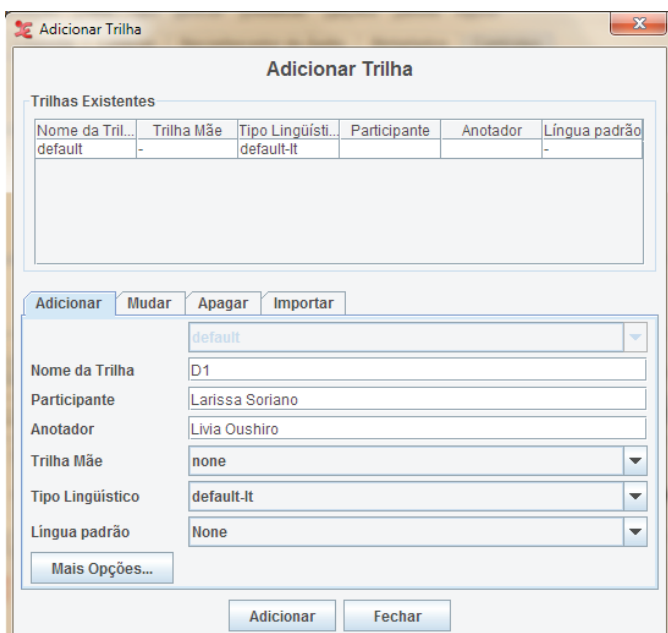


Figura 11 – Janela para criação de novas trilhas.

Os demais campos (“Trilha mãe”, “Tipo linguístico”, “Língua padrão”) podem ser deixados nos valores pré-configurados. Para cada trilha, clique em “Adicionar”. Depois disso, clique sobre a aba “Apagar” para deletar a trilha “default”. Por fim, clique em “Fechar” para sair da janela.

## 4. PROCEDIMENTOS

### 4.1. Mudança nos atalhos

Em geral, é preferível trabalhar com o teclado o máximo possível, pois o trabalho se torna muito mais rápido e menos cansativo (fisicamente) do que alternar entre o *mouse* e o teclado. Para isso, você pode usar os atalhos para a maior parte dos comandos no ELAN. Para ver (e imprimir, se quiser) uma lista de atalhos, clique em *Visualizar > Atalhos....*

O ELAN vem com um grande conjunto de atalhos pré-configurados, alguns dos quais não muito intuitivos. Você pode mudar quaisquer atalhos para uma combinação que achar mais adequada, através de *Editar > Preferências > Editar Atalhos....* Uma nova janela se abrirá (Figura 12). Selecione *Categoria* do menu, no canto inferior esquerdo, para organizar os diferentes atalhos de acordo com sua função. Selecione o

atalho a editar clicando sobre ele e clique em *Editar Atalho* (ver Figura 13). Insira a combinação de teclas que você prefere na nova janela e clique em *Apply in all modes*. Após fazer todas as alterações que desejar, clique no botão *Salvar* na janela de atalhos.

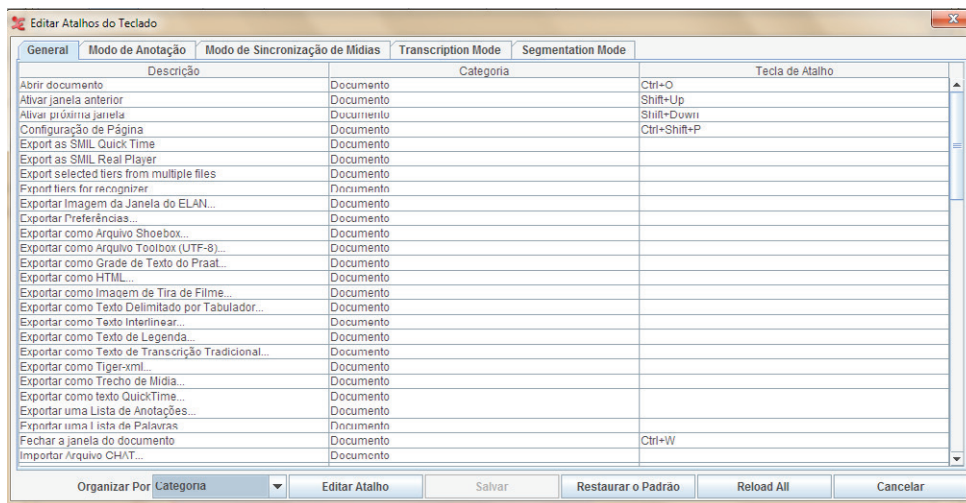


Figura 12 – Janela de edição de atalhos.

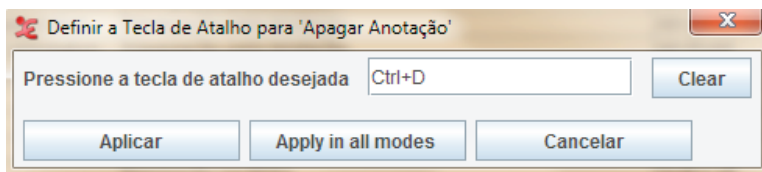


Figura 13 – Especificação de novo atalho.

## 4.2. Sugestões de atalhos

As seguintes teclas de atalho são sugeridas para facilitar a transcrição no ELAN.

<b>Modo de anotação – Edição de anotações</b>	
Deletar anotação	[Ctrl] + [Delete]
Modificar valor da anotação	[Ctrl] + [M]
Modificar o tempo da anotação	[Ctrl] + [Enter]
Nova anotação aqui	[Shift] + [Enter]
Remover valor da anotação	[Shift] + [Delete]
<b>Modo de anotação – Trilha</b>	
Ativar trilha superior	[Ctrl] + [↑]
Ativar trilha inferior	[Ctrl] + [↓]
<b>Transcription Mode – Navegação pela anotação</b>	
Ir para a próxima anotação	[Alt/Option] + [→]
Ir para a anotação anterior	[Alt/Option] + [←]
<b>Segmentation Mode – Navegação na mídia</b>	
Tocar/pausar a mídia	[Shift] + [Space]
Tocar seleção	[Ctrl] + [Space]
Definir tempo 1 segundo atrás	[Ctrl] + [←]
Definir tempo 1 segundo à frente	[Ctrl] + [→]
Ir para o <i>pixel</i> anterior	[Shift] + [←]
Ir para o próximo <i>pixel</i>	[Shift] + [→]
<b>Segmentation Mode – Seleção</b>	
Limpar seleção	[Esc]
Mover cursor para os extremos da seleção	[Ctrl] + [/]

Quadro 1 – Sugestão de novos atalhos no ELAN.

Também é recomendável mudar as preferências de edição, de maneira que para salvar as mudanças em uma caixa de anotação, baste pressionar [Enter] (a configuração *default* é pressionar [Ctrl] + [Enter]). Para fazer isso, clique em *Preferências > Editar Preferências....* No menu à esquerda, clique em *Editar e*, à direita, selecione a segunda opção: *A tecla Enter realiza as mudanças na caixa de edição alinhada*. Você também pode escolher a primeira opção: *Ao retirar a seleção da caixa de edição do texto alinhado as mudanças serão realizadas*, o que significa que, quando você sair da caixinha de edição (por exemplo, clicando fora

dela), o que foi digitado dentro da caixinha será salvo (a configuração *default* é que o ELAN descarte tais mudanças).

### 4.3. Fluxo de trabalho

Os seguintes passos são sugeridos para um fluxo de trabalho mais dinâmico na transcrição das gravações:

- 1) ligue o *Modo de Seleção* (e o *Modo de Repetição*, se quiser) clicando sobre as respectivas caixas de seleção acima da onda sonora;
- 2) o cursor estará no início do arquivo. Pressione [Shift] + [Space] para começar a tocar. À medida que o cursor se mover, ele selecionará o intervalo tocado;
- 3) deixe o cursor tocar até antes da primeira sentença. Pressione [Shift] + [Space] para pausar a gravação;
- 4) use os controles de navegação de mídia para mover o cursor exatamente para o ponto em que você quer começar um nova anotação. As teclas [Ctrl] + [←] e [Ctrl] + [→] voltam/adiantam 1 segundo e as teclas [Shift] + [←] e [Shift] + [→] voltam/adiantam um *frame* (provavelmente esses últimos serão mais utilizados);
- 5) desfça a seleção atual com [Esc].
- 6) inicie o *playback* novamente com [Shift] + [Space]. O cursor agora vai começar a selecionar o trecho de fala.
- 7) pause o *playback* logo após o final da fala com [Shift] + [Space]. Se necessário, use os controles de navegação de mídia novamente para mover o cursor exatamente para o ponto final da nova anotação. Você pode ouvir o trecho selecionado com os comandos [Ctrl] + [Space];
- 8) ative a trilha do falante correspondente onde você quer a nova anotação usando [Ctrl] + [↑]/[↓];
- 9) se você ligou o *Loop Mode*, pressione [Ctrl] + [Space] para começar a tocar a seleção;
- 10) pressione [Shift] + [Enter] para criar uma nova anotação na trilha ativa;
- 11) após transcrever a fala, pressione [Enter] para salvar a transcrição;
- 12) pressione [Ctrl] + [Space] para parar o *loop* da seleção atual;
- 13) pressione [Shift] + [Space] para recomeçar o *playback* da atual posição do cursor;
- 14) repita os passos (3) a (13).

Com um pouco de prática, o usuário perceberá que tais comandos logo se tornam automatizados. Para *Editar o conteúdo de uma anotação*, navegue para a anotação com [Alt/Option] + [←]/[→]. Isso vai mover a anotação selecionada para trás ou para frente. Pressione [Ctrl] + [M]. Uma janela de edição se abrirá. Após fazer as modificações desejadas, pressione [Enter] para salvar e fechar a janela de edição. Para *Mudar o tempo de duração* dos segmentos, certifique-se primeiro de que o *Modo de Seleção* está ligado. Navegue até a anotação com [Alt/Option] + [←]/[→]. Use [Ctrl] + [I] para colocar o cursor do lado da seleção que você deseja mudar (esquerda ou direita). Use os comandos de navegação de mídia para mudar a extensão da seleção ([Ctrl] + [←]/[→] move 1 segundo, [Shift] + [←]/[→] move 1 pixel). Após ajustar a seleção, pressione [Ctrl] + [Enter]. Para *apagar uma anotação*, navegue até ela [Alt/Option] + [←]/[→] e pressione [Ctrl] + [Delete].

## 5. BUSCAS EM MÚLTIPLOS ARQUIVOS E EXPRESSÕES REGULARES

Uma vez que o *corpus* esteja transcrito, é possível buscar certas palavras ou sequências de caracteres em um ou mais arquivos de transcrição no próprio ELAN. Essas opções se encontram em *Buscar*<sup>3</sup>. Uma das vantagens do programa é permitir a busca não só por sequências literais (por exemplo, “os”, “as”), mas também através de *expressões regulares*, que são sequências de caracteres em que alguns têm significado especial. Pode-se especificar, por exemplo, que se deseja buscar todas as ocorrências de /r/ em coda em posição final com a expressão regular [aâêéííóóú]r\b, em que “[aâêéííóóú]” cria uma classe de caracteres que podem se alternar (neste caso, qualquer vogal oral, especificadas por seus grafemas) e “\b” indica uma fronteira de palavra.

Para utilizar expressões regulares, o usuário deve selecionar a caixa *Expressão regular*, para indicar ao programa que não se deseja buscar a sequência literal. Uma lista completa de caracteres especiais no ELAN pode ser consultada em seu manual (HELLWIG; GEERTS, 2013, p. 305).

3 Para múltiplos arquivos, deve-se definir um *Domínio de busca*, que se refere a uma pasta ou a um conjunto de arquivos em que a busca deve ser realizada.

## 6. EXPORTAÇÃO DE ARQUIVOS

Como visto, os arquivos do ELAN são salvos com extensão .eaf. O programa permite a exportação desses arquivos em diversos formatos (texto delimitado por tabulador, texto de transcrição tradicional, lista de palavras, textgrid do Praat, etc.), compatíveis com outros programas como Word (.txt), Excel (texto delimitado por tabulador), R (.txt) e Praat (.textgrid). Tais opções se encontram em *Arquivo > Exportar como....* Recomenda-se que o usuário explore as diversas opções de exportação de arquivo, para que encontre aquela que melhor atenda aos seus propósitos.

## 7. ALGUMAS NOTAS SOBRE NORMAS DE TRANSCRIÇÃO

O objetivo principal de um sistema de transcrição é transpor a língua falada para o texto escrito de uma forma fiel à língua oral, mas inteligível, de modo a armazenar o material coletado em meio escrito de maneira padronizada para facilitar a sua análise. É necessário lembrar que nenhum sistema de transcrição é capaz de reproduzir a fala tal e qual e que qualquer sistema proposto sofrerá de limitações; com esse fato em mente, os critérios devem sempre levar em conta os objetivos do grupo de pesquisa e os tipos de análise que serão desenvolvidas com o material. Compare, por exemplo, diferentes normas propostas para o Projeto da Norma Urbana Oral Culta (NURC) (CASTILHO; PRETTI, 1986); o projeto Amostra Linguística do Interior Paulista (ALIP) (TENANI; GONÇALVES, s/d); o C-ORAL-BRASIL (MELLO; RASO, 2009); o Projeto SP2010 (MENDES; OUSHIRO, 2013).

Entretanto, é possível traçar algumas recomendações gerais. Quanto maior o número de critérios, mais heterogêneas tendem a ser as transcrições das gravações do mesmo *corpus*, algo que não é desejável quando um dos principais objetivos é padronizar o material coletado. O fato de que o pesquisador terá a transcrição sincronizada com a onda sonora leva à preferência por um sistema mais simples de transcrição – por exemplo, sem o uso (ou abuso) de caracteres como dois pontos, chaves, parênteses, maiúsculas e minúsculas, a anotação de dados contextuais no meio da fala dos informantes e a indicação de apagamento de segmentos, por exemplo, “pe(i)xe”, “fala(r)”. Muitos pesquisadores transcrevem suas gravações de acordo com a variável a ser analisada, adotando diferentes convenções para cada variante. No entanto, é interessante projetar que tais gravações poderão ser utilizadas em pesquisas futuras, pelo mesmo pesquisador ou por outros, de modo que se recomenda uma transcrição mais “neutra”.

É importante transcrever por extenso tudo o que se ouviu, sobretudo números (por exemplo, “terceiro” em vez de “3º”), pois tais palavras podem conter segmentos-alvo para determinados estudos. Siglas e abreviaturas podem ser grafadas de formas diferentes a depender se são pronunciadas letra a letra (por exemplo, “B.O.”, “I.N.S.S.”, “P.T.”) ou como palavras (por exemplo, “USP”, “IAMSP”, “UFSCAR”). Podem-se criar convenções para marcação de pausas, alongamentos, truncamentos, hesitações e interjeições; no entanto, é preferível que tais anotações não conduzam à criação de novas palavras (por exemplo, indicar silabação por hífen: “com-ple-ta-men-te”), para que não se “inflacione” o número total de palavras do *corpus*. Se se pretende compartilhar o *corpus* com outros pesquisadores, vale a pena revisar as transcrições, preferivelmente por outra pessoa que não o próprio transcritor (para fugir dos vícios de um texto já familiar).

Por fim, a regra mais básica é a aplicação consistente de qualquer norma criada; o pesquisador poderá, então, trabalhar com maior confiança em suas futuras explorações da amostra transcrita.

## REFERÊNCIAS

- CASTILHO, A.; PRETI, D. *A linguagem falada culta na cidade de São Paulo: materiais para seu estudo*, vol. I – Elocuções Formais. São Paulo: T.A. Queiroz, 1986.
- HELLWIG, B.; GEERTS, J. ELAN – Linguistic Annotator. Versão 4.4.0. Disponível em: <<http://www.mpi.nl/corpus/manuals/manual-elan.pdf>>. Acessado em: 11 fev. 2014.
- MELLO, H.; RASO, T. Para a transcrição da fala espontânea: o caso do C-ORAL-BRASIL. *Revista Portuguesa de Humanidades*, Portugal, v.13, n. 1, p. 301-325, 2009.
- MENDES, R. B.; OUSHIRO, L. Documentação do Projeto SP2010 – Construção de uma amostra da fala paulistana. Universidade de São Paulo, 2013. Disponível em: <<http://projetosp2010.fflch.usp.br/producao-bibliografica>>. Acessado em: 11 fev. 2014.
- NAGY, N.; MEYERHOFF, M. Extending ELAN into Variationist Sociolinguistics. Workshop apresentado no NWAV42, Pittsburgh-PA/EUA, 2013. Disponível em: <<http://www.nwav42.pitt.edu/workshops/#elan>>. Acessado em: 11 fev. 2014.
- TENANI, L. E.; GONÇALVES, S. C. L. *Manual do Sistema de Transcrição de Dados - Projeto ALiRP (Amostra Linguística de Rio Preto)*. Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, s/d.

