

ASPECTOS TÉCNICOS NA COLETA DE DADOS LINGÜÍSTICOS ORAIS

Miguel Oliveira, Jr.

INTRODUÇÃO

Existem conjuntos de práticas específicas para a coleta de dados orais que vêm sendo adotados internacionalmente em projetos de documentação linguística. Órgãos como E-MELD School of Best Practice, Open Language Archives Community (OLAC) e o Comitê Técnico da International Association of Sound and Audiovisual Archives (IASA), estudam e propõem tais práticas. É, todavia, ainda um fato recorrente a não observação dessas recomendações técnicas na coleta de dados para estudos linguísticos, seja pela falsa ideia de que o mais importante é o dado de fala em si e não a qualidade do registro, ou simplesmente pela falta de conhecimento técnico do assunto (VAUX; COOPER, 1999).

O objetivo do presente texto é apresentar de maneira concisa, usando uma linguagem acessível, técnicas recentes de recolha de dados orais para documentação linguística e listar recomendações básicas para a sua adequada realização. Não se intenta aqui explorar métodos teóricos de recolha propriamente, algo que vai depender, evidentemente, dos objetivos de pesquisa individuais, mas oferecer sugestões técnicas, baseadas em recomendações feitas por órgãos internacionais de codificação e transmissão de dados de áudio.

1. RECOLHA DE DADOS ORAIS: CONSIDERAÇÕES INICIAIS

A escolha de equipamentos de gravação de áudio para a construção de um *corpus* de fala deve ser guiada por estratégias que garantam um bom resultado do material a ser coletado. A coleta de dados é a base de um bom *corpus*.

A seguir, veremos alguns dos princípios essenciais de gravação de dados de fala que visam a melhoria de sua precisão, de sua consistência e de sua confiabilidade. O que se deve sempre considerar na coleta de dados é que o material coletado pode ser útil para diversos estudos futuros, inclusive estudos que exijam uma alta qualidade dos dados (como, por exemplo, determinadas análises acústicas). É importante, por esse motivo, sempre primar pela qualidade na coleta de dados.

2. GRAVAÇÃO DE ÁUDIO

É fundamental, por razões já expostas, garantir que o arquivo de áudio gravado contenha o maior número possível de informações do sinal original. Essa regra vale tanto para gravações quanto para digitalizações de uma gravação analógica existente.

Em princípio, um gravador deve ser capaz de capturar a resposta de frequência (<10.000Hz) e o alcance dinâmico (40dB), característicos da fala. No entanto, existem outras considerações importantes ao escolher um gravador para coleta de dados de fala.

2.1. Gravação Analógica

A gravação analógica pode capturar a resposta de frequência e o alcance dinâmico, próprios da fala, com precisão e detalhe, sobretudo se forem utilizados um bom microfone pré-amplificado e um gravador de fita magnética profissional. No entanto, os gravadores analógicos geralmente vêm acompanhados de mecanismos de transporte ruidosos, não oferecem a possibilidade de registrar o tempo, o que dificulta a análise e o manuseio dos dados, utilizam um meio (a fita analógica) bastante frágil e, o mais importante: para fazer parte de arquivos digitais (algo altamente recomendável, como veremos mais adiante), as gravações precisam ser digitalizadas, o que implica em perda de informações. É importante ressaltar que equipamentos analógicos e mídias analógicas são cada vez mais difíceis de serem encontrados no mercado, o que torna difícil esse tipo de gravação.

2.2. Gravação digital

Bastante comuns e acessíveis hoje em dia, os gravadores digitais podem ser encontrados nos mais diversos modelos, com características também diversificadas: DAT, *Minidisk*, *solid state*, CD/DVD-R, e *hard disk* são alguns exemplos. O que basicamente os distingue são o meio e o formato de gravação. Algumas dessas tecnologias, como o DAT e o *Minidisk*, são hoje consideradas obsoletas. A tendência atual é a utilização de gravadores *solid state*. Esses gravadores utilizam como mídia cartões de memória estável (*secure digital*), o que os torna menos vulneráveis a choques e vibrações, e desse modo, a interferências/ruídos no sinal registrado.

Na construção de um *corpus* de fala, a recomendação é que os gravadores registrem o áudio em formato não comprimido. Na hora de decidir pelo equipamento adequado, é importante considerar as taxas de amostragem e resolução com que trabalha (idealmente 24 bit/96kHz), as entradas para microfone (XLR), o sistema de alimentação (alimentação fantasma, pilhas, USB), a resposta em frequência de (20Hz - 20kHz) e o alcance dinâmico (> 80dB). É muito importante que o gravador faça registro em formato não comprimido (WAV, por exemplo).

Os gravadores portáteis Marantz da linha PMD, como o PMD662I, estão entre os mais utilizados em coleta de dados orais que exijam alguma qualidade. São modelos portáteis que fazem gravação em mídia SD e têm entrada XLR para microfones profissionais. Pela metade do preço do Marantz PMD66I, o gravador Sony PCM-M10 também é uma boa opção para projetos menos exigentes. A utilização de um bom microfone (como o Audix HT5, por exemplo) poderá, nesse caso, fazer a diferença. Como só há entrada de microfone de 3,5 mm, será preciso adquirir um adaptador XLR-para-3,5mm para usar microfones profissionais.

3. MICROFONES

O microfone é parte fundamental na recolha de dados orais, senão o mais importante. Escolher o microfone que melhor se adeque aos propósitos da recolha e às especificações do equipamento de gravação exige algum conhecimento básico acerca da estrutura e mecanismo de operação desses equipamentos. Os microfones são descritos pelo seu princípio transdutor, por sua direcionalidade e pelo tipo de uso a que se destinam (*design*). Assim, os microfones podem ser dinâmicos, condensadores (capacitivos), omnidirecionais, bidirecionais, cardioídes, hipercardioídes, *shotguns*, de lapela, de mão, *headsets*, etc.

Os microfones dinâmicos são bastante resistentes, não precisam de baterias ou fontes de alimentação externas (o que é uma vantagem), mas não vêm equipados

com pré-amplificadores. Por conta de suas características, podem não responder bem aos sons agudos e/ou transientes, próprios da voz, o que os torna pouco indicados para gravações da fala.

Os microfones condensadores respondem com clareza a sons transientes, o que resulta em sons mais naturais, limpos e ricos em detalhes. Além disso, pesam muito menos e podem ser muito menores. No entanto, porque requerem uma fonte de energia e têm cápsulas muito mais delicadas, podem apresentar algumas dificuldades de uso e maior cuidado de manipulação.

Os microfones também podem ser identificados por suas propriedades direcionais, isto é, pela forma como foram planejados para captar o som de diversas direções. Assim, os microfones podem ser classificados geralmente como pertencendo a um de dois grupos principais: omnidirecional e direcional. Os microfones omnidirecionais captam igualmente sons de todas as direções. Os direcionais podem ter várias configurações, sendo as mais comuns os cardioides, que captam sons provenientes da frente, e os bidirecionais, que registram sons que vêm de direções opostas. Um grande problema de microfones direcionais, como os cardioides, é o chamado efeito de proximidade. O efeito de proximidade refere-se ao aumento na resposta de frequências baixas em função da proximidade da fonte sonora, o que pode causar problemas em análises acústicas. Para evitar o efeito de proximidade, recomenda-se o uso de microfones omnidirecionais.

Em se tratando do *design*, talvez o tipo de microfone mais popular é o chamado microfone de mão: são mais naturais e dão maior liberdade aos usuários; são geralmente utilizados em situações que exigem a troca de falantes, como em entrevistas. Entretanto, são muito sensíveis à ruídos de manuseio e podem ser cansativos se usados por um longo período de tempo. Os microfones de lapela são usados em situações que requerem de seus usuários as mãos desocupadas. Oferecem maior liberdade e espontaneidade, mas são também muito suscetíveis a ruídos provocados por roupa, cabelo e batidas acidentais no microfone ou no cabo. Os microfones *headset* são utilizados quando se quer garantir uma distância fixa e consistente entre o microfone e a boca do usuário, oferecendo ao mesmo tempo uma relativa liberdade de locomoção. Têm sido considerados a melhor alternativa para documentação linguística (PLICHTA, 2004).

Antes de selecionar um microfone, é preciso também levar em conta as suas características técnicas. Além do princípio transdutor, da direcionalidade e do design é preciso atenção para: (i) a impedância (que deve ser menor que 600Ω); (ii) a resposta em frequência (que deve ser entre $20\text{Hz} - 20\text{kHz}$, plana); (iii) o nível máximo de pressão sonora ($> 120\text{dB}$); (iv) a sensibilidade (média, por volta de 6mV/Pa); (v) o ruído próprio ($< 32\text{dBA}$) e (vi) o tipo de conexão. Essas informações são disponibilizadas pelo fabricante, mas podem ser obtidas em organizações independentes, como a Microphone Data <<http://www.microphone-data.com/>>.

O DPA 4006-TL é um dos melhores microfones de mão do mercado para gravação de voz devido à sua alta sensibilidade, excelente resposta em frequência (sobretudo as frequências baixas) e baixíssimo ruído próprio. Por conta de suas características, deve ser usado em ambientes silenciosos, com suporte apropriado e, de preferência, com proteção de explosões (*pop filter*). Se a gravação exige microfone de lapela, o Sanken COS-11DBP é uma excelente opção: o microfone pode ser usado com bateria ou com alimentação fantasma, registra uma vasta gama de frequências e é bastante discreto. O uso de microfones de lapela, no entanto, deve ser considerado com bastante cautela se o objetivo for registrar áudio para análise linguística. Como já apontado acima, esse tipo de microfone capta ruídos indesejados e não registra informações acústicas importantes, por conta de sua localização em relação à boca do entrevistado (discutiremos esta questão em particular logo mais abaixo).

Considerado um dos melhores *headsets* no mercado para registro de voz, o DPA 4066 tem uma resposta em frequência ampla e plana, padrão omnidirecional (o que permite o registro em proximidades muito curtas), média sensibilidade e habilidade de lidar com altos níveis de pressão, o que possibilita a captura de sinais ricos em detalhes, com amplo *headroom*. Conecta-se a praticamente qualquer tipo de entrada, com o adaptador adequado. O Audix HT5, por sua vez, é um *headset* relativamente barato (sobretudo se comparado ao DPA 4066). Possui um *design* bastante confortável e apropriado, um adaptador que provê energia independente (com baterias do tipo AA) e média sensibilidade, o que é ideal para gravações de fala natural, evitando o uso de altos ganhos do pré-amplificador/alimentação fantasma e, assim, possíveis ruídos indesejáveis. A relação *s/n* é bastante boa, assim como a resposta em baixas frequências. Bastante usado em pesquisa acústica, o Opus 55 Mk II tem sido muito elogiado pela crítica especializada por sua habilidade de apresentar respostas muito planas (resultando em um envelope de fala bastante realista) e amplas (registrando harmônicos perto da fundamental). Ao contrário da maioria dos *headsets*, que têm baixa sensibilidade para registrar altos volumes, o Opus 55 tem sensibilidade média, o que é apropriado para voz em altura normal. Requer um adaptador para usar com alimentação fantasma.

4. FORMATO DO ARQUIVO DE ÁUDIO

Na avaliação de equipamentos digitais a serem adquiridos, é preciso sempre levar em consideração o formato com que trabalham. Os equipamentos devem registrar as sessões de coleta de dados nos formatos recomendados por arquivos digitais com reputação internacional.

Recomenda-se que a escolha do formato digital baseie-se nos seguintes parâmetros (SIMONS, 2006):

- sem perdas: o formato deve garantir o registro integral do conteúdo;
- padrões abertos: o formato deve usar um padrão computacional que seja aberto;
- transparente: o formato deve ser acessível;
- usado em muitos aplicativos: é preciso garantir que o arquivo seja aberto no futuro.

O formato não comprimido WAV tem se tornado padrão em arquivos destinados a pesquisas linguísticas, sendo o mais recomendado por órgãos de referência e, em muitos casos, o único formato aceito para preservação em bancos de dados internacionais.

5. RECOMENDAÇÕES PARA GRAVAÇÃO DE ÁUDIO

No que diz respeito à gravação de dados propriamente, um dos primeiros cuidados que se deve ter é estabelecer níveis corretos de registro de áudio. Trata-se de um procedimento que deve ser feito pontualmente, ou seja, não é possível estabelecer um nível padrão para todas as situações de gravação.

A regra geral para definir os níveis de gravação é regular o ganho do amplificador em um nível que permita uma gravação em uma altura adequada, com um amplo *headroom*, sem causar sobrecarga e/ou clipagem. Essas recomendações servem para gravações em que o microfone possa estar sempre perto da boca, ou com microfones pré-amplificados de alta qualidade.

Em gravações feitas em salas silenciosas ou acusticamente tratadas, recomenda-se gravar o mais baixo possível, deixando um amplo *headroom* que possa ser manipulado posteriormente, caso se faça necessário. Em ambientes ruidosos, com um microfone montado à cabeça, recomenda-se a gravação em um nível de -12dBFS para garantir uma relação sinal-ruído favorável e detalhes espectrais adequados. Se o microfone pré-amplificado gerar muito ruído nesse nível, recomenda-se a diminuição do ganho em 6dB.

O processo de registro de dados orais deve estar amparado não apenas no conhecimento técnico dos equipamentos utilizados no registro em si, mas também – e sobretudo – no conhecimento da metodologia empregada no registro de voz. Para isso, é fundamental seguir as orientações sugeridas por órgãos reconhecidos, como, por exemplo, o Open Archival Information System (OAIS), que é um modelo de referência, com padrão ISO 14721:2003, adotado pelos bancos de

dados linguísticos mais recentes, e o Comitê Técnico da International Association of Sound and Audiovisual Archives (IASA).

A recomendação mínima que se dá para gravação digital de áudio é a utilização de uma taxa de amostragem igual ou superior a 44.100Hz, com uma resolução superior a 16bit, em formato não comprimido (.wav). Estudos recentes em psicoacústica, no entanto, revelaram que níveis de amostragem mais altos têm efeitos perceptuais significativos, o que aponta para a recomendação de uma taxa de amostragem mínima de 96kHz (PLICHTA, 2004). As gravações devem ser feitas em apenas um canal (monoaural).

5.1. Ruído

No momento da gravação, é preciso muita atenção para as seguintes possíveis fontes de ruído: (I) nas proximidades: pessoas, animais, atividades (domésticas ou outras); (II) nas redondezas: o tráfego, geradores, aviões; (III) máquinas: geladeiras, ventiladores, computadores, telefones celulares; (IV) não audível: interferência elétrica; (V) eventos imprevisíveis: farfalhar de papéis ou de roupa, mesa ou cadeira em desequilíbrio; (VI) eventos previsíveis (que podem ser evitados): tamborilar em mesa, *feedback* do pesquisador; (VII) manuseio do equipamento: mouse, teclados, microfone, cabos; (VIII) gerados por equipamentos: cabos inapropriados ou baratos, nível de sinal de entrada inapropriado, etc.

Há situações em que o ruído é inevitável. Nelas, é preciso encontrar uma solução para minimizar a sua interferência na gravação. Para evitar o trânsito, o ideal é investigar horários de pico e gravar apenas fora deles. Se não for possível, utilizar um microfone unidirecional. Para evitar interferência de fenômenos naturais (vento, chuva, trovão), use um “gato morto”: acessório protetor de vento acoplado ao microfone. O “gato morto” pode também ser útil para lugares com motores, ou geradores intermitentes. Nesses casos, no entanto, o indicado é garantir que serão desligados ou procurar um outro lugar. Crianças e animais precisam ser monitorados. O ideal é investigar, quando possível, o melhor momento.

5.2. Distância do microfone

Um dos principais fatores para garantir um sinal de áudio de boa qualidade é a observação de uma distância mínima (e constante) entre a boca do participante e o microfone. A recomendação de distância mínima (e máxima) varia de acordo com as características e especificidades do microfone utilizado, mas, em geral, gira em torno de 30 cm a 15 cm para a maior parte dos microfones.

De acordo com a lei do inverso do quadrado, dobrando-se a distância entre o microfone e a fonte, a energia (ou intensidade) sonora registrada cai quatro vezes, aumentando as chances de se captar sinais indesejados. O ideal, portanto, é que o registro de voz deva ser feito com o microfone posicionado o mais próximo possível da boca do participante. Mas é preciso atenção para o efeito de proximidade, tal como observado acima.

Microfones do tipo *headset* ficam a uma distância relativamente curta da boca do participante (1-5cm), com a cápsula posicionada ao lado da boca (e não à frente ou abaixo). Para evitar o efeito de proximidade, a maioria dos microfones *headset* são omnidirecionais. Além disso, os melhores microfones desta categoria são equipados com *shock-mounts*, o que garante gravações livres de ruídos causados por choques mecânicos.

RECOMENDAÇÕES FINAIS

O presente texto teve por objetivo apresentar métodos recomendados para registro de dados orais em pesquisas linguística e discutir propriedades técnicas úteis para a escolha de equipamentos adequados, garantindo assim a qualidade dos dados coletados. Essas recomendações, todavia, não devem ser consideradas empecilhos para a recolha de dados em situações extremas. Como bem observa Bower (2008), uma gravação feita em uma situação distante do ideal é melhor que nenhuma gravação. É preciso considerar que as condições de coleta de dados nunca serão perfeitas. Portanto, o que se recomenda é sempre planejar cuidadosamente a recolha de dados, e na hora, lançar mão do melhor que se puder.

REFERÊNCIAS

- BOWERN, C. *Linguistic fieldwork: a practical guide*. Basingstoke: Palgrave Macmillan, 2008.
- PLICHTA, B. Data acquisition problems. *Signal acquisition and acoustic analysis of speech*, 2004. Disponível em: <http://bartus.org/akustyk/signal_aquisition.pdf>. Acesso em: 11 ago. 2013.
- SIMONS, G. F. Ensuring that digital data last: the priority of archival form over working form and presentation form. *SIL Electronic Working Papers 2006-003*. Boston: mar 2006.
- VAUX, B.; COOPER, J. *Introduction to Linguistic Field Methods*. Munich: Lincom Europa, 1999.