

## Práticas de pesquisa microetnográfica: geração, segmentação e transcrição de dados audiovisuais como procedimentos analíticos plenos

### *Microethnographic research practices: generating, segmenting and transcribing audiovisual data as full-fledged analytic procedures*

Pedro de Moraes GARCEZ (UFRGS)

Gabriela da Silva BULLA (UFRGS)

Letícia Ludwig LODER (UFRGS)

#### RESUMO

*Apresentamos a perspectiva analítica e os procedimentos de geração, segmentação e transcrição de dados que adotados e desenvolvidos ao longo de mais de dez anos de pesquisas em Linguística Aplicada no Grupo de Pesquisa Interação Social e Etnografia para a análise da ação social humana mediada pelo uso da linguagem. A produção de asserções analíticas a respeito das atividades dos participantes de um encontro social requer observação detalhada de registros audiovisuais, exigindo revisitação para o aprofundamento das investigações. Por isso, as etapas de geração, segmentação e transcrição da fala-em-interação nesses encontros, em vez de preliminares à análise, são concebidas como partes integrantes da atividade analítica em si. Destacamos ainda nossas propostas de procedimentos para transcrição multimodal e a importância de exercício em sessões de análise conjunta.*

**Palavras-chave:** *métodos de pesquisa; fala-em-interação; microetnografia; transcrição.*

## ABSTRACT

*An account is offered here of the analytic perspective and procedures for generating, segmenting and transcribing interactional data employed and developed during more than a decade of applied linguistic research within the Social Interaction and Ethnography Research Group for the analysis of social action accomplished in language use. Analytic assertions about the activities of participants of a social encounter require detailed observation of audiovisual records, demanding revisitation for careful research. Therefore generation, segmentation and transcription of talk-in-interaction are considered integral steps of analysis in and of themselves. Proposals for multimodal transcription are highlighted, and the importance of training in data analysis sessions is emphasized.*

**Key-words:** *research methods; talk-in-interaction; microethnography; transcription.*

## Introdução

Neste artigo, oferecemos um relato dos procedimentos de pesquisa que vimos praticando e desenvolvendo em projetos de investigação em Linguística Aplicada do Grupo de Pesquisa ISE – Interação Social e Etnografia<sup>1</sup> –, nos quais a geração, a segmentação e a transcrição de dados audiovisuais são tarefas centrais. Acreditamos que nossa experiência possa ser útil para pesquisadores que desejam iniciar projetos dessa natureza, bem como para os investigadores cuja pesquisa envolve a análise de fala-em-interação de maneira suplementar.

Primeiramente, explicitamos as bases gerais que sustentam a produção de conhecimento em metodologia qualitativa interpretativa e, particularmente, a nossa necessidade, como pesquisadores interessados em fala-em-interação, de trabalhar com dados de ocorrência natural, que precisam ser registrados de algum modo. Nas seções seguintes

---

1. Grupo de pesquisa dedicado ao estudo da organização da fala-em-interação social em diversos cenários interacionais – especialmente os escolares. Sob a coordenação do professor Pedro de Moraes Garcez, o ISE foi iniciado em 1998, na Universidade Federal de Santa Catarina, e, desde 2002, está vinculado à Universidade Federal do Rio Grande do Sul e registrado no Diretório de Grupos de Pesquisa do CNPq.

(2. Geração de dados, 3. Segmentação e 4. Transcrição), abordamos três etapas do trabalho do pesquisador que lida com dados desse tipo, apresentando questões teórico-práticas inerentes a esses procedimentos analíticos. Ainda na seção sobre transcrição, destacamos, em subseções, (a) um sistema de transcrição amplamente adotado internacionalmente e (b) nossas experiências e escolhas frente ao tratamento de elementos multimodais inerentes a encontros face a face. Em seguida, destacamos a necessidade de treinamento e de sessões de análise conjunta de dados para a qualidade de empreendimentos investigativos dessa natureza. Por fim, apontamos algumas contribuições de trabalhos que compartilham algumas das práticas de pesquisa microetnográfica aqui relatadas, visando exemplificar brevemente o tipo de conhecimento que pode ser produzido a partir delas.

## **1. Metodologias interpretativas para a produção de conhecimento**

O que descrevemos aqui são procedimentos de pesquisa, isto é, de “investigação e estudo, minudentes e sistemáticos, com o fim de descobrir ou estabelecer fatos ou princípios relativos a um campo qualquer do conhecimento” (Ferreira 1999:1556), ou seja, procedimentos de produção de conhecimento. Para isso, ao longo dos últimos séculos e especialmente das últimas décadas, foram desenvolvidos materiais e métodos eficazes em áreas de investigação científica dominantes, como a Física e a Química, que garantiram à humanidade os avanços tecnológicos de que hoje desfrutamos. Assim, é comum equacionar produção de conhecimento com experimentação, isto é, a testagem de hipóteses mediante a observação de um fenômeno no qual se interfere em um aspecto ao mesmo tempo em que se controlam as demais variáveis, buscando, com isso, estabelecer relações seguras de causa e efeito.

De fato, para estabelecer com precisão e autoridade “científica” o “impacto disso naquilo”, “a influência de tal aspecto sobre tal outro elemento”, ou “como tal elemento condiciona, determina, afeta um outro”, nada melhor do que um bom experimento, ou, em se tratando de populações, um bom levantamento censitário. Em qualquer dessas e outras metodologias tradicionais de produção de conhecimento, contudo, é preciso abrir mão, necessariamente, da vida como ela é,

isto é, das relações recíprocas entre os elementos sob observação tal como essas relações se colocam situadamente, sob a perspectiva dos observados, no tempo e no espaço físico.

No entanto, como linguistas aplicados, queremos produzir conhecimento sobre as ações dos membros de uma espécie social que atuam em grupos sociais e em comunidades de práticas situadas. Diante de um encontro de pessoas em um cenário de interesse, a pergunta geral que nos move é “o que está acontecendo aqui e agora?”. Para produzir conhecimento em nossa área, não queremos e não podemos prescindir de observar a vida como ela é entre aquelas pessoas, naquele encontro, naquele cenário, ali-e-então. E talvez por isso não possamos ter a pretensão de estabelecer relações de causa e efeito. Afinal, causa, na ação humana, é algo muito diferente de causa nos universos físico e químico; as pessoas fazem sentido à medida que agem e mudam o sentido que fazem de seus comportamentos em momentos e lugares diferentes. Assim, se, de fato, no mais das vezes, de nada interessam as perspectivas dos elementos “reagentes” físico-químicos sob observação na produção de conhecimento segundo a tradição do método científico, na produção de conhecimento sobre as ações humanas que estamos interessados em compreender em cenários diversos da vida social, buscamos articular o conhecimento particular, específico, da ecologia de vida e da experiência continuada dos participantes da ação social (o que pode estar opaco para eles próprios).

Estamos, portanto, tratando aqui de produção de conhecimento em metodologia qualitativa interpretativa. Segundo Erickson (1990: 106-108),

*A tarefa do pesquisador interpretativo é descobrir as maneiras específicas em que as formas locais e não locais de organização social e cultural se relacionam com as atividades de pessoas específicas no processo de fazer escolhas e conduzir a ação social em conjunto.*

...

*A tarefa do analista é expor as diferentes camadas de universalidade e de particularidade que se apresentam no caso específico sendo examinado.*

Segundo essa ótica, o critério básico de validade para o trabalho investigativo é o entendimento dos sentidos das ações conforme esses

sentidos se definem da perspectiva dos atores, o que envolve trabalho de campo: observação, participação, registro, reflexão analítica com base nos registros e relato descritivo, narrativo, persuasivo. O privilégio à perspectiva que os participantes demonstram uns para os outros acerca de ‘o que está acontecendo no seu aqui-e-agora’ é um critério determinante para a natureza dos procedimentos apresentados nesse trabalho, exigindo um elevando grau de empenho do pesquisador. Conforme apresentado a seguir, transcrevemos décimos de segundos na notação de silêncios e fazemos um esforço extraordinário para registrar certos movimentos corporais porque nossas análises da fala-em-interação indicam que esses aspectos importam para os interagentes no seu aqui-e-agora interacional, não porque sejamos fascinados por essas minúcias.

A tarefa do pesquisador interpretativo, portanto, é ser tão completo quanto possível no ato de notar e descrever a atividade cotidiana de modo a identificar a significação das ações para os participantes. Além disso, por ser deliberadamente interpretativa, a produção de conhecimento conforme concebida aqui entende que privilegiar a perspectiva dos atores passa necessariamente pela atenção crítica às perspectivas dos próprios analistas, que, para serem subordinadas analiticamente às perspectivas dos atores, não podem ser negligenciadas. Em outras palavras, a tarefa do pesquisador durante o trabalho de campo é se tornar cada vez mais consciente e reflexivo acerca dos quadros interpretativos das pessoas observadas e de suas próprias lentes interpretativas trazidas para o cenário.

Por entender que a fala-em-interação é intersubjetivamente construída, acreditamos que a sua análise pode subsidiar a produção de conhecimento acerca da organização de ocorrências reais, particulares, ecologicamente situadas, de fenômenos da vida social que, ao pesquisador e à sociedade, importe compreender. Desse modo, espera-se que a descrição de elementos da ecologia de situações, especialmente daquelas que não se pode controlar (ensino ruim ou bom, por exemplo), possa ser feita via contraste e complementação entre diversas análises de ocorrências naturais em cenários comparáveis. É o que chamamos de pesquisa microetnográfica (Garcez 2008), cujo objetivo consiste em descrever como a fala-em-interação (face a face ou mediada por tecnologias) é organizada social e culturalmente em

cenários interacionais particulares (como as salas de aula), partindo de um ponto de vista próprio sobre o uso da linguagem em encontros interacionais entre múltiplos participantes em sociedades complexas contemporâneas (ver Erickson 2004).

Para tanto, trabalhamos com registros audiovisuais de encontros sociais, gerados por meio de equipamentos de gravação, a fim de investigar em detalhe minucioso o que os interagentes fazem em tempo real à medida que coconstroem a vida cotidiana pela fala-em-interação. É preciso gerar e analisar esses registros, o que, por sua vez, envolve providências e decisões de natureza prática, cuja coerência teórico-metodológica é complexa e que passamos a detalhar nas seções a seguir.

## 2. Geração de dados

Empregamos o termo “geração” em vez do mais corrente “coleta” por entendermos que a vida social que nos interessa compreender é em si evanescente e que não pode ser captada integralmente por nenhum aparelho ou método de gravação. O que examinamos em nossas análises são registros que efetivamente geramos, desde a própria gravação, o que implica escolher um equipamento a ser disposto em algum lugar, um ângulo de diafragma que seleciona parte do campo visual disponível aos atores sociais no ali-e-então, um “operador” que ocupa lugar e participa, uma qualidade de áudio distinta daquela disponível aos atores sociais no ali-e-então. Além disso, nossos procedimentos envolvem grandes transformações dos registros gerados até chegarmos a um excerto de transcrição, que, em geral, é tudo o que o interlocutor de nossa produção acadêmica avista diretamente.

Se, até o início desta década, era comum os pesquisadores se ocuparem com recomendações técnicas detalhadas acerca de equipamentos para gravação da fala-em-interação, atualmente estão de tal modo popularizados os equipamentos de registro audiovisual em mídia digital que tais preocupações se tornaram não apenas obsoletas como também dispensáveis. Mantêm-se, contudo, algumas orientações gerais.

Em primeiro lugar, o critério geral é que não queremos estar, de saída, em condições distintas daquelas em que os participantes se

encontravam durante o encontro social. Por isso, procuramos sempre registrar eventos completos, isto é, desde a formação do encontro dos participantes até a sua dispersão e encerramento. Além disso, pela mesma razão, a escolha do tipo de registro (audiovisual ou apenas em áudio) deve ser análoga aos recursos interacionais de que os participantes dispõem. Desse modo, se eles estão em condições de monitoramento mútuo pela visão (na fala-em-interação social face a face ou mediada por videotecnologia), é certo que irão se valer de elementos multimodais vocais e não vocais (fala, gestos etc. – ver subseção d. Multimodalidade, abaixo), o que requer a geração de registros audiovisuais. O registro apenas em áudio se justifica quando os participantes estão interagindo nessa mesma condição, como é o caso de interações em chamadas de voz por telefone ou audioconferências via internet, cujo caráter multimodal do uso da linguagem se resume a elementos vocais e sonoros. Embora no passado tenhamos nos valido de alguns dados de interação face a face registrados apenas em áudio (particularmente em razão do alto custo e da conseqüente escassez de equipamentos para videogravação analógica de qualidade profissional), esse procedimento nos parece pouco justificável atualmente, face à popularização dos aparelhos para gravação audiovisual digital, como destacamos acima. A indisponibilidade de registros audiovisuais desses tipos de evento impede a análise multimodal, o que, na nossa experiência mais recente, limita o estudo detido das ações dos participantes a ponto de tornar a análise dos dados inviável em muitas ocasiões.

É verdade que há casos em que a gravação em vídeo é especialmente melindrosa, o que a nosso ver requer negociação de entrada e ambientação extensas. Por exemplo, em um projeto de pesquisa de microetnografia escolar que desenvolvemos entre 2001 e 2009 (Conceição & Garcez 2005; Garcez 2006; Garcez & Melo 2007; Schulz 2007; Conceição 2008; Almeida 2009), as gravações audiovisuais foram iniciadas somente em 2003, após dois anos de aproximação, negociação de entrada e observação participante. Assim, antes da gravação propriamente dita em cenários como salas de aula, é recomendável um período de ambientação dos participantes ao equipamento, um período em que o pesquisador apresenta fisicamente o equipamento e, se for adequado, inclusive permite que os participantes manuseiem o equipamento e satisfaçam sua curiosidade. Na microetnografia escolar aludida acima, as crianças tinham pouca familiaridade com câmeras de

vídeo. A curiosidade satisfeita pela oportunidade, reiterada à exaustão, de enxergar através da objetiva da câmera e assistir a gravações da própria turma ajudou a dissipar a intrusão da presença do equipamento, que foi se tornando parte da rotina. Por outro lado, na geração de dados de Rosa (2008) em uma sala de aula de crianças da mesma idade, não houve necessidade desse tipo de ambientação, pois as crianças manifestaram escassa curiosidade pelo equipamento de gravação.

Em nossa prática de geração de dados, procuramos ter o máximo de abrangência de visão do evento interacional, o que significa que, para um ambiente como o de uma sala de aula típica, é recomendável ter ao menos duas filmadoras dispostas em oposição. Também é recomendável haver gravadores de áudio adicionais e independentes da videocâmera, o que pode ser muito útil quando, por exemplo, uma turma, mesmo que pequena, como em um curso livre de idiomas, se divide em diversos grupos, e a gravação em áudio das videocâmeras registra um conjunto de vozes em geral impossível de distinguir e muito menos de transcrever. Mesmo numa reunião de negócios com poucos participantes (Garcez 1993; 2004), a gravação auxiliar em áudio pode ser crucial para se entender certos trechos no momento de transcrever<sup>2</sup>.

Outro critério geral de posicionamento das videocâmeras é mantê-las estacionárias (usando tripés de sustentação) e com foco parado tanto quanto possível. Isso nem sempre se mantém, pois é comum que fique patente para o operador da filmadora/pesquisador/participante que um foco de atenção inusitado se estabeleceu a certa altura entre os participantes. Se eles saírem do escopo da gravação, portanto, justifica-se usar o dispositivo de *zoom*, alterar a posição do equipamento ou mesmo movimentar-se com ele para acompanhar os participantes.

Não se justifica, sob hipótese alguma, a gravação de fala-em-interação sem o consentimento expresso dos participantes. Antes de acionar equipamentos, tomamos providências para assegurar, em documento escrito, o consentimento dos participantes e, no caso de menores de

---

2. É claro que equipamentos sofisticados para captação de áudio em excelente qualidade estão disponíveis, como gravadores de lapela individuais com transmissão via rádio, permitindo que os participantes se movam sem impedimentos (ver Rampton 2006). Não tratamos aqui de questões relativas ao uso desse tipo de equipamento porque deles não dispomos em nossas atividades habituais de pesquisa.



idade, dos seus responsáveis. A consulta a qualquer das dissertações e teses citadas aqui levará o leitor ao formulário de consentimento utilizado no trabalho, reproduzido em apêndice.

A pouca familiaridade dos participantes com a noção de pesquisa acadêmica e a falta de balizas claras e exequíveis na realidade brasileira, em especial fora da área da pesquisa em saúde, tem nos feito trilhar um caminho de simplificação dos documentos de obtenção de consentimento<sup>3</sup>. Modelos de consentimento muito detalhados impuseram restrições limitadoras (ver Abeledo 2008), e a interlocução com leitores que têm pouca escolaridade revelou que tínhamos pressupostos infundados, como o de que eles soubessem o que era “UFRGS”, por exemplo. Dadas essas limitações, não afirmamos que nossos procedimentos produzam “termos de consentimento livre e esclarecido”; nos esforçamos para que sejam, sim, e tomamos diversas medidas para que os participantes não tenham o que temer e se sintam legitimados a discutir os limites de seus consentimentos.<sup>4</sup>

### 3. Segmentação

Uma vez gerados os dados, inicia-se outra fase analítica, bastante trabalhosa, que precede a transcrição. Em geral, os eventos interacionais que estudamos são bastante longos, de modo que transcrevemos apenas segmentos específicos de determinado registro de fala-em-interação e só raramente o registro completo. A segmentação do fluxo contínuo da interação para que se possa chegar a trechos transcritos é uma operação analítica em si. Assim, empregamos o termo *segmento* para nomear uma unidade analítica e *excerto* como termo metatextual para referir a trechos de transcrição em sessões de análise conjunta ou nos relatórios de pesquisa.

---

3. Um ponto de referência relevante para nossas decisões é a experiência relatada por especialistas em bioética que atuam no Comitê de Ética em Pesquisa do Hospital de Clínicas de Porto Alegre, em Goldim (2006) e Goldim, Clotet & Ribeiro (2007).

4. Movimentos recentes de exigências de que todos os projetos de pesquisa no País se submetam à Resolução 196/1999 do Ministério da Saúde vão requerer ajustes a esses procedimentos.

Decidir o que é um segmento, onde começa e onde termina é, portanto, um procedimento analítico crucial. Como diz Moerman (1988), “impor divisas por razões de uma teoria exógena ou por conveniência é [como] estudar flores silvestres com um cortador de grama” (Moerman 1988: 72). Por isso, para a seleção de segmentos com base na perspectiva dos participantes, há um esforço a ser feito de modo a encontrar sustentação para o que se vê como marcos divisórios relevantes na interação. A segmentação do registro audiovisual é um procedimento indutivo, guiado pela atenção ao fato de que a ação social é composta de unidades que têm divisas consensuais construídas pelos participantes como parte da inteligibilidade que eles produzem. Em outras palavras, “as unidades [que a fala-em-interação e a sua análise exigem] são produzidas pelos falantes e por isso podem ser encontradas pelo analista” (Moerman 1988: 72).

Em suma, esperamos do analista que ele seja capaz de responder por que o segmento que apresenta é uma unidade organizacional segundo a perspectiva dos participantes e, se tiver uma coleção de ocorrências de algo que aponta como um mesmo fenômeno, por que cada ocorrência tem tal estatuto. Essa expectativa, por sua vez, pode constituir uma orientação bastante prática como procedimento analítico, pois conduz o pesquisador a examinar os registros interacionais com atenção analítica desde o início. E, de fato, há diversos elementos na conduta dos participantes que indiciam a visão que eles têm de que há marcos divisórios com “antes e depois”, que vão desde os mais evidentes, como lapsos de silêncio (cf.: Sacks, Schegloff & Jefferson 2003[1974]: 33 e 70) e alterações de configuração de posturas corporais, até os mais sutis, como alterações de enquadre e alterações de padrões de direcionamento de olhar. Erickson e Shultz (2002[1981]) é leitura bastante útil como guia de orientação teórico-metodológica para o procedimento de segmentação.

#### 4. Transcrição

Assim como a geração e a segmentação dos registros, tratamos a elaboração de transcrições como atividade analítica plena, conforme já tivemos oportunidade de discutir em Garcez (2002) e Loder (2008). Seguem atuais as observações de Duranti (1997: 161):

- i) a transcrição é um processo *seletivo*, que busca salientar certos aspectos da interação, de acordo com metas investigativas específicas;
- ii) não há transcrição *perfeita*, no sentido de uma transcrição que possa recapturar inteiramente a experiência de se estar na situação original, mas há transcrições *melhores do que outras*, isto é, transcrições que representam as informações de maneiras que são (mais) consistentes com as nossas metas descritivas e teóricas;
- iii) não há uma transcrição *final*, apenas versões *diferentes, revisadas*, de um texto de transcrição anterior para um propósito específico, para uma plateia específica;
- iv) os textos de transcrição são *produtos analíticos*, que precisam ser continuamente atualizados e comparados com o material a partir do qual foram produzidos (...);
- v) devemos ser tão *explícitos* quanto possível sobre as escolhas que fazemos ao representar as informações na página (ou na tela);
- vi) os formatos de transcrição variam e devem ser avaliados com relação às metas que devem atingir;
- vii) devemos estar criticamente *conscientes* das implicações teóricas, políticas e éticas do nosso processo de transcrição e dos produtos finais que dele resultam.

Bucholtz (2000) oferece também uma discussão alinhada com essas posições e com nossas práticas ao afirmar que “não existe uma transcrição desinteressada” (Bucholtz 2000: 1446) e que “não há como não tomar (...) decisões” (Bucholtz 2000: 1463). Portanto, há decisões importantes que cabem ao analista/transcritor tomar, no nosso caso considerando o compromisso com a busca por se aproximar da perspectiva dos participantes e das especificidades do aqui-e-agora de cada ecologia sob análise. Por isso mesmo, tais decisões devem ser continuamente reavaliadas a cada novo trabalho, considerando as evidentes potencialidades e limitações de cada alternativa. Em suma, mais do que mero preciosismo, a reflexão cuidadosa a respeito dessas escolhas é também parte do trabalho analítico, uma vez que tem implicações de ordem prática (a transcrição deve registrar o mais próximo e reconhecivelmente possível os fenômenos de interesse analítico central) e política (a transcrição constitui os participantes para os leitores da transcrição)<sup>5</sup>.

5. Para outros apontamentos sobre o planejamento e a geração de gravações até o processo de transcrição e algumas de suas implicações, ver Schnack, Pisoni & Ostermann (2005).

O registro gráfico para transcrever a fala dos participantes da interação<sup>6</sup>, portanto, deve ser escolhido com atenção e consistentemente empregado em cada segmento. Dentre múltiplas opções possíveis, tipicamente adotamos, em trabalhos de pesquisa alinhados à perspectiva microetnográfica, uma das seguintes alternativas para transcrever falas: ortografia, grafia modificada ou alfabeto fonético.

O uso da ortografia facilita a leitura da transcrição para um público mais amplo, na medida em que não exige “alfabetização prévia”, além de reproduzir para os leitores a percepção dos próprios participantes acerca da “normalidade” de suas falas, sem aspectos desviantes que as tornem potenciais focos de investimento interacional. Entretanto, justamente por isso, esse procedimento nivela diferenças que podem ser relevantes para os participantes (por exemplo, diferenças dialetais, fenômenos em variação, alternâncias de código ao longo da interação, brincadeiras com os sons de palavras, imitações de sotaques ou apresentação de marcas de outras línguas diferentes da língua de interação), apresentando suas falas de modo idealizado (ver Duranti 1997: 125).

O uso de grafia modificada (por exemplo, *tá* para *está*), enquanto mantém, em certa medida, os benefícios da ortografia (de leitura mais acessível), permite um registro mais próximo de como a elocução foi vocalmente proferida. Porém, nessa tentativa, se desnaturalizam fenômenos comuns na língua falada<sup>7</sup>, tornando-os possivelmente exóticos para os leitores da transcrição, que “podem fazer inferências injustificadas sobre os falantes cuja fala está representada de modo não padrão” (Bucholtz 2000: 1455) e ter a impressão (negativa) de que esses participantes têm uma fala “desviante” ou que são “desleixados” e “ignorantes”<sup>8</sup>. Além disso, dado que “nenhum princípio parece estabelecer quando a grafia não padrão deve ser usada” (Bucholtz 2000: 1458), há possibilidade de inconsistências na representação modificada

---

6. Em geral utilizamos fonte *Courier*, pois todos os caracteres ocupam o mesmo espaço, o que facilita o alinhamento das falas sobrepostas, que são registradas em linhas diferentes.

7. Por exemplo, em português brasileiro, o apagamento do -r final em verbos no infinitivo (grafar *falá* em lugar de *falar*), a monotongação em *cadeira* ou *falou* (grafar *cadera*, *falô*) etc.

8. A esse respeito, ver também Jefferson (1983), Gago (2004:97) e, em um quadro teórico distinto, Zilles & Faraco (2002).

de um mesmo fenômeno (inclusive no âmbito de uma mesma transcrição, como registra Jefferson 1996).

O uso do alfabeto fonético é bastante raro entre nós devido à dificuldade que impõe a leitores não familiarizados às suas convenções sem acarretar, em princípio, nenhum benefício adicional à análise microetnográfica da interação. Entretanto, “há situações em que a fala dos participantes (...) é melhor representada foneticamente. Elas incluem instâncias de brincadeiras com sons e instâncias de fala ininteligível” (Ochs 1979: 61). Assim, quando algum fenômeno dessa natureza é demonstravelmente relevante para os participantes, o uso de alfabeto fonético é justificado.

Nossa orientação geral, portanto, tem sido modificar a ortografia apenas nos casos em que se observa que os participantes manifestam atenção a alguma produção como algo especial. Portanto, não registramos apagamento de -r nas formas verbais ou na monotongação. Um caso particularmente difícil é a forma *vamos*, pronunciada tão comumente como *vamo* a ponto de nos fazer crer que grafar *vamos* indicaria uma fala cuidada. Há formas, contudo, muito produtivas na fala em português brasileiro, em que sílabas inteiras deixam de ser pronunciadas, sobretudo no caso do verbo *estar*. Nesses casos, adotamos a grafia consagrada em tiras e histórias em quadrinhos (*tá, tô, tive*).

As especificidades da língua de interação dos participantes, particularmente no que tange a interjeições e atividades vocais não lexicais, também exigem do analista/transcritor escolhas de como representá-las por escrito de modo a não causar estranhamento no leitor a ponto de tornar esses itens irreconhecíveis<sup>9</sup>. Um exemplo recorrente é a grafia de práticas vocais para sinalizar, por exemplo, concordância ou atenção ao interlocutor sem reivindicar o turno, isto é, continuadores. Enquanto parece haver certo consenso entre transcritores norte-americanos da fala-em-interação em inglês em torno da grafia *uh huh*<sup>10</sup>, em português

9. No âmbito da comunidade acadêmica brasileira, o artigo de Gago (2004) inaugura algumas dessas discussões.

10. Para o que norte-americanos grafam como *uh*, transcritores britânicos preferem *er*, registrado como entrada no dicionário Collins Cobuild, como interjeição, “um som feito ao hesitar na fala” [<http://www.collinslanguage.com/results.aspx?context=3&reversed=False&action=define&homonym=-1&text=er>]. *Huh* também é registrado como entrada nesse dicionário. Não temos conhecimento de registro de interjeições equivalentes em dicionários de língua portuguesa.

brasileiro, há várias tentativas (*ãhã, ãrrã, ahn hã, ahn ham*, entre outras) que, além de buscar reproduzir os sons tal como emitidos, também podem destacar a distinção entre uma batida rítmica (por exemplo, *ãhã*, representada graficamente como uma “unidade”) e duas batidas rítmicas (por exemplo, *ahn ham*, representada como duas “unidades”). A transcrição desses itens ainda não está consolidada no âmbito da comunidade acadêmica brasileira, ficando as discussões, por ora, restritas a grupos de pesquisa locais.

Por fim, é central que o analista/transcritor decida como nomear os participantes na transcrição, também balizado pela busca por aproximar-se da ótica dos participantes no ali-e-então. Uma vez que temos o compromisso de preservar a identificação dos participantes (bem como de nomes de terceiros e locais mencionados), assegurada nos termos de consentimento informado, empregar os nomes reais não é uma opção. Algumas alternativas vêm sendo adotadas.

A prática de identificar os participantes por letras do alfabeto (“A”, “B”) evita a imposição de categorias identitárias *a priori* aos participantes. Essa alternativa ocorre especialmente em esquemas e na explicação da organização genérica de algum fenômeno. No entanto, o uso de letras em ordem alfabética pode indicar um indesejado ranqueamento entre os participantes (A tendo algum privilégio ou maior importância em relação a B), além de tratá-los como indivíduos sem personalidade própria, ignorando diferenças entre eles potencialmente relevantes à interação (por exemplo, em relação ao conhecimento sobre algum assunto, à posição institucional, ao gênero etc.).

A identificação dos participantes por categorias institucionais (por exemplo, “professor”/“aluno”, “médico”/“paciente”, “juiz”/“depoente”) ao longo de toda a transcrição é opção comumente empregada em estudos dedicados a analisar as particularidades interacionais de contextos institucionais específicos<sup>11</sup>. No entanto, isso pode enviesar a leitura do interlocutor da transcrição, uma vez que enquadra o participante, desde o início, em uma única categoria identitária, negligenciando outras que possam se tornar salientes e relevantes, para os próprios participantes, ao longo da interação.

11. Ver, por exemplo, Heritage e Sefi (1992), Have (1999: 170-180) e Paoletti e Fele (2004).

A alternativa que em geral adotamos é empregar pseudônimos, com o cuidado de escolher “nomes” com o mesmo número de sílabas e, se possível, de caracteres do nome real. Essa opção não desumaniza os participantes (como o uso de letras), evita a imposição de categorias *a priori* ao longo de toda a interação (como as categorias institucionais) e se aproxima da maneira como as pessoas normalmente nomeiam-se umas às outras nas interações rotineiras (por nomes próprios)<sup>12</sup>. Ainda assim, salienta certas associações, como gênero gramatical (os nomes próprios, em geral, indiciam masculino ou feminino), o que pode eventualmente suscitar, ao leitor da transcrição, uma ênfase na identidade de gênero quando não necessariamente é central para os atores sociais naquele ali-e-então.

Na elaboração de transcrições, algumas de nossas decisões seguem as de outros pesquisadores, escolhas já estabelecidas e consolidadas no âmbito da pesquisa microetnográfica, como é o caso de nossa adoção das convenções de transcrição Jefferson, desenvolvidas pela tradição em pesquisa em Análise da Conversa Etnometodológica, a principal adesão teórico-metodológica dos membros do Grupo de Pesquisa ISE.

### **a. Convenções Jefferson**

Para além das escolhas de registro das elocuições por escrito e identificação dos participantes, são relevantes, em pesquisa microetnográfica, os aspectos entonacionais, temporais e de velocidade da produção vocal dos participantes da fala-em-interação social. Para representá-los em texto, as chamadas “convenções Jefferson de transcrição” se consolidaram como conjunto estável de convenções detalhadas, evidenciando o caráter sequencial dialógico dos encontros transcritos. Loder (2008) discute detalhadamente as convenções; por isso, aqui vamos apenas apresentá-las sinopticamente, na forma do quadro que utilizamos a partir de Atkinson e Heritage (1984: ix-xvi), Ochs, Schegloff e Thompson (1996: 461-465) e das instruções para submissão de artigos ao periódico especializado *Research on Language and Social Interaction*.

12. Para discussão a esse respeito, ver Schegloff (1999: 566) e Garcez (2002: 88).

.	(ponto final)	entonação descendente
?	(ponto de interrogação)	entonação ascendente
,	(vírgula)	entonação de continuidade
↑	(seta para cima)	mais agudo
↓	(seta para baixo)	mais grave
<b>palav-</b>	(hífen)	marca de corte abrupto
<b>pala::vra</b>	(dois pontos)	prolongamento do som (maior duração)
<b>pal<u>av</u>ra</b>	(sublinhado)	sílaba ou palavra enfatizada
<b>PALAVRA</b>	(maiúsculas)	intensidade maior (“volume” alto)
<b>°palavra°</b>	(sinais de graus)	intensidade menor (“volume” baixo)
<b>&gt;palavra&lt;</b>	(sinais de maior do que e menor do que)	fala acelerada
<b>&lt;palavra&gt;</b>	(sinais de menor do que e maior do que)	fala desacelerada
<b>hh</b>	(série de h's)	aspiração ou riso
<b>.h</b>	(h's precedidos de ponto)	inspiração audível
<b>=</b>	(sinais de igual)	elocuições contíguas, sem intervalo
<b>[ ]</b>	(colchetes)	início e fim de falas simultâneas/sobrepostas
<b>(2,4)</b>	(números entre parênteses)	medida de silêncio (em segundos e décimos de segundos)
<b>(.)</b>	(ponto entre parênteses)	micropausa de até 2/10 de segundo
<b>( )</b>	(parênteses vazios)	fala que não pôde ser transcrita
<b>(palavra)</b>	(segmento de fala entre parênteses)	transcrição duvidosa
<b>((olha para baixo))</b>	(parênteses duplos)	descrição de atividade não-vocal

A leitura de transcrições seguindo essas convenções não é imediatamente simples, requerendo familiarização semelhante a uma breve alfabetização. Os leitores acostumados com essa notação, contudo, passam a paulatinamente operar uma análise de um excerto pela leitura, sobretudo porque a prosa analítica do relatório de pesquisa se vale das



notações para construir as asserções que o pesquisador faz acerca do que está acontecendo ali. Em seguida, ilustramos tais convenções com um excerto de transcrição de conversa telefônica entre duas jovens, em que combinam detalhes de uma festa que estão organizando.

```

01 Cris: a: eu não me importo de beber cerveja.
02      (.)
03 Cris: tamb[ém.]
04 Lu:   [e:u] também nã::o.
05      (.)
06 Cris: se a gente compra u[ma marca legal é::]
07 Lu:   [°(pega uma cerveja boa)°]
08      (.)
09 Cris: m[m hm]
10 Lu:   [d e ] repente:: em vez de tu compra::::r (.)aquela::°s::°-
11       ā:::- (.) Del- (.) >que- como é que é::<
12       (0,3)
13 Lu:   >tem umas-< Del?co
14       (.)
15 Lu:   >que é ho[rrível]< compra uma um pouquinho melhor↑zinha.
16 Cris: [mm hm]
17       (0,5)
18 Cris: Belco né, >eu acho que [ é < ]
19 Lu:   [Belco.] é.=
20 Cris: =i:[sso]
21 Lu:   [e_s]sa h:::[h ā h ā . D e l c o é : ] bateria.
22 Cris: [ .hhhh >a gente pode compra::::r-<]
23       (.)
24 Cris: mh:-mm hm:::: =
25 Lu:   =.HHH=
26 Cris: =>a gente pode comprar< em:::: (.) t- em lata
27       (0,3)
28 Lu:   é e[m latinha.]
29 Cris: [não sei°::°] a gente vê o que é melhor também.

```

No excerto acima, por exemplo, é crucial a notação exata da ordenação temporal e dos detalhes da produção vocal das participantes para que se observe na transcrição a delicadeza interacional da ação de corrigir o outro (ver Garcez & Loder 2005: 295-296), que envolve mitigações (prolongamento de sons, cortes abruptos, pausas) antes que a palavra considerada pelos participantes como inadequada (*Delco*, linha 12) seja substituída pela considerada adequada (*Belco*, linha 17).

Para a análise das ações dos participantes da fala-em-interação, são imprescindíveis esses “detalhes”, uma vez que os atores sociais claramente se valem de tudo o que está notado para constituir suas ações conjuntas. Para interlocutores pouco familiarizados com as convenções, entretanto, é claro que as notações dificultam a leitura sobremaneira. Mais relevante para essa dificuldade, todavia, é a indiferença do leitor/interlocutor aos princípios teórico-metodológicos que suscitam o detalhamento na notação. Se o interlocutor não compartilha o entendimento de que o uso da linguagem é ação conjunta e que importa privilegiar a

perspectiva dos participantes da ação transcrita, então não verá razão para todo esse empenho, que de certo lhe parecerá despropositado. Se não compartilha o entendimento de que alterações de intensidade e velocidade da fala, disfluências, silêncios etc. são constituintes da ação social pelo uso da linguagem, as notações devem parecer, de fato, tolices de desocupados. Não é incomum que leitores dos nossos relatórios, com quem queremos estabelecer interlocução – colegas da área de Educação, por exemplo –, manifestem publicamente que pularam “as partes com essas coisinhas”, ou que achem que mobilizamos “a artilharia do Exército, da Marinha e da Aeronáutica para abater um pardal”.

Destarte, nossa opção tem sido apresentar transcrições simplificadas, ou mesmo vinhetas narrativas (ver Garcez & Melo 2007: 8-9, 10-11) que relatam o segmento, nas ocasiões em que, para os nossos interlocutores, são potencialmente relevantes tão somente as asserções analíticas finais de nossos estudos, suas consequências para a discussão de certos debates ou mesmo para ilustração de cenários e fenômenos. Essa também tem sido nossa alternativa no que chamamos de interlocução pedagógica, em situações de formação de professores, por exemplo, uma vez que não está em pauta a persuasão desses interlocutores de que a análise é sistemática, rigorosa ou coerente.

Para a análise acadêmica profissional da fala-em-interação social, contudo, em que o interlocutor é um coanalista, esse detalhamento é necessário e, em muitos casos, pode ser insuficiente (conforme exemplificado na subseção d. Multimodalidade, abaixo). Isso muitas vezes fica patente para os pesquisadores durante sessões de análise conjunta, em que o analista expõe sua prosa analítica ao crivo de colegas coanalistas, que vão aprofundando a aproximação situacional do que está acontecendo no segmento sob análise de tal sorte que a transcrição muitas vezes se mostra incompleta ou equivocada, com mudanças muitas vezes decisivas na análise. Assim, a atividade de produzir uma transcrição detalhada é também fruto de treinamento nessas atividades de análise conjunta (ver seção 3. Treinamento, abaixo).

Um auxílio interessante para quem não dispõe de colegas analistas com quem aprender a transcrever são os tutoriais disponíveis na internet, porém exigindo familiaridade com inglês escrito e falado. Em <http://www-staff.lboro.ac.uk/~ssca1/sitemenu.htm>, se encontra

um tutorial de caráter introdutório (não requer conhecimentos prévios aprofundados em *Análise da Conversa Etnometodológica*), desenvolvido por Charles Antaki, com transcrições no modelo Jefferson, acompanhadas dos segmentos em áudio e vídeo, de modo que se pode acompanhar a produção de uma transcrição. Em <http://www.sscnet.ucla.edu/soc/faculty/schegloff/TranscriptionProject/index.html>, são apresentadas as principais convenções Jefferson em trechos de transcrições associados a suas gravações em áudio, que o internauta pode ouvir ao acompanhar a transcrição.

Na elaboração de transcrições pelas convenções Jefferson, a inclusão de elementos multimodais adicionais tem se mostrado cada vez mais analiticamente reveladora. Nesse sentido, os pesquisadores do Grupo de Pesquisa ISE têm adotado procedimentos para criar transcrições multimodais e experimentado recursos auxiliares às transcrições, para representação espacial e temporal dos cenários interacionais, conforme apresentamos a seguir.

### **b. Multimodalidade**

A fala-em-interação é um fenômeno eminentemente multimodal, em que distintas modalidades de natureza vocal-auditiva (fala, cadência rítmica, entonação, respiração) e espaço-visual (gestos, expressões faciais, objetos com os quais os participantes estão lidando, posicionamento espacial, aspectos cinéticos – movimentos próximos e espaciais) são utilizadas para a coconstrução de cursos de ação (Stivers & Sidnell 2005). Entendemos que todos esses aspectos de naturezas distintas estão eminentemente inter-relacionados (Kendon 1990). Distanciando-nos de perspectivas que, embora interessadas no uso da linguagem, ficam circunscritas à fala e à sinalização verbal, consideramos necessário atentar, na análise de fala-em-interação social (face a face ou por videotecnologia), para o conjunto de elementos multimodais que os participantes têm à disposição e de que claramente lançam mão nas interações. Para tanto, vimos enfrentando o desafio contemporâneo de desenvolver convenções que permitam transcrever em alguma medida essa realidade<sup>13</sup>.

---

13. Os procedimentos de transcrição multimodal dos estudos do Grupo de Pesquisa ISE foram inspirados em diversos trabalhos de Charles Goodwin (ver, por exemplo, Goodwin 2003a; 2003b).

A videogravação de um evento interacional possibilita atualmente ao analista a observação – tanto em movimento quanto em quadros parados até cada centésimo de segundo – dos diversos elementos multimodais envolvidos nas complexas ecologias sociais. Na última década, os recursos se tornaram acessíveis em programas de computador de uso comum e em nível de complexidade técnica ao alcance das novas gerações de estudantes universitários. Assim, onde antes fazíamos descrições entre parênteses duplos (conforme as convenções Jefferson) para notar um simples gesto, por exemplo, nos últimos anos passamos a incorporar quadros parados (extraídos dos vídeos) às nossas transcrições. As imagens conferem precisão e legibilidade à notação de elementos multimodais, substituindo com vantagem as descrições que antes eram feitas ao longo da transcrição da fala e que, por mais vívidas que fossem, em geral exigiam muita imaginação por parte do leitor.

O acesso a tal riqueza de detalhes, no entanto, requer continência a fim de mantermos o privilégio à perspectiva dos participantes sem sobrecarregar nossos interlocutores, coanalistas, leitores das transcrições<sup>14</sup>. Assim, nossa orientação é de que o analista selecione quadros parados do registro em vídeo que capturem elementos multimodais relevantes para os propósitos analíticos em pauta. No excerto de transcrição multimodal reproduzido a seguir para ilustração, pode-se observar a relevância dos quadros parados selecionados para o entendimento do que os participantes estão fazendo juntos no cenário institucional escolar no momento em que Priscila suspende a atividade de contação de história para focar no controle social da conduta dos demais participantes.

---

14. Em Bulla (2007) e Rosa (2008), podemos observar grandes diferenças entre o número de quadros utilizados nas transcrições multimodais, fruto de discussões no Grupo ISE acerca da sustentabilidade do privilégio à perspectiva dos participantes na seleção de quadros.

## Práticas de pesquisa microetnográfica

15 Priscila: você[s nã:o conse]quem parar de fala:r

16 Luciano: [°l::: é::°]

17 Paulo: eu-

18 (0,2)

19 Paulo: eu [tô tent-]

20 Jane: [a ou:tra] turma [tu consegui:u,]

21 Paulo: [tando ouvi:] só que eu-

22 Luciano: [c-l::]

23 (0,2)

24 Paulo: [e prestar] atençã:o só que eu-

25 Luciano: [c-l:::] ((simula esfaquear pescoço))

26 Paulo: -° [nã]o [consi:go]°

27 Valter: [tr:] ((simula esfaquear Luciano))

28 Priscila: [si:m só] que se os [outros] não-

29 Luciano: [c-l:::]

30 Priscila: -pa[rarem de fala:r né:] Lucia::no

31 Breno: [eu també:m ( )]

32 (Marina): eu també:m né:,

33 Luciano: [°s:i::m°]-

34 (Paulo): [hm hm hm::]

35 (1,8)

36 Priscila: então vou tenta:r ter[mina:r] de contar a-

37 Luciano: [ô ô]

38 Priscila: -histó:ria para você:s::

39 (0,3)

(transcrição revisada a partir da que apareceu em Rosa 2008: 76-77)

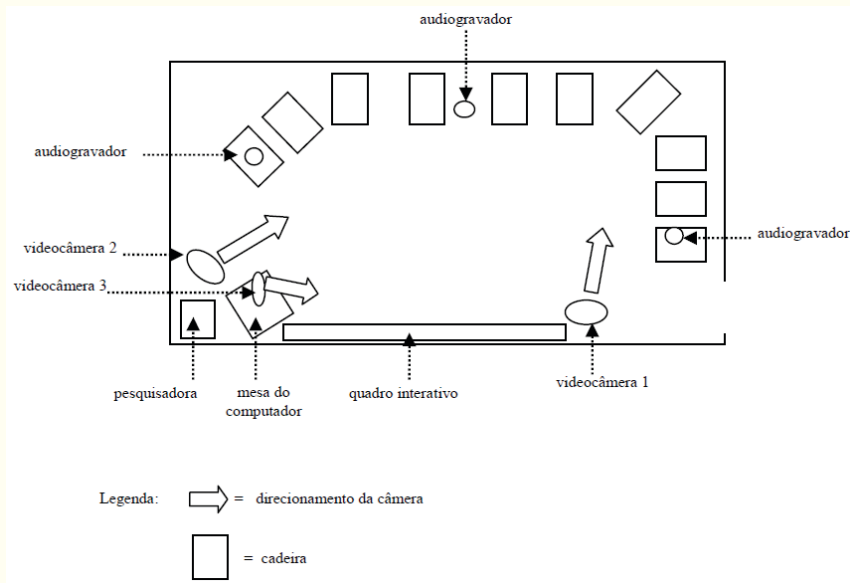
O entendimento das ações de Luciano na sequência interacional fica bastante facilitado pela combinação entre a transcrição vocal e não vocal. Por exemplo, as indicações de direcionamentos de olhar entre Luciano e Priscila (quadros 2, 4, 5 e 6) são relevantes para sabermos que, embora sejam muitos os participantes antes endereçados, neste momento, o foco de atenção interacional está claramente entre esses dois. Serve também para observarmos como um participante ratifica o outro pela fala (linha 33) combinada com o olhar (quadros 5 e 6). Esse aspecto é relevante na análise de Rosa (2008), que discute como Priscila alterna entre as atividades de contação de história, em inglês, e de controle da conduta dos participantes para que se mantenham atentos, em português.

Como se observa no excerto acima, as transcrições multimodais que produzimos no Grupo ISE<sup>15</sup> partem de transcrições dos elementos vocais-auditivos, feitas segundo as convenções Jefferson, e a elas agrega-se uma seleção de quadros parados obtidos da videogravação em meio digital. Portanto, seguimos os procedimentos de geração, segmentação e transcrição Jefferson apresentados acima e, posteriormente, inserimos quadros com imagens realizando as seguintes etapas: (1) selecionar os quadros (tendo claras as justificativas para cada seleção) e transformá-los em figuras (mediante uso de *softwares* de edição de imagens); (2) associar cada quadro à transcrição da fala, no exato momento em que fala e quadro coocorrem no registro audiovisual; (3) numerar os quadros em ordem crescente no excerto; (4) na figura, marcar setas que indiquem direcionamento de olhar ou movimentos corporais relevantes para as ações em curso; e (5) destacar com círculos os gestos ou objetos que sustentem o privilégio à perspectiva dos participantes e estejam relacionados ao fenômeno em foco na pesquisa. Também, no excerto acima, observa-se o uso do encobrimento dos rostos dos participantes (pela utilização da ferramenta *blur* ou “mascarar” em partes dos quadros), prática que vimos adotando com o objetivo de evitar a identificação visual dos participantes das interações sob análise (ver observação sobre consentimento na seção 2. Geração de Dados acima).

Outro procedimento de atenção à multimodalidade espaço-visual é a representação do posicionamento espacial dos participantes em determinado evento interacional mediante construção de um mapa do cenário (Bulla 2007; Schulz 2007; Abeledo 2008). Podem ser identificados, assim, os objetos que compõem o cenário, a localização dos participantes, o posicionamento da(s) câmera(s) e dos demais instrumentos de gravação etc., como na ilustração abaixo:

---

15. Alguns trabalhos recentes incluindo análise de transcrições multimodais são Bulla (2007), Rosa (2008), Salimen (2009), Ramos (2010), Lange (2010), Garcez & Salimen (2011) e Baumvol (2011).

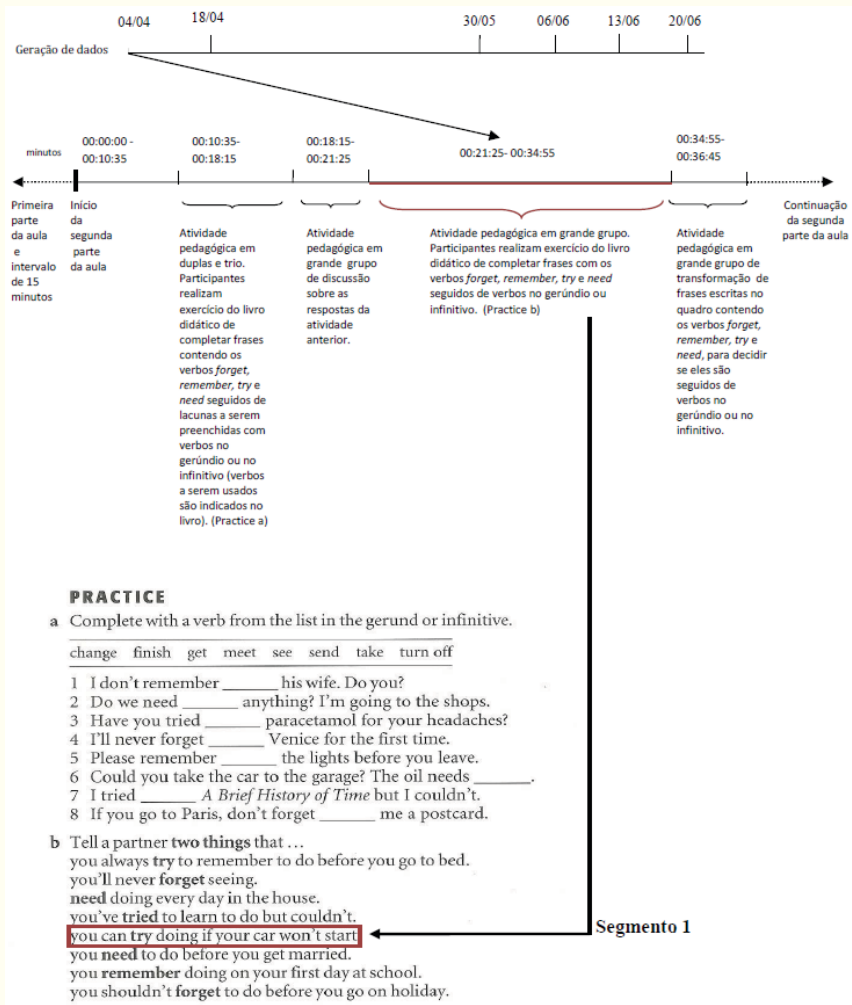


(Baumvol 2011: 49)

Em caráter complementar à transcrição Jefferson acrescida de quadros parados, vimos utilizando, como procedimento recente de atenção ao plano temporal, linhas do tempo que situam excerto ou segmento transcrito na duração do evento e no conjunto de eventos registrados em um dado cenário ou estudo (Almeida 2009; Ramos 2010; Baumvol 2011)<sup>16</sup>. O objetivo é permitir ao leitor localizar-se temporalmente no universo situacional em que foram geradas as gravações, bem como situar, na totalidade das gravações, determinado segmento a ser analisado. No exemplo a seguir, o leitor tem acesso à sequência de atividades em que determinado segmento interacional se insere, com a identificação dos dias de gravação (04/04, 18/04 etc.), das atividades desenvolvidas em um desses dias (04/04), do que veio antes e depois do segmento focal (Segmento 1), e do objeto com o qual os participantes lidam durante o segmento focal (página do livro didático de inglês, com destaque ao enunciado referido no segmento de interação analisado). As linhas do tempo, assim como o mapa do cenário (apresentado acima), não substituem transcrições Jefferson acrescidas

16. Transcrições inspiradas em Erickson & Shultz (1982), Kendon (1990) e Garcez (1996).

de elementos multimodais espaço-visuais, mas são complementos que permitem aprofundar as análises que tomam por base as transcrições de fala-em-interação.



(Baumvol 2011: 63)

Para a construção de transcrições multimodais, como as apresentadas aqui, e a análise de dados, o pesquisador pode se valer do apoio de diferentes *softwares*. Para a construção de transcrições da audiogravação, com atenção a pausas, entonação, cadência rítmica etc., utilizamos o *software* Audacity (<http://audacity.sourceforge.net/?lang=pt>), que é



gratuito. Contudo, o *software* pago SoundForge (<http://www.sonycreativesoftware.com/soundforagesoftware>) possibilita o tratamento de dados em áudio e vídeo ao mesmo tempo, além da captura de quadros parados. O *software* gratuito Transana (<http://www.transana.org>) possibilita o trabalho com o vídeo e a transcrição na mesma tela, a vinculação de ambos (empreitada deveras trabalhosa) e a construção de bancos de dados com coleções de fenômenos interacionais. O Elan (<http://www.lat-mpi.eu/tools/elan>), também gratuito e com estrutura para a criação de bancos de dados, permite a criação de transcrições em que cada elemento multimodal pode ser desmembrado em uma linha diferente e vinculado ao vídeo cronologicamente (para exemplo de transcrição construída no Elan, ver Bulla 2007: 95). Por fim, o *software* CLAN (<http://talkbank.org/CABank/>), utilizado por pesquisadores na área de Aquisição da Linguagem, tem um “modo Análise da Conversa” e, assim, é uma alternativa gratuita de suporte eletrônico para transcrição e análise de dados de fala-em-interação em áudio ou vídeo que adotem as convenções Jefferson e outros procedimentos referidos aqui.

## 5. Treinamento

Todos os procedimentos de que tratamos acima envolvem treinamento, isto é, exercício. Do nosso ponto de vista, isto se faz em parte pelo enfrentamento individual das tarefas, pela sua execução e revisão. Em larga medida, contudo, o treinamento se desenvolve não apenas ao longo do trabalho solitário do pesquisador de revisitar suas gravações e transcrições e aprimorá-las a partir de um segundo olhar, mas também mediante refinamentos alcançados pelas exigências de articulação em sessões de análise conjunta, em que o analista submete aspectos diversos de seu trabalho a seus pares. A atenção simultânea de múltiplos olhares analíticos a um mesmo segmento produz alcance imediato, por exemplo, a ângulos de visão e audição distintos dos do analista que produziu a transcrição, ângulos esses que podem imediatamente ser verificados e submetidos à deliberação em conjunto com outros investigadores. Pontos de chegada aos dados que partem de diferentes interesses, muitas vezes mais periféricos e até bem distintos dos interesses do pesquisador, também proporcionam oportunidades instantâneas de rever *de novo* dados que poderiam lhe parecer esgo-

tados. Assim, consideramos que a análise de dados interacionais deve contar com momentos de análise necessariamente coletiva.

Embora ainda incomuns no Brasil, sessões de análise coletiva de dados de fala-em-interação são comuns em eventos acadêmicos que reúnem microetnógrafos ou analistas da conversa. Tais sessões parecem imprescindíveis para o treinamento de pesquisadores capacitados a executar os procedimentos de geração, segmentação e transcrição de registros audiovisuais que descrevemos aqui. De fato, no Grupo de Pesquisa ISE, fazemos reuniões regulares para análise conjunta de dados desde pelo menos 2001. Ao longo desse período, organizamos diversas oficinas para compartilhar a resolução de problemas e a sistematização de procedimentos que, em diferentes momentos, permitiram a consolidação e o registro de nossas práticas conforme relatamos aqui.

Elevando o que discutimos aqui a um novo patamar de complexidade, a tradução de transcrições para exame em sessões de análise conjunta com pesquisadores que não compartilham a língua utilizada nos dados e apresentações de trabalho ou publicações (ver Dornelles & Garcez 2001) em outra língua tornam patente o que em estágios anteriores a esse pode ficar opaco: transcrições são produtos analíticos plenos.

## 6. Trajetórias em curso

Os procedimentos que apresentamos aqui resultam do trabalho conjunto de um grupo de dezenas de pesquisadores que, ao longo da última década, realizaram repetidas análises, que foram discutidas conjuntamente, revisadas e relatadas. Reunidos periodicamente, buscamos soluções para problemas imediatos de ordem eminentemente executiva na geração, segmentação, transcrição e análise de dados audiovisuais de fala-em-interação social, mas também procuramos deslindar questões de natureza teórico-metodológica. Mais que tudo, porém, os procedimentos apresentados resultam da busca por se aproximar, tanto quanto possível, da perspectiva dos participantes de situações e eventos situados a fim de descrever suas práticas e ações e, assim, compreender o que está acontecendo nesses cenários.

As análises de fala-em-interação produzidas mediante esses procedimentos apresentados acima resultaram em um conjunto coeso de relatórios de pesquisa, alguns dos quais citados ao longo deste artigo. Esses trabalhos flagram trajetórias em curso, que só podem ser efetivamente apreciadas pela leitura de trabalhos analíticos específicos, que esperamos que o leitor consulte. Entretanto, de modo sumário, pode-se dizer que se pautaram pela necessidade de compreender os fenômenos interacionais, primeiramente na matriz da socialidade, que é a conversa cotidiana (Garcez 2008). Possibilitaram uma visão aprofundada de fenômenos da organização elementar da fala-em-interação com base em dados de conversa em português brasileiro (Dornelles & Garcez 2001; Garcez & Loder 2005; Loder & Jung 2008), bem como de práticas e ações interacionais em cenários institucionais, sobretudo escolares (Loder & Jung 2009; Fortes 2009).

Com interesse particular de atender às demandas de Linguística Aplicada, produzimos, assim, um conjunto vigoroso de trabalhos inter-relacionados que examinam ângulos diversos da reprodução e construção conjunta de conhecimento em cenários e eventos escolares (Garcez 2006), especialmente em salas de aula e aulas de língua. Enfocando temas como gênero, letramento e escolarização (Almeida 2009; Jung 2009; Abeledo 2010), participação (Schulz 2007), atividades pedagógicas colaborativas (Bulla 2007) e a própria noção de aprendizagem na interação – ao flagrar e descrever momentos de “fazer aprender” e “fazer ensinar” (Freitas 2006; Abeledo 2008; Rosa 2009; Salimen 2009; Lange 2010; Frank 2010; Garcez & Salimen 2011; Garcez, Frank & Kanitz 2012), – nossas análises possibilitam a reflexão sobre (a) as ações dos participantes da fala-em-interação de sala de aula para fazerem “ser professor, aluno ou outras categorias”; (b) as relações entre tarefas pedagógicas e atividades pedagógicas; e (c) os modos como cada participante age para o alcance de objetivos educacionais, com destaque para o que age na capacidade de “professor”.

Na base de todas essas reflexões, no entanto, estão a geração, a transcrição e a segmentação de dados de fala-em-interação, tomadas não como procedimentos preliminares à análise, mas como fundamentos teórico-metodológicos centrais a qualquer empreitada em pesquisa microetnográfica.

Recebido em setembro de 2013  
Aprovado em março de 2014  
E-mails: pmgarcez@pq.cnpq.br  
gsbulla@gmail.com  
leticialoder@terra.com.br

## Referências bibliográficas

- ABELED, Maria de la O L. 2008. *Uma compreensão etnometodológica da aprendizagem de língua estrangeira na fala-em-interação de sala de aula*. Tese de doutorado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- \_\_\_\_\_. 2010. “¿Te parece que ella se afeitada?”: categorias identitárias de gênero e aprendizagem de vocabulário na fala-em-interação de sala de aula de língua estrangeira. *Calidoscópico* 8/2: 103-117.
- ALMEIDA, Alexandre N. 2009. *A construção de masculinidades na fala-em-interação em cenários escolares*. Tese de doutorado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- ATKINSON, John M. & John Heritage. 1984. *Structures of social action*. Cambridge: Cambridge University Press.
- BAUMVOL, Laura K. 2011. *O trabalho interacional para a reconfiguração de atividades pedagógicas instrucionais na sala de aula de inglês como língua adicional*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- BUCHOLTZ, Mary. 2000. The politics of transcription. *Journal of Pragmatics* 32/10: 1439-1465.
- BULLA, Gabriela S. 2007. *A realização de atividades pedagógicas colaborativas em sala de aula de português como língua estrangeira*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- CONCEIÇÃO, Luciana E. 2008. *Estruturas de participação e construção conjunta de conhecimento na fala-em-interação de sala de aula de Língua Inglesa em uma escola pública municipal de Porto Alegre*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- CONCEIÇÃO, Luciana E. & Pedro de M. Garcez. 2005. O revozeamento no discurso da escola pública cidadã. *Intercâmbio* (PUC-SP) 14: 1-10. (disponível em <http://revistas.pucsp.br/index.php/intercambio/article/viewFile/3941/2591>)

- DORNELLES, Clara & Pedro de M. Garcez. 2001. Making sense of nonsense: Fabrication, ambiguity, error and clarification in the organization of experience in ordinary conversation. *Journal of Pragmatics* 33/11: 1707-1730.
- DURANTI, Alessandro. 1997. *Linguistic anthropology*. Cambridge: Cambridge University Press.
- Erickson, Frederick. 1990. Qualitative methods. In: Robert L. Linn & Frederick Erickson. Orgs. *Quantitative methods; Qualitative Methods. Vol.2*. New York: Macmillan.
- \_\_\_\_\_. 2004. *Talk and social theory: Ecologies of speaking and listening in everyday life*. Cambridge: Polity Press.
- ERICKSON, Frederick & Jeffrey J. Shultz. 1982. *The counselor as gatekeeper: Social interaction in interviews*. New York: Academic Press.
- \_\_\_\_\_. 2002. 'O quando' de um contexto: Questões e métodos na análise da competência social (Pedro de M. Garcez & Clarissa Surek-Clark. Trans.). In: Branca T. Ribeiro & Pedro de M. Garcez. Orgs. *Sociolinguística Interacional. 2a Ed.* São Paulo: Loyola. [Tradução de Erickson, Frederick & Jeffrey J. Shultz. 1981. 'When' is a context: Some issues and methods in the analysis of social competence. In: Judith L. Green & Cynthia Wallat. Orgs.. *Ethnography and language in educational settings*. Norwood, NJ: Ablex.]
- FERREIRA, Aurélio B. de H. 1999. *Novo Aurélio século XXI: o dicionário da língua portuguesa*. 3ª ed. Rio de Janeiro: Nova Fronteira.
- FORTES, Melissa S. 2009. *Uma compreensão etnometodológica do trabalho de fazer ser membro na fala-em-interação de entrevista de proficiência de português como língua adicional*. Tese de doutorado. Porto Alegre: Programa de Pós-Graduação em Letras, Universidade Federal do Rio Grande do Sul.
- FREITAS, Ana Luiza. P. 2006. *With a little help from my friend: um estudo sobre o reparo levado a cabo pelo terceiro na sala de aula de língua estrangeira*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- GAGO, Paulo C. 2004. Questões de transcrição em Análise da Conversa. *Veredas* 6/2: 89-113.
- GARCEZ, Pedro de M. 1993. Point-making styles in cross-cultural business negotiation: A microethnographic study. *English for Specific Purposes* 12/2: 103-120.
- \_\_\_\_\_. 1996. *Brazilian manufacturers and U.S. importers doing business: The co-construction of arguing sequences in negotiation*. Tese de doutorado. Filadélfia: Universidade da Pensilvânia.

- \_\_\_\_\_. 2002. Transcrição como teoria: a identificação dos falantes como atividade analítica plena. In: Luiz Paulo da Moita-Lopes & Liliana C. Bastos. Orgs. *Identidades: recortes inter- e multidisciplinares*. Campinas, SP: Mercado de Letras.
- \_\_\_\_\_. 2004. Arguing sequences in cross-cultural business negotiation talk. In: Carlos A. M. Gouveia, Carmina Silvestre & Luísa Azuaga. Orgs.. *Discourse, communication and the enterprise: Linguistic perspectives*. Lisboa: Centro de Estudos Anglísticos da Universidade de Lisboa.
- \_\_\_\_\_. 2006. A organização da fala-em-interação na sala de aula: controle social, reprodução de conhecimento, construção conjunta de conhecimento. *Calidoscópico* 4/1: 66-80. (disponível em [http://www.unisinos.br/publicacoes\\_cientificas/images/stories/Publicacoes/Calid\\_v4n1/art07\\_garcez.pdf](http://www.unisinos.br/publicacoes_cientificas/images/stories/Publicacoes/Calid_v4n1/art07_garcez.pdf))
- \_\_\_\_\_. 2008. Microethnography in the classroom. In: Kendall A. King & Nancy H. Hornberger. Orgs.. *The encyclopedia of language and education. Vol.10: Research methods in language and education*. Berlin: Springer.
- GARCEZ, Pedro de M. & Leticia L. Loder. 2005. Reparo iniciado e levado a cabo pelo outro na conversa cotidiana em português do Brasil. *DELTA* 21/2: 279-312.
- GARCEZ, Pedro de M. & Paloma S. Melo. 2007. Construindo o melhor momento para tomar o turno na fala-em-interação de sala de aula na escola pública cidadã de Porto Alegre. *Polifonia* 13: 1-21. (disponível em <http://cpd1.ufmt.br/meel/arquivos/artigos/138.pdf>)
- GARCEZ, Pedro de M. & Paola G. Salimen. 2011. Pedir e oferecer ajuda para “fazer aprender” em atividades pedagógicas de encenação na fala-em-interação de sala de aula de inglês como língua adicional. In: Ana Maria F. Barcelos. Org. *Linguística Aplicada: Reflexões sobre ensino e aprendizagem de Língua Materna e Língua Estrangeira*. Campinas, SP: Pontes Editores.
- GARCEZ, Pedro de M., Ingrid Frank & Andréia Kanitz. 2012. Interação social e etnografia: sistematização do conceito de construção conjunta de conhecimento na fala-em-interação de sala de aula. *Calidoscópico* (Unisinos) 10/2: 211-224. (disponível em <http://www.unisinos.br/revistas/index.php/calidoscopio/article/view/cld.2012.102.08/986>).
- GOLDIM, José Roberto. 2006. Consentimento e informação: a importância da qualidade do texto utilizado. *Revista do HCPA & Faculdade de Medicina da Universidade Federal do Rio Grande do Sul* 26: 117-122.

- GOLDIM, José Roberto, Joaquim Clotet & Jorge P. Ribeiro. 2007. Adequacy of informed consent in research carried out in Brazil. *Eubios Journal of Asian and International Bioethics* 17: 177-180.
- GOODWIN, Charles. 2003a. Pointing as situated practice. In: Sotaro Kita. Org. *Pointing: Where language, culture and cognition meet*. Mahwah, NJ: Lawrence Erlbaum.
- \_\_\_\_\_. 2003b. The semiotic body in its environment. In: Justine Coupland & Richard Gwyn. Orgs. *Discourses of the body*. New York: Palgrave/Macmillan.
- HAVE, Paul ten. 1999. *Doing conversation analysis*. Londres: Sage.
- HERITAGE, John & Sue Sefi. 1992. Dilemmas of advice: aspects of the delivery and reception of advice in interactions between health visitors and first-time mothers. In: Paul Drew & John Heritage. Orgs. *Talk at work*. Cambridge: Cambridge University Press.
- JEFFERSON, Gail. 1983. Issues in the transcription of naturally occurring talk: caricature versus capturing pronunciation particulars. *Tilburg papers in language and literature* 34: 1-12.
- \_\_\_\_\_. 1996. A case of transcriptional stereotyping. *Journal of Pragmatics* 26: 159-170.
- KENDON, Adam. 1990. *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
- LANGE, Catilcia P. 2010. *Formulação e ensino-aprendizagem na fala-em-interação de sala de aula de Inglês como língua adicional na Educação de Jovens e Adultos*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- LODER, Letícia L. 2008. O modelo Jefferson de transcrição: convenções e debates. In: Letícia L. Loder & Neiva M. Jung. Orgs. *Fala-em-interação social: introdução à Análise da Conversa Etnometodológica*. Campinas, SP: Mercado de Letras.
- LODER, Letícia L. & Neiva M. Jung. Orgs. 2008. *Fala-em-interação social: uma introdução à Análise da Conversa Etnometodológica*. Campinas, SP: Mercado de Letras.
- \_\_\_\_\_. 2009. *Análises de fala-em-interação institucional: a perspectiva da Análise da Conversa Etnometodológica*. Campinas, SP: Mercado de Letras.
- MOERMAN, Michael. 1988. *Talking culture: Ethnography and conversation analysis*. Filadélfia: University of Pennsylvania Press.
- OCHS, Elinor. 1979. Transcription as theory. In: Elinor Ochs & Bambi B. Schieffelin. Orgs. *Developmental pragmatics*. New York: Academic Press. (disponível em <http://www.sscnet.ucla.edu/anthro/faculty/ochs/articles/ochs1979.pdf>)



- OCHS, Elinor, Emanuel A. Schegloff & Sandra A. Thompson. 1996. *Interaction and grammar*. Cambridge: Cambridge University Press.
- PAOLETTI, Isabella & Giolo Fele. 2004. Order and disorder in the classroom. *Pragmatics* 14/1: 69-85.
- RAMOS, Ingrid F. 2010. *Constituição e superação de momentos desconfortáveis em sequências de convites à participação: a construção do engajamento na fala-em-interação de sala de aula*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- RAMPTON, Ben. 2006. *Language in late modernity: Interaction in an urban school*. Cambridge: Cambridge University Press.
- ROSA, Aline P. 2008. “É o professor quem diz quando se fala?” *A tomada de turnos de fala em atividades diferentes em uma turma de 1a. série em educação bilíngue*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- SACKS, Harvey, Emanuel A. Schegloff & Gail Jefferson. 2003\*. Sistemática elementar para a organização da tomada de turnos para a conversa. *Veredas* 7(1-2), 9-73. [Tradução de Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematic for the organization of turn-taking for conversation. *Language* 50: 696-735. \*Efetivamente publicada em 2005.]
- SALIMEN, Paola G. 2009. *A atividade pedagógica de encenar em grupos na sala de aula de língua estrangeira: pedidos de ajuda, ofertas de ajuda e aprendizagem*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- SCHEGLOFF, Emanuel A. 1999. Schegloff's text as Billig's data: a critical reply. *Discourse and Society* 10/4: 558-572.
- SCHNACK, Cristiane M., Thaís D. Pisoni & Ana Cristina Ostermann. 2005. Transcrição de fala: do evento real à representação escrita. *Entrelinhas* 2/2.
- SCHULZ, Lia. 2007. *A construção da participação na fala-em-interação de sala de aula: um estudo microetnográfico sobre a participação em uma escola municipal de Porto Alegre*. Dissertação de mestrado. Porto Alegre: Programa de Pós-Graduação em Letras, UFRGS.
- STIVERS, Tanya & Jack Sidnell. 2005. Introduction: Multimodal interaction. *Semiotica* 156/1: 1-20.
- ZILLES, Ana Maria S. & Carlos Alberto Faraco. 2002. Considerações sobre o discurso reportado em corpus de língua oral. In: Paulino Vandresen. Org. *Variação e mudança no português falado na Região Sul*. Pelotas, RS: Educat.