

## A inspiração etnográfica e o papel da multimodalidade na análise de uma sequência de fala-em-interação institucional

Ethnographic inspiration and the role of multimodality in the analysis of an institutional talk-in-interaction sequence

Roberto Perobelli<sup>1</sup>  
Luana Santos Lemos<sup>2</sup>

**Resumo:** A partir de uma sequência de fala-em-interação de um professor conversando com a mãe de uma estudante em um atendimento numa escola pública, é possível notar como alguns aspectos da interação social que estão além das expressões formulaicas (e.g. “ela vai sair dessa”) ganham contornos de preocupação real e legítima a partir de outras ações corporificadas simultaneamente implementadas no momento da fala. Essa constatação analítica se torna possível em razão dos recursos que se utilizam, uma vez que ajudam a compor o quadro de observação participante do campo investigado, tal como é esperado de pesquisadores que fazem um trabalho etnográfico. Com o objetivo de mostrar como a Análise da Conversa pode ajudar a compreender as ações humanas em uma situação social, sobretudo se for levada em consideração a mobilização de recursos de naturezas semióticas distintas, este artigo conclui que os avanços tecnológicos e econômicos que permitiram às câmeras e aos microfones se tornarem cada vez mais portáteis e acessíveis, novas possibilidades de investigação etnográfica se consolidam com o uso desses aparelhos. Além disso, discute-se uma nova perspectiva de trabalho que suplanta a linearidade do tempo.

**Palavras-chave:** multimodalidade; temporalidade; etnografia; fala-em-interação institucional; Análise da Conversa.

**Abstract:** A talk-in-interaction sequence of a teacher talking to a student’s mother in a public school is important to notice how some aspects of social interaction that go beyond formulaic expressions (e.g., “she will be fine”) gain contours of real and legitimate concern within other embodied actions simultaneously implemented at talk. The resources used in this analysis help to compose the participant observation framework that characterize the investigated field. Since Conversation Analysis can help to understand human actions in a social situation, especially if the mobilization of resources of different semiotic natures is considered, this article concludes that the technological and economic advances that allowed cameras and microphones become increasingly portable and accessible, new possibilities for ethnographic investigation are consolidated with the use of these devices. In addition, a new perspective of work that overcomes the linearity of time is discussed.

**Keywords:** multimodality; temporality; ethnography; institutional talk-in-interaction; Conversation Analysis.

---

1 Universidade Federal do Espírito Santo, Departamento de Línguas e Letras, Centro de Ciências Humanas e Naturais, Programa de Pós-Graduação em Linguística, Vitória, ES, Brasil. Endereço eletrônico: [robertoperobelli@gmail.com](mailto:robertoperobelli@gmail.com).

2 Universidade Federal do Espírito Santo, Centro de Ciências Humanas e Naturais, Programa de Pós-Graduação em Linguística, Vitória, ES, Brasil. Endereço eletrônico: [luagora@yahoo.com.br](mailto:luagora@yahoo.com.br).

### **Introdução: a inspiração etnográfica da análise multimodal de interações**

Às vezes, um atendimento institucional vai além de seus limites. Uma conversa com o professor da filha, por exemplo, pode revelar uma série de questões que podem ir além da relação institucional que a escola estabelece. Para além de saber como estão as notas da filha na escola, há um grande contingente de traumas e inibições que não se diagnostica nas notas trimestrais. Além disso, a conversa de uma responsável com o professor da filha pode também deixar transparecer que há na própria responsável algumas angústias que precisariam ser ouvidas, antes mesmo de a situação escolar da filha ser concretamente discutida. Nesse sentido, um olhar sensível para as trajetórias de ação que os participantes implementam em suas interações pode ser uma boa oportunidade de se pensar sobre a importância da proposição de mecanismos de escuta ativa aos/as profissionais da educação, não para que substituam profissionais de outros campos (como psicólogos/as, por exemplo), mas para que saibam como agir quando são alçados ao posto de alguém em quem se confia muito.

Tendo em vista que o dossiê temático especial em que este artigo se insere tem o objetivo de agregar pesquisas que coloquem em perspectiva a inspiração etnográfica presente em estudos da linguagem em seus diferentes aspectos, pretendemos mostrar como a Análise da Conversa pode ajudar a compreender as ações humanas em uma situação social, sobretudo se for levada em consideração a mobilização de recursos de naturezas semióticas distintas. A observação dos gestos e dos contornos prosódicos dos participantes pode ser um elemento importante quando esses próprios participantes precisam tomar decisões sobre como devem agir diante de uma situação contextual que ultrapassa a agenda de trabalho projetada para um encontro conhecido como *plantão pedagógico*, cujo objetivo principal é a entrega, pelo professor de cada componente curricular, das notas trimestrais aos responsáveis dos estudantes. Com dados gerados no contexto de fala-em-interação desses atendimentos pedagógicos<sup>3</sup> às famílias de estudantes de uma escola pública, na região metropolitana de Vitória (ES), a análise busca investigar como certas ações são interacionalmente propostas, negociadas e redirecionadas, considerando não só o que é expresso verbalmente, mas também o que é realizado corporificadamente.

---

3 Em alinhamento ao que apresenta Couto (2021), optamos por denominar os atendimentos pedagógicos dessa maneira, pois eles são assim identificados pelos próprios participantes da pesquisa. Sendo assim, pretendemos dar destaque para como esses encontros são reconhecidos pelos indivíduos em suas próprias práticas. Essa denominação é reconhecida pela escola como uma prática realizada pelos indivíduos que fazem parte da comunidade escolar para tratar de assuntos relacionados ao universo educacional como prestar apoio pedagógico aos pais, mediar conflitos, dar aconselhamento aos estudantes, tratar de assuntos administrativos, entre outros.

Ao lançarmos nosso olhar para os atendimentos pedagógicos como encontros sociais, percebemos que essas interações possuem sistemas de troca de fala que são institucionalmente específicos, que ocorrem entre os indivíduos responsáveis pelos estudantes da instituição de ensino e membros da unidade escolar. Tomando de empréstimo a visão da Sociolinguística Interacional, um encontro social pode ser definido, de acordo com Garcez e Ostermann (2002, p. 260), como “um empreendimento em orientação conjunta, ou comprometimentos de face, em que duas ou mais pessoas em uma situação social ratificam conjuntamente uma(s) à(s) outra(s) como cossustentadoras autorizadas de um único foco de atenção cognitiva e visual, ainda que móvel.” Nesses ambientes, os participantes, ou pelo menos um deles, demonstra(m) que a cena em que se está atuando difere da conversa cotidiana e que o sistema de tomada de turno<sup>4</sup> é fortemente restrito por procedimentos demarcados no espaço escolar.

Uma vez que essas peculiaridades da fala-em-interação institucional são observadas, uma descrição desses cenários<sup>5</sup> carrega consigo a necessidade de que essa observação não seja feita apenas sobre as palavras que se dizem, mas principalmente sobre as ações que são realizadas com as palavras que se dizem. Diferentemente da linha adotada por Austin (1990 [1962]), não procuramos aqui catalogar ou classificar atos que são performados quando certas elocuições são ditas. Nosso intuito é, portanto, priorizar as ações em curso no aqui e agora da interação, para além do que se diz, e assim, procurar descrever essas ações da maneira mais próxima possível da perspectiva de quem vive a situação interacional. Dessa forma, consideramos importante mencionar que é em nome dessa descrição que somos levados a fazê-la com o máximo de detalhamento, tendo em vista que a não observação desses detalhes pode comprometer a análise. Isso significa que nosso olhar analítico não privilegia nenhum dos recursos que são levados a cabo pelos participantes de uma interação, porque a ecologia dos sistemas semióticos<sup>6</sup> presentes na interação institucional possui características peculiares que exibem, elas mesmas, as ordens sociais e institucionais de uma interação. Para que uma análise multimodal tenha relevância, portanto, é necessário que ela consiga flagrar os principais

---

4 Tomada de turno aqui é o termo utilizado por Sacks *et al.* (2003 [1974]) para se referir ao sistema de trocas de fala entre participantes de uma determinada interação. Considera-se turno o que comumente também se conhece por *vez de falar*, na mesma medida em que a tomada de turno seria, portanto, a conquista da vez de falar por um participante da interação em curso. Segundo os autores, “[o]bviamente, trata-se de um tipo proeminente de organização social, cujos exemplos estão implicados em uma vasta gama de outras atividades. Nas atividades organizadas socialmente, a presença de ‘turnos’ sugere uma economia, com turnos como algo a que se confere valor – e com meios para alocá-los, que afetam sua distribuição relativa, tal como em sistemas econômicos.” (SACKS *et al.*, 2003 [1974], p. 12)

5 *Cenário* aqui deve ser entendido como a “estrutura social local” (SCHEGLOFF, 1992, p. 111).

6 Para um melhor entendimento da noção de ecologia dos sistemas semióticos, recomendamos a leitura de Goodwin (2010) e Cruz, Ostermann, Andrade e Frezza (2019).

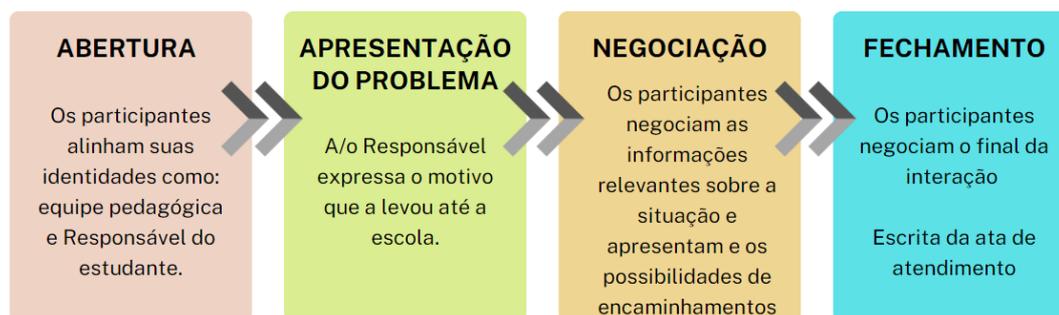
movimentos dos interagentes (vocalizados ou corporificados, indistintamente) e, assim, se constituir como mais um recurso etnográfico voltado para a realização de uma análise qualitativa das ações sociais produzidas por uma determinada comunidade de prática.

### **Informações etnográficas e seu caráter bordeador da análise multimodal**

Nas interações cotidianas, as características estruturais incluem atividades específicas como sua abertura e o fechamento, e o restante da interação é coconstruído pelos participantes sem uma ordem determinada. Entretanto, algumas interações institucionais apresentam uma organização estrutural geral específica, com fases ou atividades que emergem caracteristicamente em uma ordem em particular (fig. 1). Um trabalho importante para as pesquisas que se dedicam ao estudo da fala-em-interação institucional foi o livro editado por Paul Drew e John Heritage (1992), *Talk at Work: interaction in institutional settings*, que apresentam em sua introdução as noções basilares sobre a fala institucional que se produz, principalmente, pelo fato de que a identidade institucional ou profissional dos/as participantes se faz relevante, de alguma forma, para as atividades de trabalho nas quais estejam engajados (DREW; HERITAGE, 1992, p. 4). Para Drew e Heritage (1992), é por meio da fala-em-interação que muitos profissionais atingem vários objetivos práticos e é também por meio da fala-em-interação institucional que se constitui o meio central através do qual as atividades do trabalho diário são conduzidas. Vale lembrar que a institucionalidade da interação não é determinada pelo local em que a conversa ocorre, mas na medida em que os participantes ou profissionais tornam sua identidade profissional de alguma maneira relevante às atividades do trabalho em que estão engajados.

Do ponto de vista de descrição etnográfica e de sua contribuição para os estudos linguísticos, destacamos que os atendimentos pedagógicos, em sua organização macroestrutural, no cenário analisado, não possuem uma estrutura engessada, porém notamos que existe uma organização estrutural global comum no que se refere à sequência de ações institucionalizadas pela gestão da escola para esses atendimentos. Analisando os atendimentos pedagógicos da escola pesquisada, observamos que esses encontros possuem algumas ações rotinizadas pela equipe pedagógica que demarcam a sequencialidade e organização da interação, que balizam a característica da fala-em-interação institucional em uma determinada organização e com algumas estruturais gerais, como é possível ver a seguir:

Figura 1 - Organização macroestrutural dos Atendimentos Pedagógicos às Famílias



Fonte: elaboração própria dos autores

Uma análise aprofundada das sequências conversacionais no contexto desses atendimentos pedagógicos evidencia seu caráter de fala-em-interação institucional, na medida em que descreve um padrão específico de trocas de turnos que difere do padrão formatado por Sacks, Schegloff e Jefferson (2003 [1974]) para a conversa cotidiana. Um dos elementos que todos os sistemas de troca de falas institucionais têm em comum é que eles são adaptados para atender a uma função ou um conjunto de funções em contextos particulares e especializados (WATSON; GASTALDO, 2015 p.131). Na esteira dessa multiplicidade, as interações institucionais são pautadas pela necessidade de realização de uma (ou mais) tarefa(s) específica(s) e, deste modo, tendem a desenvolver formatos organizacionais que visam ao cumprimento dessa tarefa. Del Corona (2009) expõe que vários pesquisadores se dedicaram a analisar especificamente interações institucionais não apenas em termos da organização da tomada de turnos, mas também em termos de outros fenômenos interacionais, contribuindo, dessa forma, para caracterizá-las, na especificidade de cada cenário, em contraponto com as interações cotidianas.

Na escola analisada, os encontros promovidos institucionalmente podem acontecer por solicitação dos responsáveis dos estudantes ou por necessidade da escola em conversar com a família. Neste último, a família (ou os/as responsáveis) é (são) convocada(os/as) por meio de bilhete por escrito ou por contato telefônico. Existem também eventos específicos previstos em calendário escolar, em que, em data previamente agendada, as famílias dos estudantes podem se dirigir à Unidade Escolar para conversar com os professores e com a equipe pedagógica. Não existe uma duração específica para a realização desses atendimentos, pois podem variar bastante dependendo do assunto a ser tratado e seus desdobramentos.

### **Metodologia de geração e análise dos dados: inspiração etnográfica**

O dado a ser analisado neste artigo trata de uma interação entre a responsável por uma estudante (RES) e um professor (PRO), durante um plantão pedagógico, previsto no calendário escolar, que aconteceu em 12 de dezembro de 2019. A dinâmica desses encontros consistia na entrega das notas finais pelos professores e no diálogo entre a família e a escola sobre o aproveitamento escolar dos estudantes. Quando tratamos do espaço da escola, as interações sociais que ali acontecem estão pautadas no reconhecimento de que pelo menos um dos participantes tenha objetivos ou metas-fim que orientam suas ações, na fala-em-interação institucional, voltadas para a aprendizagem dos estudantes. Dessa forma, no contexto dos atendimentos pedagógicos o representante da escola (professor/a, pedagogo/a, coordenador/a, diretora/a) possui alguns objetivos específicos, como, por exemplo, prestar contas de uma situação pedagógica que envolve o estudante, informar a família sobre o rendimento escolar, buscar entendimento sobre alguma dificuldade do estudante junto à família, tratar sobre questões de comportamento, entre os mais diversos assuntos que fazem parte do cotidiano escolar. Neste atendimento, RES procura PRO com o objetivo de receber o boletim com as notas da estudante e conversar sobre como foi o desenvolvimento da estudante ao longo do ano letivo. Encontram-se na sala da esquerda para a direita, na figura 2, o professor (PRO) e a responsável pela estudante (RES).

Figura 2 - Disposição dos participantes na sala



Fonte: elaboração dos próprios autores

A geração de dados desta pesquisa foi realizada em colaboração com outra pesquisadora (COUTO, 2021). Atualmente, esses dados também fazem parte do banco de dados do grupo de

pesquisa Linguagem, Interação e Etnometodologia (GLIE)<sup>7</sup>, do qual a pesquisadora citada e os autores deste trabalho participam. É importante ressaltar que a expressão *coleta de dados*, muito comum em trabalhos etnográficos, foi aqui substituída por *geração de dados*, uma vez que, por se tratar de uma pesquisa qualitativa/interpretativa, refuta-se a ideia de que o pesquisador possa ser completamente neutro durante sua *geração de dados* no mundo social em análise. Essa substituição do termo *coleta* por *geração* foi sugerida por Garcez, Bulla e Loder (2014), para quem:

[...] 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. (GARCEZ, BULLA, LODER, 2014, p. 262).

Foram realizadas gravações de vídeo e áudio dos atendimentos pedagógicos com o objetivo de captar as ações realizadas nesses encontros sociais com o mínimo de interferência possível, isto é, mesmo com todo um aparato técnico (câmeras e microfones), a intenção era buscar registrar o que estava acontecendo ali de uma forma que os/as participantes se sentissem minimamente incomodados e buscassem realizar suas ações como se não estivessem sendo gravados. Para os estudos de base etnometodológica, essa postura é muito importante, porque privilegia uma geração naturalística de dados. Sendo assim, documentamos 16 encontros que contabilizaram aproximadamente seis horas de gravação.

Para a transcrição dos dados, utilizamos o modelo de transcrição desenvolvido por Gail Jefferson (2004) e a perspectiva de transcrição multimodal (MONDADA, 2018). A metodologia de análise precisava se expandir, uma vez que, inicialmente, os dados na área da AC eram gerados apenas em áudio. Com o avanço da tecnologia, na elaboração de transcrições pelas convenções Jefferson, os dados gerados nessas pesquisas passaram a contar com o recurso das gravações em vídeo, dessa forma, a inclusão de elementos multimodais adicionais tem se mostrado cada vez mais analiticamente reveladora. Desse modo, nesse contexto, como será

---

7 Grupo de pesquisa liderado pelo professor Roberto Perobelli, iniciado em 2018, na Universidade Federal do Espírito Santo, com a proposta de integrar graduandos dos cursos de Letras e pós-graduandos do Programa de Pós-Graduação em Linguística da Ufes interessados em desenvolver pesquisas sobre fala-em-interação cotidiana e/ou institucional.

possível observar a seguir, o campo da multimodalidade destacou-se como a solução para dar conta dessas novas questões que foram emergindo.

### **A multimodalidade na fala-em-interação institucional**

Tradicionalmente, os linguistas dão prioridade ao nível verbal da comunicação, porém, em estudos mais recentes, esse cenário tem mudado. Levinson (2012) aponta que pensar em sequências de ações não-verbais pode nos ajudar a entender as verbais, destacando ainda que gestos e vocalização são indissociáveis. No entanto, a perspectiva que se adota neste trabalho é diferente na medida em que não prevê uma hierarquia entre ações verbalizadas e ações corporificadas. Pelo contrário, a interpenetração dessas ações umas nas outras é uma condição de análise que pretende considerar radicalmente como as próprias pessoas no curso de suas ações compreendem e lidam responsivamente na construção de seus turnos de fala. Este trabalho, então, compõe uma sequência de trabalhos de pesquisa e se alinha epistemologicamente ao trabalho de Perobelli (2022, p. 176), que, também por meio da análise multimodal de uma entrevista televisiva, “pretende mostrar em minúcias que o que acontece em uma sequência de pergunta e resposta pode ser muito revelador da organização da fala-em-interação em uma entrevista como um todo”.

Segundo Mondada (2018), os estudos de base interacional que envolvem a multimodalidade, com análise de material filmado, se dão ainda no início dos anos 1970. Segundo a autora, Charles e Marjorie Goodwin e Christian Heath têm papel especial nessa história por terem iniciado a tradição de lançar mão de tais recursos nos estudos sobre a conversa em espaços institucionais. Mondada destaca que

[...] o que é diferente sobre o uso de vídeo pela AC é a atenção cuidadosa e precisa prestada a detalhes organizados temporal e sequencialmente de ações que explicam como os coparticipantes orientam-se sobre a conduta uns dos outros e montam-na de maneiras significativas, momento por momento. Mesmo a fala sendo o foco central da AC, seu objeto principal não é linguagem (ver Sacks, 1984), mas sim a ação, para a qual a linguagem é um recurso importante, embora não o único, nem essencial. Isso tem consequências para a importância dos vídeos na AC, pela forma como são especificamente produzidos, transcritos e analisados. (MONDADA, 2018, p. 86)<sup>8</sup>

---

8 *What is distinctive about CA's use of video is the careful and precise attention paid to temporally and sequentially organized details of actions that account for how coparticipants orient to each other's conduct and assemble it in meaningful ways, moment by moment. Even where talk has been the central focus of CA, its primary object is not language (see Sacks, 1984) but rather action, for which language is an important, but neither the only nor the essential, resource. This has consequences for the importance of videos in CA, for how they are specifically produced, transcribed, and analyzed.* (MONDADA, 2018, p. 86)

Nesse sentido, para além da materialidade verbal, interessam igualmente à Análise da Conversa as condutas não verbais, a que chamamos de multimodalidade ou perspectiva corporificada da interação (STREECK; GOODWIN; LEBARON, 2011). Desse modo, para apoiar o quadro metodológico desta pesquisa, convocamos, principalmente, as categorias analíticas propostas pelos estudos da multimodalidade em interações sociais que se configura a partir da linguagem e movimentação corporal nas ações comunicativas, pelo processo cognitivo, pela cultura, conhecimento de mundo, identidades e relações sociais, espacialidade e temporalidade (MONDADA, 2016a).

Os recursos multimodais são caracterizados por uma série de fatores: (i) sempre se relacionam com a organização da ação, mas não fazem sentido fora dela; (ii) a noção de multimodalidade inclui recursos linguísticos e corporais, tratando-os, em princípio, da mesma forma; (iii) os recursos multimodais não se referem apenas aos recursos convencionais, como gramática e alguns tipos de gesto, mas também para recursos ocasionados situados dependendo das características locais da ecologia da atividade de ambos, habilitando e restringindo o que os participantes tratam como um recurso significativo; (iv) são caracterizados por uma temporalidade específica que combina várias linhas sucessivas e simultâneas de conduta; e (v) são combinados em várias configurações, ou gestalts multimodais, dependendo da atividade, sua ecologia e suas restrições materiais (MONDADA, 2018, p. 86).

Mondada (2016a) defende a investigação multimodal nos estudos de Análise da Conversa pela metodologia que representa na investigação da sequencialidade de atos comunicativos e na representação social. Nessa perspectiva, a multimodalidade é estudada dentro da Análise da Conversa, sendo que algumas análises enfocam principalmente a organização de ambientes específicos, principalmente os institucionais, a fim de compreender contextos espaço-materiais complexos para a ação, incluindo o uso de tecnologias e artefatos.

Mondada (2016b) demarca que as convenções são concebidas para anotar todas as ações corporais possivelmente relevantes, como gesto, olhar, postura corporal, movimentos etc. que acontecem simultaneamente para falar ou durante momentos de ausência de conversa. Segundo a autora, a convenção é baseada em dois princípios:

- a) Caracterização da trajetória temporal: cada ação corporificada está precisamente localizada temporalmente no curso da atividade multimodal e é delimitada por dois símbolos - notando à esquerda seu surgimento e à direita sua conclusão.
- b) Caracterização da ação corporificada: cada ação corporificada conforme

descrito resumidamente. As convenções são baseadas na economia e consistência: colocando não mais anotações do que o necessário e escolhendo anotações que são interpretáveis univocamente. (MONDADA, 2016b, p.1-2)<sup>9</sup>

As transcrições multimodais fazem uso de descrições de texto, bem como imagens, que são integradas à transcrição. É muito importante especificar sempre o momento exato do vídeo a que a imagem se refere para sincronizá-lo com respeito às linhas de fala vocalizada e de ação corporificada. No excerto de transcrição que apresentamos neste artigo, isso está feito tanto através de um símbolo (#) quanto com uma indicação na linha da transcrição dedicada à imagem congelada do vídeo (fig).

Com a transcrição multimodal, é possível observar de que maneira o corpo e seus movimentos são de extrema importância para a compreensão das ações implementadas no cenário em que elas acontecem, neste exemplo, marcamos, por meio de símbolos diversos, os movimentos realizados e a preparação para tais movimentos. Abalizamos o momento em que tais movimentos acontecem, coordenados temporalmente com outros movimentos e com a fala produzida pelos participantes. Todas essas marcações são realizadas com o auxílio de símbolos, tais como “@”, “\*” para sinalizar as ações corporificadas dos participantes e “#” para demarcar o momento exato em que a figura foi capturada.

É importante ressaltar que a imagem está precisamente relacionada ao instante específico durante a conversa (representada pelo símbolo #) e em relação às trajetórias dos outros movimentos. Isso localiza a captura de tela em relação a outras anotações e mostra como vários fluxos de ação corporificada e verbal são articulados em um ponto específico. Sobre as imagens na descrição Mondada (2018, p. 90) afirma que:

[...] as imagens não estão apenas contribuindo para a representação dos movimentos textualmente descritos na transcrição, mas também para sua composição holística e ecologia na qual elas acontecem (sua posição em relação aos objetos relevantes, os corpos em interação, o ambiente e sua materialidade etc.). Nesse sentido, as imagens são um complemento poderoso e indispensável ao que a transcrição textual pode fazer.<sup>10</sup>

---

9 a) *Characterization of the temporal trajectory: each embodied action is precisely temporally located within the course of the multimodal activity, and it is delimited by two symbols – notating on the left its emergence and on the right its completion.* b) *Characterization of the embodied action: each embodied action is shortly described. The conventions are based on economy and consistency: by putting no more annotations than necessary, and by choosing annotations that are interpretable univoquely.* (MONDADA, 2016b, p. 1-2)

10 [...] *images are not only contributing to the representation of the movements textually described in the transcript but also to their holistic composition and the ecology in which they happen (their position relatively to relevant objects, the interacting bodies, the environment and its materiality, etc.). In this sense, images are a powerful and indispensable complement to what textual transcription can do.* (MONDADA, 2018, p. 90)

A transcrição multimodal evidencia a orientação do corpo, os gestos, a entonação e o volume da voz, para além do que está sendo falado. Esse tipo de transcrição pode trazer mais informações para o/a analista sobre o que está acontecendo. Dessa forma, Mondada (2018) explica que as anotações multimodais dizem respeito a dois aspectos fundamentais das condutas corporificadas que não se limitam a gestos, mas dizem respeito a todos os tipos de movimento, são elas:

- (a) sua temporalidade, isto é, sua trajetória emergente e consequente, incluindo preparação e retração, situada precisamente no giro e na ação; e
- (b) sua forma, que é o que torna o movimento reconhecível e descritível. Este último ponto levanta a questão prática e analítica de como descrever esses movimentos, com o objetivo de captar de forma relevante o que a pessoa está fazendo e o que o coparticipante pode vê-la fazendo, dentro de uma perspectiva êmica (ou seja, a perspectiva dos participantes). Dentro da estrutura da AC, essa descrição evita duas armadilhas opostas: imputação de intenções ou estados cognitivos e redução de ações a movimentos fisiológicos. (MONDADA, 2018, p. 90)<sup>11</sup>

Além disso, e mais importante, esses descritores estão presentes não apenas na forma textual, mas também na forma icônica: as capturas de tela. O/A analista deve decidir como distribuir essas descrições entre transcrição, imagens e texto analítico. Nesse sentido, as glosas na transcrição são abreviações para algo que é elaborado em outro lugar; por sua vez, as informações específicas que a transcrição fornece dizem respeito aos detalhes temporais, posições, trajetórias e arranjos dos movimentos. Sendo assim, Mondada (2018, p. 91) destaca que

[...] tempo e forma são necessários para entender o que o corpo está fazendo. Nessa concepção de multimodalidade, o significado de um movimento não é redutível à sua forma, mas está relacionado ao momento em que é produzido; um momento que é significativo em relação ao seu ambiente sequencial e sua posição na ação em curso<sup>12</sup>.

---

11 *A. their temporality, that is their emergent and unfolding trajectory, including preparation and retraction, precisely situated within the turn and the action; and B. their shape, that is what makes the movement recognizable and describable. This latter point raises the practical and analytical question of how to describe these movements, with the aim of relevantly capturing what the person is doing and what the coparticipant can see her/him doing, within an emic perspective (that is, the perspective of the participants). Within the framework of CA, this description avoids two opposing pitfalls: imputing intentions or cognitive states and reducing actions to physiological movements.* (MONDADA, 2018, p. 90)

12 [...] *time and shape are necessary to understand what the body is doing. Within this conception of multimodality, the meaning of a movement is not reducible to its form but is related to the moment in which it is produced; a moment that is meaningful in relation to its sequential environment and its position in ongoing action.* (MONDADA, 2018, p. 91)

Tendo salientado o caminho teórico-metodológico adotado nesta pesquisa, reforçamos a relevância dessas práticas de transcrição multimodal como uma forma de ressaltar que a etnografia da linguagem adotada na análise está intrinsecamente ligada à forma como analistas flagram as ações em curso de modo a tornar possíveis as descrições e análises que são feitas nessa perspectiva. A descrição densa (GEERTZ, 2008), neste caso, se faz com a ajuda de equipamentos audiovisuais e de transcrições minuciosas, na medida em que possibilitam a reconstituição das interpretações demonstradas e demonstráveis nas ações dos próprios participantes. Na próxima seção, apresentamos a análise de um excerto de um atendimento pedagógico, como um exercício etnográfico que ilustra o que defendemos nesta subseção. Importante destacar que, na ótica da Análise da Conversa Multimodal, o foco não está apenas na expressão verbal, mas também nas ações corporificadas, e que, na interação, as ações são constituídas por vários recursos semióticos para além da fala, do olhar, dos gestos, da movimentação do corpo, entre outros.

### Ilustrando o fenômeno: a importância da multimodalidade

Na sequência interacional que escolhemos analisar neste trabalho, a participante categorizada como responsável (RES) informa ao professor (PRO) que sua filha tinha apresentado problemas de visão que a prejudicaram no desenvolvimento das atividades escolares ao longo do ano e relatou também que ela enfrentou questões emocionais que a afetaram diretamente em seu aproveitamento. RES reclama que alguns profissionais não tiveram sensibilidade para lidar com a situação de sua filha. Após essa reclamação, PRO toma o turno, como é possível observar a seguir:

#### Excerto 1 – “ela vai sair de:ssa”

01 PRO \*#apesar de @to:dos os problemas que ela@ teve na vida dela:\*  
res \*.....mexe nos dedos da mão-----\*  
pro @blç MD; desenha circ no ar @  
fig #Fig.3



Fig.3

02 (0,6)  
03 PRO \*ela vai sair de:ssa\*  
res \*cruza dedos da mão \*  
04 \* (0,5) \*  
res \* balança a cabeça \*

05 PRO >entendeu<  
06 (.)

No início do excerto, PRO toma o turno de fala (“apesar de todos os problemas que ela teve na vida dela”, ℓ. 01) ao mesmo tempo em que faz um movimento com as mãos (fig. 3), enquanto RES mexe também os dedos de suas mãos (fig. 3). A movimentação de RES, associada à expressão de seu rosto, dá indícios de uma postura emocionalmente abalada. A pausa (ℓ. 02), ao marcar uma descontinuidade na fala, associada aos movimentos corporais destacados, remonta a uma abordagem cautelosa do tópico<sup>13</sup>.

Apesar de ter se aberto uma oportunidade para que RES tomasse o turno após a fala de PRO, ela não o faz, e ele, então, pelas possibilidades de tomada de turno descritas em Sacks *et al.* (2003 [1974]), continua falando até que se abra um novo lugar relevante para transição. Sua UCT<sup>14</sup> seguinte (“ela vai sair de:ssa”, ℓ. 03), inclusive, tem continuidade sintática com o que foi dito anteriormente (cf. ℓ. 01), o que só reforça a percepção de que ele vem proferindo sua elocução entre pausas de modo a imprimir um ritmo mais lento e, portanto, adaptado à delicadeza do tópico.

Na sequência dessa última UCT, uma nova descontinuidade acontece (ℓ. 04) e continua ditando o ritmo lento da interação. Enquanto isso, as ações corporificadas de RES (ℓℓ.03 e 04) não ocorrem no mesmo ritmo, pois a movimentação das mãos e da cabeça remetem a uma inquietação que não é revelada com palavras. Reflexo desse desajuste entre palavras vocalizadas e ações corporificadas é o pedido de confirmação que PRO faz logo em seguida (“>entendeu<”, ℓ. 05). Essa ação demonstra uma orientação do participante no sentido de checar como seus turnos *estão chegando* para RES. Após esse pedido de confirmação, PRO continua falando:

---

<sup>13</sup> Para os fins deste trabalho, e em razão da limitação de espaço, não há a intenção de discutir o conceito de tópico, que, sabemos, é controverso, especialmente se confrontarmos mais de uma política de investigação. Portanto, solicitamos ao leitor e à leitora que interprete *tópico* aqui, à luz do senso comum, como sendo o *assunto* que está sendo tratado pelos participantes neste encontro social específico. Sobre a associação do termo com o adjetivo *delicado*, recorremos a Erickson e Shultz, que afirmam: “a condução da interação de momento a momento envolve um equilíbrio tão *delicado* de cooperação entre os parceiros interacionais que a ação social inadequada de um dos parceiros confunde os outros parceiros” (1982, p. 70, *itálico adicionado*. Nossa tradução para: *The conduct of interaction from moment a moment involves such a delicate balance of cooperation among the interactional partners that inadequately social action by one of the partners throws the other partners off.*)

<sup>14</sup> A sigla UCT é uma referência à Unidade de Construção de Turno. Esse conceito é formulado por Sacks *et al.* (2003 [1974]) para explicar que “[h]á vários tipos de unidade com as quais um falante pode começar a construir um turno” (p. 16). Essas unidades definem uma possível projeção para a troca de falantes, na medida em que, ao fim de cada unidade (cada UCT, portanto), abrem-se espaços de oportunidade para que os demais interlocutores comecem a falar após cada UCT, ou para que o falante corrente continue falando, caso ninguém tome para si o turno de fala.



interação como a sequencialidade pode ser melhor compreendida graças à multimodalidade. O estudo da multimodalidade revela ações corporificadas que são interacionalmente relevantes e visivelmente orientadas para tal, realizadas em diferentes temporalidades e por diferentes partes do corpo.

Nesse sentido, destacamos que os detalhes multimodais mobilizados pelos participantes para organizar suas ações são potencialmente infinitos, pois ultrapassam as formas etnográficas de análise convencionais e incluem recursos situados e *ad hoc* que dependem do tipo de atividade e de sua ecologia específica. Sendo assim,

[...] conseqüentemente, o sistema de transcrição multimodal tem que ser flexível o suficiente para acomodar todos os tipos possíveis de detalhes relevantes feitos localmente e, ao mesmo tempo, representá-los de uma forma semelhante, coerente e robusta que é essencial para análises sistemáticas. (MONDADA, 2018, p. 95)<sup>15</sup>

Considerando que a Análise da Conversa foi desenvolvida com base em atividades em que a fala-em-interação constitui o principal recurso, os dados de vídeo permitem auxiliar o trabalho de uma observação etnográfica, cujas pesquisas possibilitam investigar a movimentação de corpos em um determinado espaço, independentemente de ter havido ação vocalizada ou não, uma vez que, para os participantes presentes, o curso de ação principal é organizado por uma diversidade de recursos multissemióticos disponíveis a quem participa da ação e também a quem observa para fins de pesquisa.

Neste artigo, procuramos mostrar que não são apenas as palavras vocalizadas que produzem ações analiticamente relevantes nas interações sociais. Como foi possível observar, o *desencontro* entre o que se diz e o que se faz com expressões faciais e corporais é muito mais representativo dos propósitos dos participantes na interação do que o conteúdo que se abordou no encontro entre RES e PRO analisado nos excertos 1 e 2 anteriores. Os turnos de fala de PRO só ganham contornos de que ele está de fato disposto a contribuir com RES em relação às dificuldades relatadas sobre a estudante/filha a partir de suas ações corporificadas e de seus contornos prosódicos. Já em relação à RES, o desconforto inicial perceptível pelos movimentos das mãos e dos dedos no início da sequência dá lugar a movimentos concordantes com a cabeça, sinalizando afiliação às ações de PRO.

Observamos que a sequencialidade é uma preocupação fundamental para AC, porque,

---

15 *Consequently, the multimodal transcription system has to be flexible enough to accommodate all possible kinds of locally made relevant details and at the same time represent them in a similar, coherent, and robust way that is essential for systematic analyses.* (MONDADA, 2018, p. 95)

junto aos estudos multimodais, revelam sua complexidade, levantando questões do tempo êmico das ações, as temporalidades múltiplas das condutas, a interação de sucessividade e simultaneidade e a coordenação de componentes de ação relacionados, mas distintos. Nesse sentido, a multimodalidade auxilia na análise da fala-em-interação institucional em que o contexto dos locais de trabalho, com suas configurações espaciais complexas, a difusão de atividades mediadas tecnologicamente e a presença de documentos e outros artefatos só poderiam ser investigados por meio de gravações de vídeo. Segundo Mondada (2018, p. 98), “essas atividades apresentam desafios interessantes para a transcrição multimodal e nos convidam a dar um passo adiante na elaboração de anotações multimodais que não dependem de fala”.<sup>16</sup>

Além disso, a autora também salienta que as múltiplas temporalidades também são articuladas em simultaneidades ordenadas e sequencialmente mostram que, dentro de uma atividade interacional emergente, várias outras ordens sequenciais podem ser alcançadas simultaneamente por coparticipantes, que podem atender toda ou parte delas. Assim, as interações corporificadas são caracterizadas por várias projeções ocorrendo ao mesmo tempo, iniciadas em vários momentos e respondidas de forma ordenada, mais cedo ou mais tarde.

Por fim, é preciso levar em consideração que as pesquisas qualitativas, segundo Agrosino (2009), são aquelas que deixaram de ser somente de ordem *não quantitativa*, em detrimento da observação estrita do registro de práticas interacionais e posterior análise do material gerado a partir do convívio direto e sistemático com pessoas reais de carne e osso, e não mais com a pecha de *informantes*. Isso quer dizer que, para produzir análises socialmente relevantes nessa abordagem, outros parâmetros de aproximação do campo precisam ser adotados. Com isso, somado aos avanços tecnológicos e econômicos que permitiram às câmeras e aos microfones se tornarem cada vez mais portáteis e acessíveis, essas novas possibilidades se consolidam com o uso desses aparelhos. O tratamento desses dados permite, pois, que as notas de campo, antes produzidas apenas com base na memória, agora possam ser enriquecidas ou mesmo confrontadas com o que se observa nos vídeos. E desse modo, possam compor recursos teórico-metodológicos que caminham juntos em benefício de um entendimento muito mais diversificado e multitemporal de como a sociedade funciona para além das formulações verbais que cada ator social pode produzir.

---

16 *These activities present interesting challenges for multimodal transcription and invite us to take a step further in the elaboration of multimodal annotations that do not depend on talk.* (MONDADA, 2018, p. 98).

## Referências

- AGROSINO, M. Etnografia e Observação Participante. In: FLICK, U. **Pesquisa Qualitativa**. Bookman, Artmed, 2009.
- AUSTIN, J. **Quando dizer é fazer**. Porto Alegre: Artes Médicas, 1990 [1962].
- COUTO, C. S. L. **O Choro na interação: o gerenciamento de ações em episódios nos quais o choro é tornado relevante**. Vitória, ES. Dissertação de Mestrado. Universidade Federal do Espírito Santo. UFES, 2021. 113 p.
- CRUZ, F. M.; OSTERMANN, A. C.; ANDRADE, D. N. P.; FREZZA, M. O trabalho técnico-metodológico e analítico com dados interacionais audiovisuais: a disponibilidade de recursos multimodais nas interações. **DELTA**, n. 35, v. 4, pp. 1-36, 2019. Disponível em: <https://doi.org/10.1590/1678-460X2019350404>. Acesso em: 10 out. 2022.
- DEL CORONA, M. Fala-em-interação cotidiana e fala-em-interação institucional: uma análise de audiências criminais. In: LODER, L. L. e JUNG, N. M. (orgs.), **Análises em fala-em-interação institucional: A perspectiva da análise da conversa etnometodológica**. Campinas, Mercado de Letras, 2009, p. 13-44.
- DREW, E.; HERITAGE, J. Analyzing talk at work: An introduction. In: DREW, P; HERITAGE, J. (Orgs.). **Talk at work: Interaction in institutional settings**. Cambridge: Cambridge University Press, 1992, p. 3-65.
- ERICKSON, F.; SHULTZ, J. **The Counselor as Gatekeeper: Social Interaction in interviews**. New York: Academic Press, 1982.
- GARCEZ, P. M.; OSTERMANN, A. C. Glossário conciso de Sociolinguística Interacional. In: RIBEIRO, B. T.; GARCEZ, P. M. (Org.). **Sociolinguística Interacional**. São Paulo: Loyola, 2002. p. 257-264.
- GARCEZ, P. M.; BULLA, G. S.; LODER, L. L. Práticas de pesquisa microetnográfica: geração, segmentação e transcrição de dados audiovisuais como procedimentos analíticos plenos. **D.E.L.T.A.**, 30.2, p. 257-288. 2014. Disponível em: <http://www.scielo.br/pdf/delta/v30n2/0102-4450-delta-30-02-0257.pdf>. Acesso em: 10 out. 2022.
- GEERTZ, C. **A interpretação das culturas**. Rio de Janeiro: LTC, 2008, 323p.
- GOODWIN, C. Action and embodiment within situated human interaction. **Journal of Pragmatics**, v. 32, n. 10, p. 1489-1522, 2010.
- JEFFERSON, G. Glossary of transcript symbols with an introduction. In: LERNER, G. H. (Org.). **Conversation Analysis: Studies from the First Generation**. Amsterdã: Benjamins, 2004, p. 13-31.
- LEVINSON, S. C. Action Formation and Ascription. In: SIDNELL, J.; STIVERS, T (Eds.), **The Handbook Of Conversation Analysis**. Hoboken, NJ: Wiley, 2012, p.103-130.

MONDADA, L. Challenges of multimodality. Language and the body in social interaction. **Journal of Sociolinguistics**, v. 20, p. 261–395. 2016a

MONDADA, L. **Conventions for multimodal transcription**. Disponível em: [https://franzoesistik.philhist.unibas.ch/fileadmin/user\\_upload/franzoesistik/mondada\\_multimodal\\_conventions.pdf](https://franzoesistik.philhist.unibas.ch/fileadmin/user_upload/franzoesistik/mondada_multimodal_conventions.pdf). p. 1-8. 2016b.

MONDADA, L. Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. **Research on Language and Social Interaction**. 51:1, 85-106. 2018.

PEROBELLI, R. Análise sequencial multimodal do gerenciamento da fala sobreposta em uma entrevista televisiva. In: TOMAZI, M. M; SESSA, A. (Orgs.). **Discursos contemporâneos: saúde, educação, política e interseccionalidades**. Campinas: Pontes, 2022, pp.175-199.

SACKS, H; SCHEGLOFF, E. A.; JEFFERSON, G. Sistemática Elementar para a organização da tomada de turnos para a conversa. **Veredas**, v. 7, n. 1-2, p. 9-73. Trad. de SACKS, H.; SCHEGLOFF, E. A.; JEFFERSON, G. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, v. 50. 2003.

SCHEGLOFF, E. A. Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. **American Journal of Sociology**, n. 97, v. 5, p. 1295–1345, 1992.

STREECK, J.; GOODWIN, C.; LEBARON, C. Embodied interaction in the material world: an introduction. In: STREECK, J.; GOODWIN, C.; LEBARON, C. (Ed.). **Embodied Interaction: language and body in the material world**. New York, NY: Cambridge University Press. p. 1 – 26, 2011.

WATSON, R.; GASTALDO, É. **Etnometodologia & Análise da Conversa**. Petrópolis, Rio de Janeiro: Vozes. Editora PUC- Rio, 2015. 183 p.

### Sobre os autores

*Roberto Perobelli* (Orcid iD: <https://orcid.org/0000-0002-8130-1061>)

Professor adjunto no Departamento de Línguas e Letras e no Programa de Pós-Graduação em Linguística da Universidade Federal do Espírito Santo (UFES), na linha de pesquisa Linguística Aplicada. Líder do Grupo Linguagem, Interação e Etnometodologia (GLIE), cadastrado no Diretório de Grupos do CNPq desde 2018, que reúne pesquisas e trabalhos realizados no campo da Análise da Conversa e da Etnometodologia nos mais diversos cenários, entre os quais destacam-se o escolar e o midiático.

*Luana Santos Lemos* (Orcid iD: <https://orcid.org/0000-0002-6475-1458>)

Formada em Letras pela Universidade Federal do Espírito Santo, Mestre e Doutora em Estudos Linguísticos pela UFES. Atualmente está como Subsecretária de Gestão Pedagógica pela Secretaria de Educação do Município de Vitória, capital do Espírito Santo, e atua como professora convidada no Programa de Pós-Graduação Profissional em Ciências Contábeis e Administração, na Linha de Pesquisa Gestão Escolar, da Fundação de Pesquisa e Ensino (FUCAPE). É professora efetiva de Língua Portuguesa da Rede Estadual de Educação do Espírito Santo (SEDU) e tem experiência no Ensino

Superior na área de Linguística, Ensino de Língua Portuguesa, Literatura Brasileira, Educação a Distância (EAD) e Gestão Escolar.

Recebido em outubro de 2022.

Aprovado em dezembro de 2022.