

# Documentação multimodal de interações com crianças com Transtorno do Espectro do Autismo: corpo, língua e mundo material

## Multimodal documentation of interactions with children with Autism Spectrum Disorders: Body, language and material world

Fernanda Miranda da Cruz<sup>1</sup>

fernanda.miranda.cruz@gmail.com

Universidade Federal de São Paulo

**RESUMO** - Este artigo analisa as relações entre percepção do ambiente físico material e a emergência de momentos de iniciativa espontânea de fala por parte das crianças com Transtorno do Espectro do Autismo (TEA) em situações de copresença com o outro. Nosso enfoque recai sobre (i) os recursos multimodais (verbais e não verbais) estruturantes dessas interações e (ii) sistematização de alguns procedimentos de pesquisa relativos à documentação e investigação de interações sob uma perspectiva corporificada das práticas interativas. Essa sistematização tem dois objetivos: (i) oferecer um estatuto teórico-conceitual ao trabalho de constituição de *corpora* audiovisuais para investigações linguísticas de perspectiva multimodal e (ii) dar visibilidade às implicações da constituição desses tipos de *corpora* para a investigação e teorização da interação social em quadros de TEA.

**Palavras-chaves:** multimodalidade, interação corporificada, videoanálises, TEA, mundo material.

**ABSTRACT** - This article analyzes the relationships between the perception of the physical environment and the emergence of spontaneous moments of talk initiatives produced by children with Autism Spectrum Disorder (ASD). Our focus is on the multimodal resources (verbal and nonverbal) displayed in these interactions and its representation, transcription, and analysis. This systematization has two aims: (i) to offer a theoretical-conceptual status to the work of generating audiovisual *corpora* for linguistic investigations from a multimodal perspective and (ii) to discuss the implications of the constitution of corpora audiovisuais of naturally-occurring-interactions for the investigation and theorization of the interaction in ASD.

**Keywords:** multimodality, embodied interaction, videoanalysis, ASD, material world.

### Introdução

Podemos supor que nas interações envolvendo sujeitos com Transtornos do Espectro do Autismo, das quais participam pessoas com e sem diagnóstico de TEA, encontramos formas diversas de interação, incluindo tanto padrões interativos mais conhecidos e praticados socioculturalmente como outros padrões que emergem das circunstâncias situadas e locais decorrentes do autismo, incluindo aqui as competências, as limitações e as (im)possibilidades de todos os participantes envolvidos. Os Transtornos do Espectro do Autismo (TEA) correspondem a um grupo relativamente heterogêneo de condições que afetam o desenvolvimento neurocognitivo e que comprometem, de formas e graus distintos, o desenvolvimento da linguagem e a participação dos sujeitos em interações sociais (Lai *et al.*, 2014; Cunha *et al.*, 2015). Embora o TEA possa caracterizar-se por comprometimentos nesses dois eixos centrais, interação/comunicação social e comportamentos/

interesses restritos e repetitivos, a forma como os TEAs afetam cada sujeito autista é singular, caracterizando cada caso e cada sujeito autista como únicos. No que diz respeito ao desenvolvimento da linguagem verbal, por exemplo, algumas crianças autistas não desenvolvem a fala, não fazendo uso da linguagem verbal, ou ainda fazendo um uso muito limitado, considerado na literatura clínica como casos de TEA não-verbais (Walenski *et al.*, 2006). Dentro de uma abordagem clínico-diagnóstica dos TEA, conforme aponta Tamanaha e Perissinoto (2015, p. 9), “o grau de gravidade varia de pessoas que apresentam um quadro leve e com total independência e discretas dificuldades de adaptação, até aquelas pessoas que serão dependentes para as atividades de vida diárias (AVDs), ao longo de toda a vida”.

Do ponto de vista da interação social, um comprometimento no engajamento social decorrente dos TEA afeta, de distintas maneiras, as formas pelas quais tanto os sujeitos com TEA quanto seus interlocutores, sejam eles

familiares, próximos ou profissionais, organizam e estruturam suas interações. Estudos sobre o autismo atentos à toda organização da interação social e não apenas às produções, ações e comportamentos dos sujeitos diagnosticados com TEA são relativamente recentes, mas crescentes e realizados sobretudo, mas não só, a partir de trabalhos como os de Ochs, Solomon e Sterponi (2005); Ochs e Solomon (2010); Sterponi, Kirby e Shankey (2014).

Em nossa investigação, do ponto de vista analítico-metodológico, as tentativas de compreender e descrever as formas pelas quais os sujeitos com TEA interagem com o mundo social e, portanto, com toda uma ecologia de sistemas e significados socioculturais, mentais, cognitivos tem como ponto de partida considerar que, nas interações de co-presença entre sujeitos com e sem TEA, estamos diante de situações nas quais todos os envolvidos experienciam, no limite, um difícil de trabalho de coordenação de social com os outros. Esse conjunto de possibilidades de coordenação social ou de sociabilidades distintas, dentre as quais a sociabilidade autista (Ochs e Solomon, 2010) e Solomon (2008), seria configurado por fatores que incluem as experiências de vida de cada indivíduo, a maturidade desenvolvimental, seus talentos e habilidades, suas condições neuropsicológicas, físicas e sociais (Ochs e Solomon, 2010, p. 71).

A formulação acima já nos inscreve em um certo modo de conceber tais interações, a saber, elas assumem a forma que assumem através de um trabalho de coordenação, nem sempre fácil, de todos os envolvidos. As evidências, observações e análises produzidas no quadro dos trabalhos de Ochs e Solomon (2010) resultam em um conjunto de condições (um algoritmo) que indicam como a sociabilidade autista “cresce e diminui de acordo com as condições sociais e interacionais” (Ochs e Solomon, 2010, p. 86).

Atentar para a organização da interação social em contextos envolvendo a participação de crianças com TEA tem oferecido uma contribuição importante para repensarmos e problematizarmos, analítica e metodologicamente, a investigação, a interpretação e a caracterização da linguagem nesses contextos. Sterponi *et al.* (2014), Sterponi e Shankey (2014) e Sterponi e Kirby (2016) têm mostrado, por exemplo, como fenômenos tipicamente associados à linguagem autista (como ecolalias; respostas inadequadas durante uma conversa ou topicamente irrelevantes; construções sintaticamente incongruentes; falas vazias ou não comunicativas, e expressões muito literais ou formulaicas) podem ser reexaminados a partir de uma perspectiva multidimensional da linguagem no autismo. Uma perspectiva multidimensional evoca, de forma articulada, a concepção de linguagem como um trabalho interacional, como ação social e como um modo de experiência do sujeito no mundo, envolvendo tanto uma dimensão subjetiva, relativa aos processos relacionados à constituição do *self*, quanto intersubjetiva, relativas ao

outro. A linguagem, entendida como um trabalho interacional, emerge dentro de uma dinâmica de monitoramentos e ajustes mútuos entre falantes e ouvintes (Goodwin, 1981). Assim também a fala autista e suas características linguístico-pragmáticas emergem como um produto de um contínuo e constante processo colaborativo entre os participantes envolvidos e não como uma produção linguística do indivíduo com TEA, estritamente relacionada a um déficit. Uma perspectiva multidimensional da linguagem no TEA, intimamente ancorada nas práticas linguísticas construídas sociointeracionalmente, oferece-nos a possibilidade de repensarmos a forma pela qual as produções linguísticas dos indivíduos com diagnóstico de TEA podem ser teorizadas e analisadas.

É dentro dessa perspectiva interacional aplicada a contextos específicos de interações envolvendo crianças com TEA que temos conduzido o Projeto *Gestos Mínimos* (Cruz *et al.*, 2017, Comitê de Ética em Pesquisas da Universidade Federal de São Paulo, processos 0052/2015 e 1172/2016). O projeto *Gestos Mínimos*, iniciado em 2014, tem interesse no contexto interacional envolvendo a participação de crianças com diagnóstico de TEA, com ênfase particular no papel dos gestos, da fala, do corpo e do ambiente material, envolvendo objetos ou espaço físico, na configuração e construção dessas interações. No trabalho de documentação multimodal a ser descrito detalhadamente neste artigo, nosso objetivo é mostrar procedimentos metodológicos e analíticos que nos permitam: (i) identificar e dar visibilidade à diversidade de recursos semióticos (verbais e não verbais); (ii) sistematizar as relações com ambiente físico e com objetos co-presentes no espaço utilizados por todos os participantes presentes na interação e, (iii) analisar se e como tais recursos semióticos estão espacial e temporalmente coordenados para construção local e situada da interação. Uma importante contribuição resultante de uma análise fundamentada nos pressupostos e métodos dos estudos interacionais baseados na geração de *corpora* audiovisuais está nas possibilidades de oferecermos maior inteligibilidade aos recursos mobilizados pelas crianças com TEA para se engajarem em uma ação com o outro e problematizarmos a forma como uma produção atípica pode ser analisada para além da ideia de déficit ou disfunção (Sterponi *et al.*, 2014).

O estudo das interações envolvendo crianças com TEA reportado neste artigo fundamenta-se em uma perspectiva corporificada da interação (*embodied interaction*, Streeck *et al.*, 2011). Essa perspectiva nos convida também a tomar contato com discussões em torno dos procedimentos metodológicos implicados na investigação e na análise dos aspectos corporais e materiais nessas interações.

Inicialmente, apresentaremos as bases de uma perspectiva sociointeracional e corporificada aplicada à análise de interações envolvendo crianças com diagnóstico de TEA. Em seguida, apresentaremos uma descrição dos procedimentos para concepção e construção do *corpus*

audiovisual *Gestos Mínimos* e discutiremos algumas implicações gerais da constituição de *corpora* audiovisuais de interações naturalísticas para a investigação e teorização da interação social e TEA com base nas reflexões produzidas no campo da viodeanálise (Knoblauch *et al.*, 2006). Finalmente, proporemos uma análise que evidencie nossa atenção especial dada: (i) ao papel que os recursos multimodais mobilizados pelos sujeitos (autistas e não-autistas) desempenham na estruturação dessas interações e, (ii) às relações estabelecidas com o ambiente físico-material. Para isso, propomos a análise de três excertos em que as iniciativas de falas de crianças com TEA com o outro participante co-presente é antecedida e estruturada por relações visuais e gestuais com objetos co-presentes no ambiente em que se encontram.

### Uma perspectiva corporificada da interação social

Uma perspectiva corporificada da interação social não se guia, em sua investigação, por uma primazia dos recursos verbais frente aos recursos gestuais e materiais (Streeck *et al.*, 2011; Mondada, 2016). Os estudos sobre a construção e organização da interação social produzidos a partir de uma perspectiva multimodal e corporificada podem ser abrigados dentro de um movimento identificado como uma virada corporificada (*embodied turn*, Nevile, 2015; Mondada, 2016), que reúne as várias tentativas teórico-epistemológicas de integração dos estudos já existentes sobre linguagem e de novos estudos sobre corpo, gestos e mundo material permitido o surgimento e a consolidação de diversas e diferentes abordagens sobre a multimodalidade da ação, da cognição e da interação (Mondada, 2016). Nos estudos sobre a multimodalidade da interação produzidos a partir de uma perspectiva corporificada, fundamentalmente, considera-se que construímos os espaços interacionais de forma multimodal. Dessa forma, uma ação, verbal ou não, é construída graças a uma ecologia de sistemas semióticos, estruturalmente distintos entre si, mas intrinsecamente relacionados (Goodwin, 2000). Uma perspectiva multimodal oferece assim a possibilidade de fazermos uma descrição desses vários tipos de recursos multimodais mobilizados na interação dentro de um enquadre analítico (o sociointeracional) que reconhece a diversidade e a especificidade sistêmica desses recursos semióticos usados pelos participantes na interação, ao mesmo tempo em que se preocupa em compreender e descrever como tais recursos interagem uns com os outros para construirmos localmente nossas ações, dentre elas, a fala-em-interação. Esses recursos seriam de naturezas semióticas distintas: linguísticos (aspectos gramaticais, prosódicos, sintáticos, entonacionais e léxicais, por exemplo); corporais (posturas corporais, gestos, direcionamentos do olhar, mímica facial, movimentos da testa, postura do corpo, *etc*) e materiais (relações múltiplas

que temos de manuseio, referência, percepção de objetos, sensorialidade com elementos do espaço físico).

Trazendo essa discussão para nossas investigações em particular baseadas nas interações documentadas no *corpus Gestos Mínimos*, formulamos que se interagir com o outro significa também interagir a partir de uma ecologia material de objetos, corpos e práticas presentes em um mundo social, coletiva e materialmente organizado (Streeck *et al.*, 2011), temos que compreender essa organização coletiva e material pode: (i) nos oferecer pistas importantes sobre as possíveis (e potenciais) formas de interagir; (ii) nos oferecer uma descrição das produções verbais e gestuais das crianças com TEA situadas, em termos sequenciais e temporais, aos contextos de emergências dessas produções no curso da interação com o outro e configurados pelo ambiente físico-material local.

Os procedimentos para a condução de uma pesquisa multimodal das práticas interativas dentro de uma perspectiva corporificada da interação social estão fundamentalmente comprometidos com as possibilidades metodológicas e analíticas de se identificar e sistematizar a coordenação temporal e espacial entre os recursos multimodais para os quais os participantes orientam-se nas atividades constitutivas da fala-em-interação. Mantendo esse princípio, apresentaremos a seguir os elementos que compõem nosso trabalho de documentação multimodal, nos moldes sugeridos por Mondada (2008a), envolvendo a concepção, constituição, o tratamento e a exploração de um *corpus* audiovisual de interações das quais participam crianças com TEA com vistas a dar conta de uma análise corporificada dessas interações.

### O *corpus* audiovisual *Gestos Mínimos*: concepção, constituição e exploração multimodal

O *Corpus Gestos Mínimos* é composto, no momento, por um total 20 horas de registro audiovisual visualizadas na íntegra e parcialmente transcritas. Descreveremos aqui os critérios de concepção, constituição e exploração do *corpus*. Tal descrição visa a compartilhar algumas questões metodológicas e analíticas envolvidas na condução de pesquisas multimodais da interação social aplicada a um contexto específico que são as interações das quais participam crianças com TEA.

Uma investigação multimodal da interação requer, antes de tudo, a possibilidade de termos registros em vídeo das interações. A discussão sobre questões técnicas, práticas, éticas, metodológicas, analíticas, teóricas e epistemológicas envolvidas no trabalho prático com vídeo é um tema relativamente recente nos estudos interacionais. Knoblauch *et al.* (2006), ao sistematizarem algumas dessas questões, propõem que esse trabalho seja em si um tópico de reflexão teórico-analítica. Também Mondada (2008b, 2012), ao discutir as implicações da abordagem

praxeológica das práticas de vídeo, mostra como o vídeo se torna um lugar de preservação de elementos e detalhes importantes da organização da interação ao mesmo tempo em que configura os fenômenos de análise. Para Mondada (2003, 2008b) as práticas de registro e de transcrição de dados gerados a partir de vídeo seguem um princípio de disponibilidade: se os detalhes multimodais são relevantes antes de tudo para os participantes (perspectiva êmica), então, para análise, é importante que tais detalhes sejam disponibilizados através das formas de realização dos registros (escolhas de dispositivos, ângulos, por exemplo) e através das formas pelas quais o material é transcrito, representado e apresentado.

Os aspectos mencionados acima e presentes na literatura sobre o tema têm implicações diretas para a forma como concebemos, produzimos e exploramos os dados de interações no TEA a partir de uma documentação multimodal. Assim, consideramos importante, dada a crescente reflexão sobre as metodologias de uso e análise de vídeo em pesquisas interacionais, trazermos uma descrição mais detalhada seguida de algumas reflexões relativas ao tratamento do material registrado em vídeo para os fins analíticos a que nos pretendemos.

Para iniciarmos, adotamos a distinção entre dados primários (relacionados ao registro) e dados secundários (dados transcritos, formas de notação, formas de representação dos dados). As escolhas de como realizar um registro audiovisual incluem: local de instalação dos equipamentos; formas de operação dos equipamentos durante o registro (com ou sem a presença de um operador); a quantidade e os tipos de equipamentos; os tipos de zoom ou movimentos de captura realizados ou não durante as gravações, etc. Escolhas dessa natureza não são feitas sem consequências importantes para o que se pode ver com a câmera (Mondada, 2006), para a transcrição e para a análise do material registrado. Assim, compartilhamos da consideração de que o registro em vídeo “pode gerar dados primários extremamente complexos, sujeitos a muitas transformações” (Mondada, 2008a, p. 150). Dentre essas transformações, consideramos as técnicas de anonimização, as formas de anotação e de alinhamento entre as distintas modalidades (corporais e verbais), o auxílio de ferramentas (softwares) e a adoção de uma convenção de transcrição multimodal.

Para a construção do *Corpus Gestos Mínimos* e para o registro propriamente dito, adotamos o método de instalação de uma ou duas câmeras fixas em locais de rotinas interacionais das crianças do estudo. Os detalhes do contexto interacional dos três excertos analisados neste artigo serão descritos mais abaixo. No que diz respeito à transcrição, tomamos como base um princípio geral para a transcrição multimodal realizada no enquadre analítico dos estudos interacionais (Goodwin 1986, 2000; Erickson, 2010; Mondada, 2008a, 2014, 2016; Heath, 1986): o de que os recursos multimodais objetos de transcrição são

aqueles relacionados à temporalidade da fala e da ação, cujo sentido apenas se dá em uma relação de sincronização e coordenação de diferentes ações e modalidades do mesmo participante e em uma relação de coordenação dessas ações e modalidades com as ações dos demais co-participantes presentes.

É assim que para o estudo apresentado aqui, dentre as possibilidades de ferramentas para anotação multimodal existentes, recorremos ao uso do software ELAN (Wittenburg *et al.*, 2006). O uso de software em pesquisas sociointeracionais e psicolinguísticas têm permitido uma vasta documentação de interações em situações naturalísticas. O ELAN possui recursos interessantes, sobretudo para aplicação em pesquisas multimodais cujo cerne está na sincronização e coordenação temporal e espacial de modalidades de naturezas distintas, percebidas auditiva e visualmente. Em nossas investigações, estamos interessados em ver, como anunciamos, como as interações constroem-se multimodalmente e em como os distintos recursos semióticos coordenam-se espacial e temporalmente no fluxo da interação conjuntamente construída. Embora não seja obrigatório, o recurso a um software pode ser muito útil ao trabalho de visualização, de anotações e de alinhamento temporal das distintas modalidades, dentre elas a fala, que servirão de base à posterior transcrição multimodal dos dados audiovisuais. Assim, o trabalho de realização de anotações e de transcrição multimodal pode variar consideravelmente segundo a escolha em se usar ou não tais software. Por isso tal escolha também gera questões importantes para o próprio campo de pesquisas baseadas em vídeo, como apontam (Knoblauch *et al.*, 2006), uma vez que interfere nas possibilidades e nos resultados analíticos. Nas interações envolvendo crianças com TEA analisadas neste artigo, a interpretação de suas ações interacionais e de suas produções verbais dependem necessariamente da possibilidade de alinhamento temporal realizado com o ELAN entre distintas modalidades mobilizadas, gestuais e verbais, coordenadas à referência a objetos presentes na situação, como mostraremos a seguir.

Um aspecto aparentemente pouco importante que merece menção é que todas as ferramentas disponíveis para alinhamento temporal de modalidades gestuais e verbais exigem um certo tempo de dedicação por parte do pesquisador e, obviamente, não realizam transcrições automáticas. Na literatura do campo da viodeonálise, Knoblauch *et al.* (2006) reconhecem que “isso pode causar um certo delay no trabalho analítico, uma vez que um tempo considerável é consumido em mero ‘craftsmanship’” (Knoblauch *et al.*, 2006, p. 15, tradução nossa). Tem havido uma tendência cada vez mais presente de se trazer a público os rascunhos, esboços desse artesanato que tornam visíveis passos metodológicos importantes do trabalho com vídeo em pesquisas linguísticas (Mondada, 2008a).

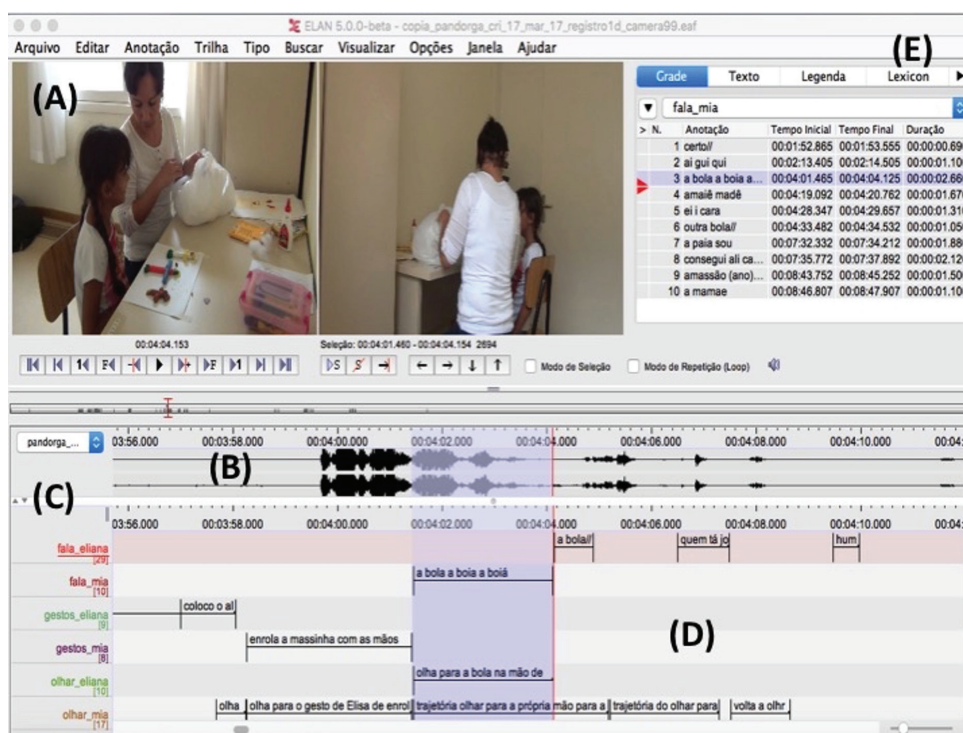
Se o uso de uma ferramenta técnico-tecnológica pode influenciar significativamente não apenas o tempo,



mas o próprio trabalho de transcrição, representação e análise dos dados resultante desse uso, podemos então depreender que a relação entre o saber técnico-tecnológico e a geração de dados multimodais transcritos não é apenas um detalhe do tema, mas senão um aspecto sobre a própria produção de conhecimento no campo dos estudos linguístico-interacionais. Consideramos que tornar explícitos tais aspectos práticos (os rascunhos) e metodológicos em uma empreitada investigativa das ações verbais ou corporais de crianças com TEA parece colaborar também para elucidar as possibilidades analíticas que se abrem a partir das escolhas feitas pelo pesquisador para os campos de investigação da linguagem e da interação em quadros de TEA. Levando essa discussão preliminar em consideração, descrevemos abaixo os procedimentos de visualização, anotação e transcrição do corpus feitos com o auxílio do ELAN. Para ilustrar algumas de suas funções e o trabalho de tratamento para exploração multimodal dos dados, inserimos abaixo (Figura 1) uma captura de tela do software correspondente a um pequeno trecho transcrito de um dos excertos analisados mais adiante (Excerto 2). Esse tratamento prévio com ELAN, que antecede e permite-nos chegar à versão final da transcrição, como nos três excertos deste artigo, se aplica a todo o *corpus* já transcrito ou em fase de transcrição.

Em (A) temos a mídia vídeo e em (B) temos a mídia áudio do registro (dados primários). Neste exemplo,

temos dois vídeos referentes ao registro da situação interativa sob dois ângulos distintos feito por duas câmeras, sincronizadas entre si e com a mídia áudio (B) através da função *Sincronizador de mídias* do ELAN. Sobre os rostos de uma das participantes (a que está em posição frontal ao ângulo da câmera) aplicamos sobre a imagem uma técnica de anonimalização possível para os formatos de apresentação pública do material visual. Outras técnicas de anonimalização são possíveis. Em (C), temos as trilhas de anotações. As trilhas são fundamentais para o trabalho de anotação e posterior transcrição multimodal dos dados. O pesquisador pode optar por criar trilhas de acordo com a natureza dos dados, suas práticas de visualização e seus fins de análise. Para nosso estudo, optamos por criar trilhas que nos permitissem dar conta de ocorrências que são descritas na literatura sobre o TEA (como ausentes, como alteradas ou como presentes) e ainda de ocorrências emergentes no *corpus* em questão. Vejamos um exemplo. O direcionamento do olhar tem sido considerado, desde as primeiras descrições do TEA (Kanner, 1943; Asperger, 1991 [1944]), como um traço autista em termos de movimentos de desvios, aversões, padrões atípicos ou mesmo ausentes. Korkiakangas e Rae (2014) têm identificado e descrito, a partir de suas investigações baseadas em dados naturalísticos, as funções que tais direcionamentos do olhar desempenham nas interações e as contingências interacionais sob as quais sua ocorrência espontânea pode



**Figura 1.** Captura de tela ELAN ilustrando alinhamento e anotação de gestos e fala.

**Figure 1.** ELAN screenshot illustrating alignment and annotation of gestures and speech.

emergir. O mesmo raciocínio se aplica às ocorrências de riso nessas interações, cujo padrão também teria características atípicas ou pouco apreensíveis a partir de padrões sociointeracionais mais conhecidos. Auburn e Pollock (2013), em suas investigações sobre o riso na interação envolvendo crianças com autismo, mostram como uma análise interacional deixa ver algumas funções interacionais do riso. Temos ainda que ações de alinhamento com o outro (que podem ser construídas por meio do olhar ou do riso) são indícios importantíssimos dos momentos de atenção mútua e potencial construção conjunta de ações e de engajamento social dos sujeitos com autismo. Assim, atribuir trilhas específicas para olhar e riso pode, por exemplo, permitir identificar, através de buscas (E), tais ocorrências no *corpus*. A partir das visualizações prévias do material audiovisual do *corpus* e com objetivo de dar conta especificamente de uma análise multimodal, criamos as seguintes trilhas, para cada participante presente: (i) fala de cada participante; (ii) gestos de cada participante (excluindo-se da categoria gestos e ações, os risos e os direcionamentos de olhar, que receberam uma trilha própria cada um); (iii) risos; (iv) direcionamento do olhar.

Finalmente, em (D), temos as anotações realizadas de forma a deixar ver a sincronização e a coordenação temporais e espaciais das modalidades verbais e gestuais, também temporalmente delimitadas e alinhadas com as fontes de áudio e vídeo. Após os alinhamentos realizados, as anotações produzidas em ELAN são exportadas em formato “transcrição tradicional”, uma das opções de exportação disponíveis no software e, finalmente, transcritas a partir de uma convenção de transcrição multimodal.

No que diz então respeito à adoção de uma notação de transcrição em formato texto, adotamos a proposta de convenção de transcrição multimodal de Mondada (2012, 2014, 2016), parcialmente apresentada no Quadro 1, em anexo neste artigo. Podemos nos perguntar se, com tantas possibilidades de recursos tecnológicos e técnicos de representação e tratamento de material visual (em vídeos e imagens), “ainda” é necessário lançar mão de uma transcrição textual em formato lista. Esse é um aspecto que mereceria uma atenção maior e que aqui foi lançado porque faz parte dos fóruns sobre transcrições, do próprio trabalho prático de transcrever e, muitas vezes, da recepção que o material em vídeo tem, já que as anotações realizadas pelo ELAN (e mesmo por outros softwares) podem, por exemplo, ser extraídas como legendas e inseridas no próprio vídeo. Então, para responder à pergunta formulada acima (e presente em nosso trabalho prático-metodológico com os dados de vídeo) sobre a necessidade ou não de transcrevermos em formato texto (ou lista), nos alinhamos à ponderação anunciada por Knoblauch *et al.* (2006) sobre a indispensabilidade da transcrição em texto. Ao introduzir essa discussão em termos da relação entre texto-imagem na apresentação no livro “*Video analysis methodology and methods: Qualitative audiovisual data analysis in*

*sociology*” (2006), dedicada a metodologias e métodos para análises qualitativas baseadas em dados audiovisuais interessadas na forma como as pessoas constroem suas ações sociais na interação, os autores formulam:

No caso da videoanálise, entretanto, este tema exibe um aspecto muito prático: a transcrição de dados inscreve em sua forma particular como o visual é tratado pelo analista, de forma que qualquer desenvolvimento da análise do vídeo dependerá da forma pela qual os dados estão sendo transcritos ou ainda tornados acessíveis para a análise. A análise será cada vez mais capaz de se basear na representação visual, tendo como consequência que as transcrições escritas podem perder sua importância até certo ponto para possivelmente abrir o caminho para uma “mentalidade visual” analítica – um modo de análise que depende menos da palavra escrita do que da visualização e da imaginação. As mudanças tecnológicas em curso também podem afetar os meios (e já estão afetando as formas) em que os estudos estão sendo apresentados (ver, por exemplo, Büscher 2005). No entanto, por enquanto, ainda confiamos nas formas bastante convencionais de transcrições e capturas de imagens que são usadas neste livro. Considera-se que a transcrição de dados não é apenas uma fase preliminar de análise. Ela consiste em uma parte essencial da análise. Transcrever gera observações que são fundamentais para inferências analíticas. Como em pesquisas baseadas em atividades interativas naturalísticas ou em entrevistas, a transcrição de dados de vídeo é simplesmente indispensável (Knoblauch *et al.*, 2006, p. 16, tradução nossa).

E, finalmente, no que diz respeito à apresentação e à publicização dos dados para o formato impresso, como é o caso do formato de artigos, apresentamos abaixo três excertos em forma de transcrição textual, com recurso ao alinhamento temporal de imagens capturadas da tela do vídeo e tratadas com a aplicação de um filtro que permite manter a visibilidade dos traçados dos participantes e do ambiente, anonimizando suas identidades ou reconhecimento.

### O ambiente físico e as iniciativas de fala de crianças com TEA

No *corpus* identificamos que algumas iniciativas espontâneas de fala de crianças diagnosticadas com TEA são antecedidas por movimentos corporais que indicam sua percepção e orientação corporal para objetos co-presentes no ambiente interacional. Analisaremos a forma como essas iniciativas se dão e se desenvolvem na interação. Todos os três excertos analisados são antecedidos por uma sequência de imagens com a descrição dos posicionamentos corporais dos participantes no momento imediatamente anterior à iniciativa de fala (Figuras 2, 3, 4, 5, 6).

O primeiro excerto refere-se a registros audiovisuais de situações naturalísticas relativos às interações das quais participa Luiza, uma criança com diagnóstico de TEA, em interações cotidianas em seu ambiente familiar. Os registros foram realizados pela pesquisadora Caroline Cots (Cots, 2016; Cruz *et al.*, 2017). Luiza interage com recursos verbais e gestuais, ainda que seu



**Figura 2.** Sequência I: Posições corporais de Neide e Luiza.

**Figure 2.** Sequence I: Body positions of Neide and Luiza.

repertório linguístico-lexical seja comprometido se comparado ao de crianças de sua faixa-etária, que era entre 10 anos e 10 meses e 12 anos e 03 meses, no período dos registros realizados. No Excerto 1, Luiza e sua mãe, Neide, estão na cozinha. Neide encontra-se de pé, em frente à pia, lavando a louça. Luiza encontra-se sentada de frente para Neide, conforme a Sequência de imagens I (Figura 2) com a descrição dos posicionamentos corporais das participantes no momento imediatamente anterior à iniciativa de fala Luiza.

Na imagem 1, temos Neide voltada pra pia e Luiza sentada no banco, com o corpo voltado para frente com relação à Neide e ao ângulo da câmera (posição corporal 1).

Na imagem 2, Luiza direciona o olhar para um objeto que percebe no chão, no canto direito da imagem e fora do enquadre da captura do vídeo original (posição corporal 2). Na imagem 3, Luiza direciona seu corpo inteiro para o objeto (posição 3). Na posição 3, Luiza introduz seu turno de fala referente ao objeto que percebe no chão (*que isso aqui//, linha 1, Excerto 1*). Luiza sustenta essa posição corporal 3 até o final da sequência interacional relativa à ação de chamar a atenção de Neide para o objeto que está no chão da cozinha. No excerto transcrito abaixo, os gestos de Luiza estão indicados por (\*), os gestos de Neide estão indicados por (@) e as imagens alinhadas à transcrição estão indicadas por (#).

**Excerto 1: CELA\_11\_2015\_caf.**

```

01    LUI      *@que isso aqui//                      *
02    lui      >>* volta o torso e o olhar p/o chão-->*
03    nei      >>@voltada p/ a pia--> (l. 14)
04          *(1.0) + **(0.2)
05    lui      *..volta o torso p neide e a toca com mão dir.-->
06    lui      **..olha para neide----->
07          é seu//
08          (0.3) * + (0.7)**
09    lui      --,,,>*
10          -----,,,>**
11          MA-mãe**
12          **..--olha p/ o chão----->
13          (.) @@ **
14    nei      @@olha p/ o chão--->
15    lui      ---->**
16    NEI      oi//**
17    lui      **olha p/ neide-->
18    LUI      I:s**so@ aqui** é seu//
19          -->**..... **---olha p/ o chão--->
20    nei      ----->@..--volta o torso para a pia--->>
21          (0.9)** + **(0.5)

```

22           lui                 --> \*\* + \*\*----olha p neide--->  
23   **NEI**           eu \*\*nu:m sei o [que que é]\  
24   lui           -->\*\*..olha para o chãõ-->  
25   **LUI**                                 [o que é isso]//  
26                         (0.9)  
27   **LUI**           ma-mãe/ \*\*que isso [aqui ]  
28                         ----,,>\*\*..---olha p neide--->  
29   **NEI**                                 [é \*\*li:]\*\*xo)\ tira o pé/ #tá  
30   im   #im4  
31                         \*su\*jo::#\  
32   im                         #im5  
33   lui           ----->\*\*....\*\*olha para o chãõ-->\*  
34                         \*volta o torso p/ neide  
35                         \*(0.4)  
36   lui           \*...volta o torso p/ frente...---->  
37   **LUI**           [(nhêee)]  
38   **NEI**           [não po:]de por o PÉ#/ (.) no li:xo sujo\*  
39   lui           ----- ,,,,,,,,,,,,,,  
40   im                                 #im6 detetive.



imagem 6

Luiza inicia uma conversa introduzindo um pedido de informação sobre o objeto que percebera no chão (imagem 2): (*que isso aqui*//, linha 01) e (*é seu*//, linha 07). O atendimento a esse pedido de informação depende da visibilidade, por parte de Neide, do objeto a que se referem a forma demonstrativa [*isso*] e a forma dêitica espacial [*aqui*]. Nesse excerto, a forma dêitica espacial introduzida por [*aqui*] é antecedida por uma posição corporal de Luiza direcionada para o chão, local onde se encontra o objeto referido (linha 02). No momento em que Luiza formula sua pergunta, Neide está corporalmente voltada para a pia, realizando a atividade de lavar a louça. Uma ausência de fala de 1.2 segundos (linha 04) antecede um incremento (*é seu*//, linha 07) à primeira pergunta já introduzida (*que isso aqui*//, linha 01). Essa temporalidade dos gestos de Luiza é significativa. A trajetória de seus gestos corporais está articulada e sincronizada à construção multimodal de seu turno de fala e, portanto, ao formato que emerge

ali, contingenciado pelas ações que estão em curso. Para visualizarmos essa coordenação é importante seguirmos a representação da trajetória (a emergência, a sustentação e a retração) dos gestos Luiza (sua postura corporal, seu movimento com as mãos e seu olhar). Entre as linhas 1 e 9, vemos como Luiza sustenta o gesto de chamar a atenção de Neide com a mão e de manter seu olhar voltado para Neide durante a construção de seu turno de fala. Luiza mobiliza recursos multimodais que conferem, a suas ações para obter a atenção de Neide, uma sequência de turnos de fala que vai se incrementando à medida que Neide responde ou não, por direcionamento de olhar ou verbalmente, a suas chamadas. Assim, a construção do turno introdutório de Luiza se ajusta às posições corporais de sua co-participante. Identificamos no formato do turno (*qu'isso aqui*)// (1.0) + (0.2) é *seu*// (0.3) + (0.7) ) certos procedimentos como ausências de fala articulada com as ações não-verbais até que Neide oriente-se para o objeto



apontado. Luiza relança então seu pedido de atenção, dessa vez reformulando seu turno como um novo começo. Para isso, Luiza introduz um vocativo (*Ma-mãe*, linha 11), recurso verbal clássico para chamar a atenção do outro participante co-presente. Com isso, Luiza, verbalmente e de forma enfática, seleciona mais vez Neide como sua interlocutora e torna publicamente visível (audível), através dessas ênfases, que endereça a ela as perguntas feitas anteriormente sobre o objeto em questão. Note que Luiza já havia selecionado Neide, introduzindo inclusive uma forma dêitica pessoal (*é seu//*, linha 07). Apenas na segunda seleção, feita com o vocativo verbal, Neide responde verbalmente (*oi//*) e corporalmente, direcionando seu olhar para o chão, na direção onde está o objeto apontado por Luiza. Estabelecida a atenção visual e alinhadas, Luiza então reformula a pergunta feita anteriormente: (*I:sso aqui é seu//*).

Quando a solicitação de informação sobre o objeto é atendida por Neide (*[é li:]xo//*, linha 29), ambas iniciam um processo de fechamento dessa sequência interacional. Neide adverte que não se deve por o pé no lixo e Luiza volta seu corpo para a posição inicial (posição 1) em que estava antes de iniciar a sequência interacional, ou seja, antes de chamar a atenção de Neide para o objeto. A sequência de imagens 4, 5, 6 (linhas 29, 31, 38) temporalmente alinhadas ao turno de Neide, nos deixam ver a construção conjunta e multimodal do fechamento desta sequência interacional das ações de chamar a atenção de Neide e solicitar a esta uma informação sobre um objeto percebido no espaço em que estão. Quando essa ação se completa, Luiza retorna à posição 1, conforme a segunda sequência de imagens (imagens 4, 5, 6) inseridas no excerto e alinhadas temporalmente às produções de fala de Neide.

Os direcionamentos corporais de Luiza, ora voltada para Neide (posição 1), antes de iniciar seu turno de fala; ora voltada para o objeto (posição 3), enquanto inicia e desenvolve seu turno de fala; e ora voltando-se novamente para Neide (imagens 4, 5, 6), indicam como seus

movimentos corporais estão sincronicamente coordenados aos momentos de abertura e fechamento da sequência interacional da ação de pedir informação sobre o objeto em questão. Aberturas e fechamentos de sequências interacionais são ações estruturantes da interação social e demonstram uma competencial social e interacional de reconhecer e operar multimodalmente, com recursos corporais e materiais, na organização sequencial e estrutural da fala-em-interação. Procedimentos técnicos e metodológicos que nos permitam dar conta da temporalidade e da sequencialidade da interação permitem uma análise das produções verbais e não-verbais de crianças com TEA no curso e na organização da ação social. Atentar, a partir de uma perspectiva multimodal (língua, corpo e mundo material), para o curso temporal das ações, como iniciar e fechar uma conversa, pode revelar dimensões importantes para a inteligibilidade e propósitos da participação das crianças com TEA, bem como para competências que nos escapariam em uma abordagem da fala autista voltada apenas para as produções verbais individuais desses sujeitos, como apontara Sterponi *et al.* (2014).

Em um segundo excerto, gostaríamos de explorar um pouco mais as orientações corporais e de olhar de uma outra criança com TEA (Mia) para os objetos co-presentes e para ações de sua interlocutora (Eliana) e, com isso, analisar a correlação dessas orientações com sua produção verbal e com uma complexa relação semântico-discursiva que ela produz através da coordenação de distintos recursos semióticos. A interação se passa em uma atividade socioeducativa que acontece em uma instituição de convivência de crianças com TEA. Nessa instituição, há diferentes situações interativas: momentos livres e coletivos de convivência entre 12 crianças diagnosticadas com TEA, com características linguístico-interacionais e corportamentais muito variadas entre elas, e 04 educadoras; momentos de refeições conjuntas; momentos de atividades semi-orientadas propostas a pequenos subgrupos de 2 a 3 crianças.

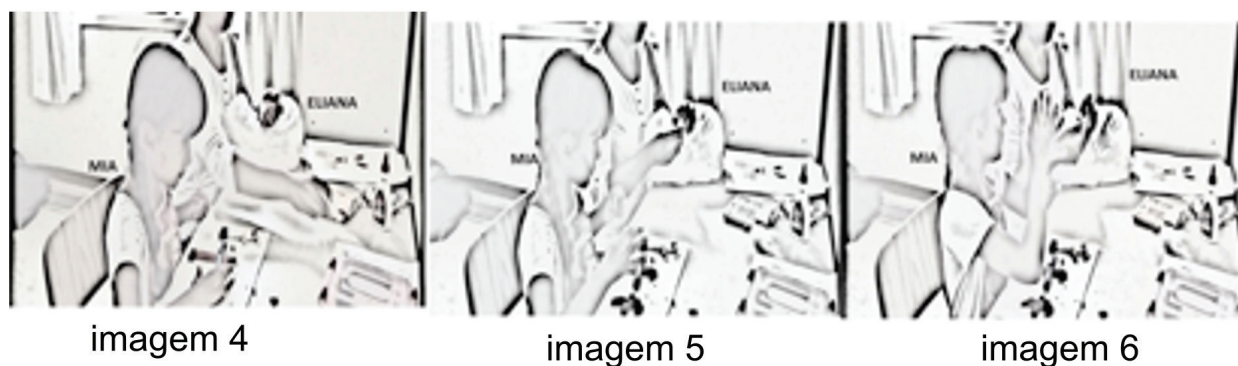


imagem 1

imagem 2

imagem 3

**Figura 3.** Sequência I: Posições corporais de Mia e Eliana.  
**Figure 3.** Sequence I: Body positions of Mia and Eliana.



**Figura 4.** Sequência II: Posições corporais de Mia e Eliana.

**Figure 4.** Sequence II: Body positions of Mia and Eliana.

Na situação em análise, Eliana, educadora, e Mia, criança autista atendida pela instituição, estão sentadas em cadeiras posicionadas uma ao lado da outra e diante de uma mesa. Mia tinha 7 anos de idade no período dos registros, entre março e dezembro de 2017. Em termos de perfil interacional, Mia faz pouco uso da linguagem verbal e mobiliza muitos recursos gestuais durante as interações na instituição. Mia faz um desenho com massa de modelar em uma folha de papel sobre a mesa. Simultaneamente, Eliana enrola pedaços de algodão e junta-os em um canto da mesa para, posteriormente, montar um coelho em uma folha de papel. Abaixo, apresentamos uma sequência de três imagens de capturas de tela do vídeo registrado (Sequência I, Figura 3). A sequência de imagens refere-se aos direcionamentos de olhar de Mia e Eliana relativos às duas atividades de enrolar que estão em curso, mencionadas acima.

As duas participantes realizam, simultaneamente, uma ação de enrolar e formar bolas (imagens 1, 2, 3). Cada uma delas mantém o direcionamento de olhar e as posturas corporais para sua própria atividade. Em uma segunda sequência de imagens (Sequência I, Figura 4), no momento imediatamente posterior à essa primeira sequência, Mia orienta-se, via direcionamento do olhar, para a sincronia de ações de enrolar (vide direcionamento do olhar de Mia nas imagens 4, 5, 6).

As sequências I e II correspondem ao momento imediatamente anterior ao momento em que Mia iniciará uma conversa (Excerto 2). Não há, nos momentos representados por essas sequências, nenhuma produção verbal. No entanto, esses momentos iniciais são fundamentais para visualizarmos como Mia faz referência a essa sincronia (imagem 6), para a qual Eliana não se orienta visualmente, introduzindo a forma lexical [*bola*], linha 1 (Excerto 2). Na transcrição, os gestos de Mia são indicados pelo símbolo (\*), os gestos de Eliana pelo símbolo ( $\Delta$ ) e as imagens alinhadas à transcrição são indicadas por (#).

Em cada uma das três vezes que Mia introduz o item lexical *bola*, notamos, seguindo as imagens (#7, #8, #9), como ela orienta seu corpo para objetos distintos que serão referentes de sua construção linguístico-discursiva. Em [*a bo:la*], (linha 01), Mia orienta-se para o desenho (#imagem 1). Em [*a bóia*], (linha 06), Mia orienta-se para a bola de massa de modelar em sua mão (#imagem 2). Em [*a bóia*], (linha 09), Mia orienta-se para a bola de algodão na mão de Eliana (#imagem 3). A forma como Mia estrutura multimodalmente seu turno de fala compõe uma unidade com coerência semântica-sintático-gestual (Kendon, 2004) a partir de uma espécie de economia funcional que integra aspectos linguísticos, ações corporais e espaço físico. Ao repetir uma série com três vezes o item lexical *bola*, coordenados, sequencial e temporalmente, aos direcionamentos de olhar e de corpo para, respectivamente, (i) seu desenho sobre a mesa; (ii) a bola de modelar em sua mão e, (iii) a bola de algodão na mão de Eliana, Mia introduz não apenas um referente [*bola*], mas uma atividade referencial mais complexa realizada multimodalmente. Mia parece referir-se à sincronia entre as duas de enrolar bolas (de algodão e de massa de modelar) que percebera, como demonstramos (capturamos) nas Figuras 3 e 4, através do direcionamento de olhar para as ações das mãos de Eliana e dela mesma.

Em um terceiro e último excerto, analisaremos uma situação em que Luiza (participante também no Excerto 1) e sua irmã Clara estão se preparando para assistir a um filme na televisão. Luiza está sentada no sofá, comendo pipoca. A televisão está ligada. Clara faz um movimento para ajustar uma coberta e se cobrir ao sentar-se (imagem 2). Neste momento, o movimento de Clara derruba no chão, sem que ela perceba, a pipoca que estava no canto do sofá em que Luiza está sentada (indicado na imagem 1, Figura 5, por um círculo desenhado). Luiza segue com o olhar a trajetória do movimento da coberta e alguns segundos depois Luiza percebe a ausência da pipoca (imagem 2, Figura 5). Após esse momento, Luiza inicia um turno

## Excerto 2: PAND\_cri\_17\_mar\_registro1d.eaf

```

01      MIA      #a ** bo:* ΔΔla/
02              ->** olha p/ o desenho ---->
03              -----> *
04      eli      ----->ΔΔ--olha p/ o desenho mia-->
05      im       #imagem 7
06              #bo:**ia/
07              ---->**olha p/ a massa de modelar---->
08      im       #imagem 8
09              bóΔΔia#\
10      eli      --ΔΔ olha p/ massa de modelar----->
11      im       #imagem 9

```



Figura 5. Sequência: Posições corporais de Clara e Luiza.

Figure 5. Sequence: Body positions of Clara and Luiza.

de fala direcionado à Clara contando o que aconteceu (Excerto 3). Clara, por sua vez, não se orienta para o fato de que foi seu movimento que derrubara a pipoca, mas propõe que quem deixou cair a pipoca foi Luiza.

Luiza inicia sua fala [tsa::u], olha para Clara e continua seu turno [pa ↑m/], fazendo dois tipos de gestos

para cada unidade lexical: (i) o de balançar a mão no ar e (ii) bater a mão no ar sincronizados à fala. Os enunciados linguísticos produzidos por Luiza, embora não contenham uma palavra ou unidade morfológica reconhecível do léxico em português, contém ordenamento sintático; pausa; fenômenos segmentais reconhecíveis como alongamento,



## Excerto 3: CELA004\_2015.eaf

```

01      LUI      *   tsa::u\#           *
02      lui      *-balança a mão -> *
03      im              #imagem 3
04              **(.)
05              **--olha p/ Clara sorrindo-->(1. 12)
06              *   #pa ↑m\  *
07              *bate a mão*
08      im              #imagem 4
09              (1.0)
10      CLA      @ <((rindo))você derrubou no **chão> //@
11      cla      @----->olha p/ o chão----->@
12      lui                               -->***--olha p/ o chão-->

```



imagem 1



imagem 2

entonação ascendente ou descendente; acentuação de ênfase tônica sobre a sílaba núcleo. Sincronicamente coordenado às produções verbais *tsau*, *pam*, Luiza realiza um conjunto de referências corporificadas como o olhar e a posição do corpo para o local onde estava a pipoca (imagem 2) e realiza os dois tipos de gestos, o gesto de balançar a mão e o gesto de bater a mão. Luiza não produz verbalmente o referente pipoca, mas constroi corporalmente a posição, na sequência sintática, do referente pipoca no turno construído multimodalmente. Assim, sugerimos que [*tsau*], seguido de gestos de lançar a mão no ar, funciona como o predicado e se refira à queda da pipoca e que a outra parte do enunciado, [*pam*], seguido do gesto de bater com a mão refira-se ao ato de bater (com o cobertor) na pipoca.

Os Excertos 2 e 3 parecem evidenciar que sem um olhar mais atento para a íntima correlação estrutural entre sistemas semióticos distintos (fala e gesto, por exemplo) e sua indexicalidade e sistematicidade envolvendo toda uma ecologia do espaço e dos objetos, muito provavelmente não seria possível atribuir significado a certas ações ou produções verbais. Uma das implicações dessa via analítica para uma abordagem dos gestos em situações interativas envolvendo patologias, alterações ou comprometimentos

linguístico-interacionais, é a de que os recursos corporais, gestuais e materiais podem ser concebidos não como compensatórios ou complementares a uma ausência ou limitação dos recursos linguísticos, mas integrados entre si, afinando-se a uma concepção radicalmente anti-logo-cêntrica de integração e de coordenação fina de práticas encarnadas e situadas, oferecendo à natureza constitutiva de nossas ações, uma dimensão multimodal primordial (Mondada, 2016). Nesse sentido, como sugere Levinson e Holler (2014), a própria concepção de multimodalidade humana pode ser vista como uma perspectiva mais ampla para se conceber as relações entre linguagem, cognição humana e mundo.

### Ainda algumas considerações finais

Procuramos neste artigo dar visibilidade aos procedimentos de uma investigação multimodal aplicada a interações envolvendo crianças com TEA. Consideramos que esse tipo de investigação possa oferecer um inventário de recursos multimodais mobilizados pelos participantes, autistas e não-autistas, durante as interações, permitindo explorar uma dimensão corporificada da interação social. Uma das implicações teóricas mais diretas que podemos



destacar para a análise dos recursos corporais mobilizados pelos sujeitos autistas e não-autistas, durante a construção de suas formas de interação, é que os recursos não-verbais podem ser concebidos não como compensatórios frente a uma falta ou limitação linguístico-verbal, mas constitutivos da interação humana. Assumir a integração de recursos multimodais, incluindo aqui a modalidade verbal como mais uma dentre as possíveis, pode contribuir para uma descrição das interações que dê conta da dinamicidade, da plasticidade, da adaptabilidade dos recursos multimodais (verbais ou não) à ecologia da interação e de uma possível sistematicidade (Mondada, 2014) de tais usos durante a construção das práticas intersubjetivas.

Consideramos que as análises dos dados propostas aqui corroboram as interpretações propostas por Sterponi *et al.* (2014) e Sterponi e Kirby (2016), que propõem tratar as características linguísticas de sujeitos com TEA, ainda que específicas, para além de uma manifestação subjacente a um déficit (Sterponi *et al.*, 2014, p. 8). Como apontam, uma perspectiva do déficit, que olha para as características atípicas da linguagem no TEA como inerentemente disfuncionais ou patológicas, pode ser revisada por uma perspectiva voltada para a funcionalidade da linguagem autista, com as características que elas apresentam de forma situada e contingente à situação interativa e ao contexto sociointeracional. Assim, para além de uma descrição de certas produções linguísticas consideradas, a partir de uma perspectiva do déficit, como exclusivamente atípicas ou alteradas, uma investigação da linguagem a partir de uma perspectiva multidimensional convida-nos a considerar o efeito do próprio contexto interacional e o papel dos interlocutores sobre a performance comunicativa das crianças com TEA. Essa perspectiva concebe as características linguísticas das falas das crianças com TEA não exclusivamente como sintomáticas da condição do indivíduo afetado, mas também produzidas por condições interacionais circuncanciadas e, em alguns casos, como vimos aqui, linguístico-interacionalmente funcionais.

Por fim, consideramos que, ao darmos um enfoque às produções verbais e gestuais circuncanciadas e situadas a determinadas condições interacionais, podemos ter subsídios e evidências para um aprofundamento sobre o engajamento social e construção conjunta e intersubjetiva de ações afetados nos quadros de TEA (Carpenter e Tomasello, 2000; Tomasello, 1999).

## Referências

- ASPERGER, H. 1991 [1944]. 'Autistic Psychopathy' in childhood. In: U. FRITH (org.), *Autism and Asperger Syndrome*. 13ª ed., Cambridge, Cambridge University Press, p. 37-62.
- AUBURN, T.; POLLOCK, C. 2013. Laughter and competence: Children with severe autism using laughter to joke and tease. In: P. GLENN; E. HOLT (orgs.), *Studies of Laughter in Interaction*. London, Bloomsbury, p. 135-160.
- CARPENTER, M.; TOMASELLO, M. 2000. Joint attention, cultural learning, and language acquisition: Implications for children with autism. In: A.M. WETHERBY; B.M. PRIZANT (eds.), *Communication and language issues in autism and pervasive developmental disorder: A transactional developmental perspective*. Baltimore, Brookes, p. 31-54.
- COTS, C.P. 2016. *Quando a linguagem verbal se ausenta, o que há?* São Paulo, SP. Relatório Final de Pesquisa de Iniciação Científica enviado à Fundação de Amparo à Pesquisa do Estado de São Paulo (processo 14/15206-6). Universidade Federal de São Paulo, 113 p.
- CRUZ, F.M.; COTS, C.P.; LUIZ, R. 2017. A linguagem em micro-acontecimentos: corpo, gestos e fala explorados em uma análise multimodal de interações envolvendo uma criança autista. *Revista Intercâmbio*, XXXIV:34-57.
- CUNHA, G.R.; BORDINI, D.; CAETANO, S.C. 2015. Autismo, trans-torno do espectro do autismo. In: S. CAETANO; M.C. LIMA-HERNANDES; F. PAULA; B.D. RESENDE; M. MÓDOLO (orgs.), *Autismo, Linguagem e Cognição*. Paco Editorial, p. 13-24.
- ERICKSON, F. 2010. The neglected listener: Issues of theory and practice in transcription from video in interaction analysis. In: J. STREECK (ed.), *New Adventures in language and interaction*. Amsterdam, John Benjamins, p. 243-256. <https://doi.org/10.1075/pbns.196.11eri>
- GOODWIN, C. 1981. *Conversational organization: Interaction between speakers and hearers*. New York, Academic Press, 195 p.
- GOODWIN, C. 2000. Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32(10):1489-1522. [https://doi.org/10.1016/S0378-2166\(99\)00096-X](https://doi.org/10.1016/S0378-2166(99)00096-X)
- GOODWIN, C. 1986. Gestures as a resource for the organization of mutual orientation. *Semiotica*, 62(1):29-49. <https://doi.org/10.1515/semi.1986.62.1-2.29>
- HEATH, C. 1986. *Body movement and speech in medical interaction*. Cambridge, Cambridge University Press, 200 p. <https://doi.org/10.1017/CBO9780511628221>
- KANNER, L. 1943. Autistic disturbances of affective contact. *Nervous Child*, 2:217-250.
- KENDON, A. 2004. *Gesture: Visible actions as utterance*. Cambridge, Cambridge University Press, 400 p. <https://doi.org/10.1017/CBO9780511807572>
- KNOBLAUCH, H.; SCHNETTLER, B.; RAAB, J.; SOEFFNER, H. (orgs.). 2006. *Video analysis methodology and methods: Qualitative audiovisual data analysis in sociology*. Berlin, Peter Lang, 217 p.
- KORKIAKANGAS, T.; RAE, J. 2014. The interactional use of eye-gaze in children with autism spectrum disorders. *Interaction Studies*, 15(2):233-259. <https://doi.org/10.1075/is.15.2.12kor>
- LAI, M.C.; LOMBARDO, M.; BARON-COHEN, S. 2014. Autism. *The Lancet*, 383:896-910. [https://doi.org/10.1016/S0140-6736\(13\)61539-1](https://doi.org/10.1016/S0140-6736(13)61539-1)
- LEVINSON S.C.; HOLLER J. 2014. The origin of human multi-modal communication. *Philosophical Transactions of the royal society B*, 369:20130302. <https://doi.org/10.1098/rstb.2013.0302>
- MONDADA, L. 2003. Observer les activités de la classe dans leur diversité: Choix méthodologiques et enjeux théoriques. In: J. PERERA; L. NUSSBAUM; M. MILIAN. (orgs.), *L'educacio linguistica en situacions multiculturals i multilingues*. Barcelona, Institut de Ciencia e de la educacion de la Universitat de Barcelona, p. 49-70.
- MONDADA, L. 2006. Video recording as the reflexive preservation and configuration of phenomenal features for analysis. In: H. KNOBLAUCH *et al.* (orgs.), *Video analysis methodology and methods: Qualitative audiovisual data analysis in sociology*. Berlin, Peter Lang, p. 51-68.
- MONDADA, L. 2008a. Documenter l'articulation des ressources multimodales dans le temps : La transcription d'enregistrements vidéos d'interactions. In: M. BILGER (org.), *Données orales: Les enjeux de la transcription*. Perpignan, Presses Universitaires de Perpignan, p. 127-156.
- MONDADA, L. 2008b. L'analyse de "collections" de phénomènes multimodaux en linguistique interactionnelle: à propos de l'organisation systématique des ressources gestuelles en début de tour. *Cahiers de praxématique*, 50:21-66.

- MONDADA, L. 2012. Organisation multimodale de la parole-en-interaction: pratiques incarnées d'introduction des référents. *Langue française*, **175**(3):129-147. <https://doi.org/10.3917/lf.175.0129>
- MONDADA, L. 2014. Conventions for multimodal transcription, versão 3.0.1. Disponível em: [https://franz.unibas.ch/fileadmin/franz/user\\_upload/redaktion/Mondada\\_conv\\_multimodality.pdf](https://franz.unibas.ch/fileadmin/franz/user_upload/redaktion/Mondada_conv_multimodality.pdf). Acesso em: 02/08/2017.
- MONDADA, L. 2016. Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, **20**(30):336–366. [https://doi.org/10.1111/josl.1\\_12177](https://doi.org/10.1111/josl.1_12177)
- NEVILE, M. 2015. The embodied turn in research on language and social interaction. *Research on Language and Social Interaction*, **48**(2):121-151. <https://doi.org/10.1080/08351813.2015.1025499>
- OCHS, E.; SOLOMON, O. 2010. Autistic Sociality. *ETHOS*, **38**(1):69-92. <https://doi.org/10.1111/j.1548-1352.2009.01082.x>
- OCHS, E.; SOLOMON, O.; STERPONI, L. 2005. Limitations and transformations of habitus in child-directed communication. *Discourse Studies*, **7**(4):547-583. <https://doi.org/10.1177/1461445605054406>
- SOLOMON, O. 2008. Autism, language and childhood: An ethnographic perspective. *Annual Review for Applied Linguistics, Special Issue Language and the Brain*, **28**:150–169. <https://doi.org/10.1017/S0267190508080148>
- STERPONI, L.; KIRBY, K. 2016. A multidimensional reappraisal of language in autism: Insights from a discourse analytic study. *Journal of autism and developmental disorders*, **46**(2):394-405. <https://doi.org/10.1007/s10803-015-2679-z>
- STERPONI, L.; KIRBY, K.; SHANKEY, J. 2014. Rethinking language in autism. *Autism*, **19**(5):517-526. <https://doi.org/10.1177/1362361314537125>
- STERPONI, L.; SHANKEY, J. 2014. Rethinking echolalia: Repetition as interactional resource in the communication of a child with autism. *Journal of child language*, **41**(2):275-304. <https://doi.org/10.1017/S0305000912000682>
- STREECK, J.; GOODWIN, C.; LeBARON, C. (eds.). 2011. *Embodied Interaction: Language and body in the material world*. New York, Cambridge University Press, 308 p.
- TAMANAH A.C.; PERISSINOTO, J. 2015. Avaliação de linguagem nos transtornos do espectro do autismo. In: S. CAETANO; M.C. LIMA-HERNANDES; F. PAULA; B.D. RESENDE; M. MÓDOLO (orgs.), *Autismo, Linguagem e Cognição*. Jundiaí, Paco Editorial, p. 25-28.
- TOMASELLO, M. 1999. *The cultural origins of human cognition*. Cambridge, Harvard University Press, 256 p.
- WALENSKI, M.; TAGER-FLUSBERG, H.; ULLMAN, M.T. 2006. Language in autism. In: S.O. MOLDIN; J.L.R. RUBENSTEIN (eds.), *Understanding autism: From basic neuroscience to treatment*. London, Taylor & Francis Books, p. 175-203.
- WITTENBURG, P.; BRUGMAN, H.; RUSSEL, A.; KLASSMANN, A.; SLOETJES, H. 2006. ELAN: A professional framework for multimodality research. In: International Conference on Language Resources and Evaluation, 5<sup>th</sup>, Wittenburg, 2006. *Proceedings...* **3**:1556-1559.

Submetido: 13/08/2017

Aceito: 26/04/2018

## Anexo

**Quadro 1.** Convenção de transcrição multimodal baseada em Mondada (2014).

**Chart 1.** Conventions for multimodal transcription based on Mondada (2014).

Aspectos linguístico-verbais	
Xxx	Segmento ininteligível.
.	(Micro pausas, inferiores a 0,3 segundos, não medidas).
(0.4)	Pausas medidas com ajuda do software ELAN versão 5.0.
:	Alongamento silábico.
.h	Marca a inspiração do locutor.
/ e \	Entonação ascendente e entonação descendente.
//	Entonação ascendente acentuada (de pergunta).
<u>Sublinhado</u>	Ênfase particular (intensidade, acento).
maÍÚSCULA	Volume forte de voz.
o o	Volume baixo, murmúrio de voz.
Trajetória dos gestos	
+----+	Delimitação do início e o fim da ação/gesto.
+, *	A fala transcrita comporta os símbolos gráficos indicadores de gestos (+, *, por exemplo) posicionados no momento em que são realizados com relação à fala.
(1.0) +* (0.2)	Se um gesto começa no meio de uma pausa, segmenta-se a pausa (ex. 1.2 segundos = 1,0 + 0,2) a depender do momento de início ou fim do gesto e de sua duração realizados durante a ausência de fala.
---->121	Continuação do gesto até a linha 21 (exemplo) do excerto.
---->+	Se um gesto continua nas linhas seguintes, sua descrição é seguida de uma flecha que remete ao símbolo que delimita o seu fim.
---->>	Continuação do gesto até o fim do excerto.
>>	O gesto descrito começa antes do início do excerto.
Trajetória temporal dos gestos	
.....	Emergência do gesto (preparação).
-----	Fim do gesto (retração).
»»»»»»»»	Manutenção do gesto (núcleo/ápice).