

# Interação situada em foco: um estudo de caso

## *Situated Interaction in Focus: a Case Study*

Gabriella Souza Oliveira

Universidade de São Paulo (USP)

São Paulo | SP | BR

gabriellaoliveira@usp.br

<https://orcid.org/0000-0002-4604-598X>

**Resumo:** Este estudo explora a ideia de língua como um processo que se estabelece e se desenvolve momento a momento entre os interlocutores de uma interação situada. Assumindo uma perspectiva crítica em relação a teorias linguísticas hegemônicas que priorizam língua como produto, esta investigação tem como objetivo aproximar a análise linguística dos estudos do corporeamento, ou ainda, do entendimento de que linguagem e cognição não podem ser consideradas à parte do corpo. Para tanto, analisou-se um trecho de uma entrevista gravada ao vivo entre a jornalista Leda Nagle e o cantor Tim Maia, retirada do programa de TV “Sem Censura” e disponível em ambiente *online*. A análise foi desenvolvida utilizando-se o *software* ELAN (Wittenburg *et al.*, 2006) e o resultado dessa exploração ilustra o modo como o processo interacional é o meio fundamental pelo qual a linguagem humana se organiza.

**Palavras-chave:** interação situada; corporeamento; linguagem humana.

**Abstract:** This study explores the idea of language as a process that is established and is developed from moment to moment between the interactants in a situated interaction. Assuming a critical perspective over the hegemonic linguistic theories that prioritize language as a product, this investigation aims to approximate the linguistic analysis to the embodiment studies, or, to the understanding that language and cognition cannot be considered apart from the body. For that, an excerpt was analysed from a recorded live interview between the journalist Leda Nagle and the singer Tim Maia, taken from the TV program *Sem Censura*, and available online. The analysis was developed by the use of the software



ELAN (Wittenburg *et al.*, 2006) and the result of this exploration illustrates how the interactional process is the fundamental way through which the human language is organised.

**Keywords:** situated interaction; embodiment; human language.

## 1 Introdução

Do início do entendimento da Linguística enquanto ciência, no começo do século XX, aos dias atuais, uma das preocupações da tradição linguística hegemônica era a de ser uma área, em princípio, teorizadora, em que língua é entendida como um construto abstrato, muitas vezes, até mesmo desvinculada dos falantes que a usam. Ainda que algumas áreas se voltem para a língua em uso, como a pragmática ou sociolinguística, vemos que muitos estudos priorizam investigar o produto linguístico, e não necessariamente o processo semiótico por meio do qual a língua ocorre.

McCleary *et al.* (2015, 2017) advertem que, surpreendentemente, até mesmo vertentes da Linguística Cognitiva que se fundamentam no uso parecem evitar contemplar uma visão mais dinâmica da linguagem, muitas vezes entendendo que sentido é algo que se dá na mente individual de cada falante, e não como um processo emergente da interação. Ainda no século XX, uma exceção pode ser feita para Sacks, Schegloff, Jefferson que, em 1974, publicam, na revista *Language*, um artigo que investiga os detalhes da conversação da vida real. Tendo como intuito descobrir a organização da conversa por meio da qual a linguagem se manifesta, esse estudo não apenas se torna o pioneiro em analisar língua como prática, ação, ao esmiuçar a maneira como a fala é realizada em tempo real, como também introduz a área hoje consolidada como análise da conversa.

Essa área cresceu para além da sua fase inicial, que inicialmente priorizava a comunicação verbal humana oralizada, e desde então, contribuiu para que, em outras investigações, língua fosse entendida não apenas como um sistema abstrato, mas como uma prática comunicativa, enquadrando a comunicação em seu sentido mais amplo. Nesses estudos, estão os estudos dos gestos (McNeill, 1992), da comunicação multimodal co-presencial<sup>1</sup> (Mondada 2016; Goodwin, 2018) e da interação de humanos com não-humanos, como animais domésticos (Amha, 2013) ou ainda, mais recentemente, com máquinas (cf. Rasenberg *et al.*, 2023). Contudo, se, por um lado, esses estudos ainda não são priorizados na Linguística hegemônica, por outro, vem se consolidando a ideia de que não se pode pensar língua sem interação.

<sup>1</sup> Tanto Goodwin (2018) como Ameka e Terkourafi (2019) nos alertam para o enviesamento do termo “face-a-face”, uma vez que uma interação comunicativa pode ocorrer por telefone, por intermédio de uma máquina, ou mesmo quando os interlocutores estão dispersos em um ambiente em que estão visualmente inacessíveis uns para os outros. Ameka e Terkourafi (2019) também comentam que, em muitas culturas, as pessoas não se posicionam face-a-face nem se olham nos olhos para se comunicar. Por essas razões, este trabalho adotará a sugestão de Ameka e Terkourafi (2019), preferindo o uso da expressão “comunicação co-presencial”.

A maneira hegemônica mecanicista e abstrata de conceber língua, no entanto, tem raízes um pouco mais profundas. A constituição da Linguística enquanto ciência e seus desdobramentos no século que se seguiu é ancorada em uma forte perspectiva ocidentecentrada (Ameka; Terkourafi, 2019; Ngué Um, 2020) que se fundamenta na emancipação da Era Moderna (cf. Toulmin, 1990), em que são priorizados a mente, a razão, a certeza, as generalizações, e o rigor lógico. Não coincidentemente, alguns fenômenos linguísticos passaram a ser caracterizados como altamente específicos entre línguas e relegados à margem dos estudos, entendidos como a “epítome da ‘anti-generalização’”<sup>2</sup> (Joseph, 1997, p. 93), como é o caso de interjeições ou ideofones, que muitas vezes sequer são considerados na análise linguística (Ameka, 2001, 2020; Dingemanse, 2017, 2018).

Algumas propostas, em vias de contornar esse ocidentecentrismo pervasivo na análise linguística, estão dispostas a questionar a abordagem dos modelos linguísticos mais tradicionais. Uma delas é a pragmática emancipatória (Hanks *et al.*, 2009) que se fundamenta no entendimento de que as diferentes línguas existem somente enquanto “uma modalidade de ação” (Hanks *et al.*, 2009, p. 3) e que é preciso descrevê-las “nos seus próprios termos” (Hanks *et al.*, 2009, p. 2). Nessa perspectiva, as línguas não devem ser concebidas como meras abstrações ou como um fenômeno monolítico, *a priori*, pré-definido e individual, mas, sim, como um processo contínuo, inerentemente diverso, coletivamente construído e historicamente situado. Anchimbe e Janney (2011) e Ameka e Terkourafi (2019) também chamam a atenção para uma mudança de perspectiva em que, nas investigações, seja centralizada a prática linguística situada, para que se faça uma análise linguística mais justa e sofisticada. Os *insights* provocados por essas pesquisas ilustram que não estudar o uso linguístico em contexto pode enviesar a pesquisa e constitui uma limitação ao entendimento do que é a linguagem humana.

No início desta década, Dingemase *et al.* (2023) publicam um manifesto intitulado *Beyond single-mindedness*, em que reiteram a proposta de recentralizar a interação nos estudos das Ciências Cognitivas e da Linguística, entendendo a interação como nosso “tecido conectivo”: é somente nela que língua e mente podem existir. E criticam essa tendência de ambas as áreas, em geral, terem tomado um caminho que privilegia não apenas o produto sobre o processo, mas a informação sobre as relações, os indivíduos sobre as interações; e, como resultado, isso obscurece o modo como a interação co-constitui a cognição humana.

Este artigo, por sua vez, tem como objetivo fazer uma análise de um trecho de uma interação situada entre interlocutores de uma conversa co-presencial. O intuito é o de aproximar a análise linguística dos estudos da corporeidade, ou, da perspectiva de que linguagem e cognição não podem ser consideradas à parte do corpo, tema esse central em outras áreas das Ciências Cognitivas, como na fenomenologia e nas teorias enativistas.

Partindo da ideia de que a cada minuto da nossa vida estamos constantemente (re) construindo o mundo em que vivemos na interação com o ambiente e com o(s) outro(s) e que as nossas ações se co-constituem mutualmente (Enfield, 2013; Goodwin, 2018), ao examinar o papel sequencial e contingencial da interação humana, e demonstrar o modo como a significação emerge desse processo, é possível expandir os limites da investigação linguística e descortinar novas possibilidades de concepção da natureza e do funcionamento das línguas e da cognição humana.

<sup>2</sup> Não surpreende que o objeto inicial da investigação linguística, a *langue*, era concebida como tendo propriedades homogêneas, estáveis e regulares (cf. Saussure, 2012), tomando como ponto de partida o estudo das línguas indo-europeias.

## 2 Fundamentos da análise

Quando duas ou mais pessoas se encontram, presencialmente, em carne e osso, esse encontro é inescapavelmente social e inescapavelmente corporal. Quer elas usem ou não a língua verbal, elas, primeiramente, e durante toda a interação, se comunicam, observando os movimentos umas das outras: como orientam seus corpos, para onde olham, como movimentam seus braços e mãos, quais expressões aparecem em suas faces. Essa é a condição *sine qua non* da interação social e da sociabilidade, e a base para a emergência de toda a tecnologia desenvolvida para servir à comunicação humana, inclusive a própria língua.

(McCleary; Viotti, 2017, p. 21)

A citação acima ilustra a questão central do que é chamado de ‘semiótica de corpos em ação’, o fato de qualquer encontro entre duas ou mais pessoas ser um encontro de corpos que se percebem mutualmente e estão envolvidos em um processo contínuo de construção de significação nas interações cotidianas, situadas em um tempo e em um lugar específicos; “em ocasiões de enunciação, em sua plenitude física, cognitiva e social” (McCleary; Viotti, 2017, p. 3). Isso quer dizer que tanto a cognição quanto a língua que usamos não são independentes do nosso corpo, especialmente de nossas capacidades perceptuais e motoras (ver Varela *et al.*, 1991; Mondada, 2016; Goodwin, 2018; Dingemanse *et al.*, 2023; Levinson, 2025 e outros): nossas conceitualizações estão intrinsecamente ancoradas no sistema sensório-motor, que é a fonte de nossas percepções, de nossa movimentação e de nossa forma de lidar com o mundo. Nossa conhecimento, portanto, depende do fato de sermos e estarmos em um mundo que é inseparável do nosso corpo, da nossa linguagem e da nossa história social – do nosso *corporeamento* (cf. Varela *et al.*, 1991). Essa concepção é consonante com a perspectiva de conceitualizar “a cognição não como um domínio de mentes singulares, mas como uma conquista interacional de agentes corporeados” (Dingemanse *et al.*, 2023, p. 2, tradução própria), visão que se fundamenta na ideia de *Umwelten*, inicialmente proposta pelo etologista von Uexküll nos anos 1920s (cf. Favareau, 2010; Rasemberg *et al.*, 2023). Nessa perspectiva, não existe um único mundo “objetivo”, mas diferentes mundos, habitados por diversos organismos, que constantemente se sobrepõem e se co-constituem.

Outra proposta que oferece fundamentação para a dinâmica de como nós, enquanto agentes corporeados (e não apenas mentes isoladas), estamos sempre situados em um ambiente e somos co-constituídos pelos sucessivos encontros co-presenciais uns com os outros, vem da fenomenologia, que, há muito, se debruça sobre a questão da intersubjetividade. Segundo essa perspectiva, nós nos encontramos intersubjetivamente interagindo com esses outros e os entendendo de modo contextualmente socializado. Para Schütz (1967), proeminente filósofo da área, os “movimentos corpóreos de uma [determinada] pessoa são percebidos por outra não somente como eventos físicos, mas também como um sinal de que esta outra pessoa está tendo uma experiência da experiência a qual [aquela primeira] pessoa está experienciando através desses movimentos” (Schütz, 1967, p. 101, tradução própria). Nossa acesso aos outros aconteceria porque há uma inter-relação entre corpos, os nossos e os dos outros, mediadas pelas percepções que temos das ações contingenciais.

A construção da ação em um estado de co-presença mútua, portanto, é o espaço central para a constituição dinâmica e processual da sociabilidade humana (Goodwin, 2018, p. 248); é o lugar em que uma série de fenômenos, incluindo a língua e outras formas de semiose entrelaçadas a ela, emergem no mundo social real. Schütz atrela este estado de co-presença mútua à *we-relationship* (Schütz, 1967, p. 115). Enquanto compartilhamos um espaço e um tempo com os outros, cuja “experiência está fluindo lado a lado com a minha”, nossas faixas de consciência se interseccionam (Schütz, 1967, p. 163). Tal presença simultânea cria uma experiência especial de tempo bem diferente do tempo divisível e quantificável proposta por algumas correntes Física. Ele vai dizer: “eu e você... temos uma genuína *durée* que se experien-*cia*, que é contínua, que é múltipla e que é irreversível”. Nós experienciamos “o fenômeno de ficarmos velhos juntos” (Schütz, 1967, p. 103, tradução própria). Isso implica que o eu e o outro – o nós, ou *we* – estamos, de uma maneira específica, “simultâneos”, que nós “coexistimos” e que nossas respectivas faixas de consciência se intersectam momento a momento na interação. Para Schütz, portanto, a intersubjetividade e essa “relação-*we*” (*we-relationship*) são a base de todas as categorias da existência humana.

Nessa concepção, o tempo e espaço compartilhados através da *we-relationship* e que possibilitam o estabelecimento da ação intersubjetiva não é algo alcançado somente pela co-presença dos participantes de qualquer interação, mas é um empreendimento mutualmente contínuo, uma vez que não seria possível que alguém saiba com precisão a experiência subjetiva do outro (Schütz, 1967, p. 99). Dessa forma, o sentido pretendido de uma ação é sempre, em princípio, subjetivo e acessível apenas ao agente daquela ação (Schütz, 1967, p. 115) e os participantes precisam agir mutualmente, co-operativamente, para negociarem a inteligibilidade da interação. Outros estudos sugerem que, na interação co-presencial, portanto, os seres humanos mobilizam ordenadamente e situadamente uma grande variedade de recursos verbais, auditórios, visuais, ou ainda, multimodais, com o intuito tanto de produzir ações inteligíveis como o de interpretar publicamente ações mútuas visualmente disponíveis (Mondada, 2013, p. 219). Goodwin (2018) também acrescenta que formas específicas de presença vão garantir a abertura necessária para que determinado fenômeno emerja por meio dos arranjos corporais co-presentes na interação.

Isso somente é possível devido ao fato de que, ao interagirem, os co-participantes de uma interação criam uma arena, um terreno comum (*common ground*), para compartilhar e calibrar suas experiências. Um lugar onde fluxos separados dessa consciência interceptada são tornados públicos, por meio de formas de semiose que abrangem tanto a língua falada quanto as capacidades expressivas dos corpos individuais (Goodwin, 2018, p. 147). É nesse espaço em que é possibilitado a cada interlocutor performar as transformações acumulativas das ações uns dos outros, de modo *online*, emergente e processual, por meio de formas semióticas que expressam a experiência corporeada das suas próprias percepções e que levam em consideração tanto as respostas dos co-participantes como as contingências do contexto interacional. Sem esse terreno comum intersubjetivo, não há significação.

## 2.2 A ação co-operativa

Partir do princípio de análise de se investigar a ação co-operativa de corpos em interação significa estudar como os humanos criam ação e compartilham conhecimento em consonância com outros, ao reusarem, com transformação, recursos obtidos previamente de outros participantes e de outros eventos com os quais interagiram. Nas palavras de Goodwin (2018, p. 11, tradução própria): “nós habitamos reciprocamente as ações de uns e outros”, assim como outrora sugeriu Schütz. Nessa perspectiva, a habilidade humana de construir sentido e ação em consonância com outros não se dá em algum lugar anterior – *a priori* – e não é constituída em modalidade única; pelo contrário, ela reúne estruturalmente diferentes tipos de recursos semióticos de construção de sentido que, ao mesmo tempo, se elaboram, se limitam (Goodwin, 2018, p. 336) e se co-constituem.

O hífen em *ação co-operativa* é intencional e serve para ilustrar, primeiramente, que essa co-constituição é diferente de cooperação, como se estuda na antropologia biológica, uma vez que não se restringe à ajuda mútua. Em segundo lugar, ele enfatiza a importância das performances das operações específicas, de decomposição e reúso, com transformação, dos materiais providenciados pelos participantes durante a interação (Goodwin, 2018, p. 6). Mais crucialmente, a ação co-operativa fornece, no meio da própria ação, um mecanismo sistemático de acumulação progressiva (com modificações) em todas as escalas de organização social, desde os gestos, da fala até a diferenciação de grupos sociais. O que está sendo focalizado são as práticas sociais públicas que seres humanos mobilizam, de modo pervasivo, para construir, em consonância com outros, as ações que permitem ser possíveis e que sustentam suas atividades.

Dessarte, entender a constituição da organização humana, e desse modo, a própria interação comunicativa como uma forma de ação co-operativa, nos permite nos atentarmos para o local onde fundamentalmente se localiza a emergência e o desenvolvimento de língua, entendida como parte da ação humana. A realização da ação inclui, na sua organização, não apenas os materiais criados por aqueles que os usaram anteriormente, como também a transformação desses materiais de modo preciso para a adaptação necessária às circunstâncias do momento. Isso só é possível dado que cada participante não apenas está, mas ativamente participa, da organização detalhada da ação de cada um, que se desenrola no tempo (Goodwin, 2018, p. 9), e por isso, os participantes da interação não são apenas participantes, mas sim, co-participantes, ou mesmo, interactantes.

A seguir, examinemos um trecho de uma interação gravada em vídeo de uma conversa em tempo real que ocorre durante a entrevista entre a jornalista Leda Nagle e o cantor Tim Maia.

## 3 A interação

O vídeo escolhido para análise foi um trecho da entrevista do programa *Sem Censura* da jornalista Leda Nagle, com Tim Maia que ocorreu em fevereiro de 1997. A entrevista está disponível no canal Leda Nagle no *YouTube*<sup>3</sup>. Do bloco de 20:28 minutos, disponibilizado no canal, foi

<sup>3</sup> O vídeo na íntegra pode ser acessado em: <https://www.youtube.com/watch?v=HgIldtL6jc8>. A plataforma garante os direitos autorais do vídeo.

selecionado um trecho de aproximadamente 2 minutos, em que Tim Maia conta para os jornalistas que está sendo processado por músicos, interação que será explorada mais adiante. O recorte especificamente nesse trecho se dá, porque nele aparece um ideofone, um dos fenômenos muitas vezes relegados às margens dos estudos linguísticos. Ao centralizar a emergência desse fenômeno, é possível observar a maneira como ele é fundamental para o entendimento do desenrolar da interação, e, portanto, essencial na criação de significação da interação.

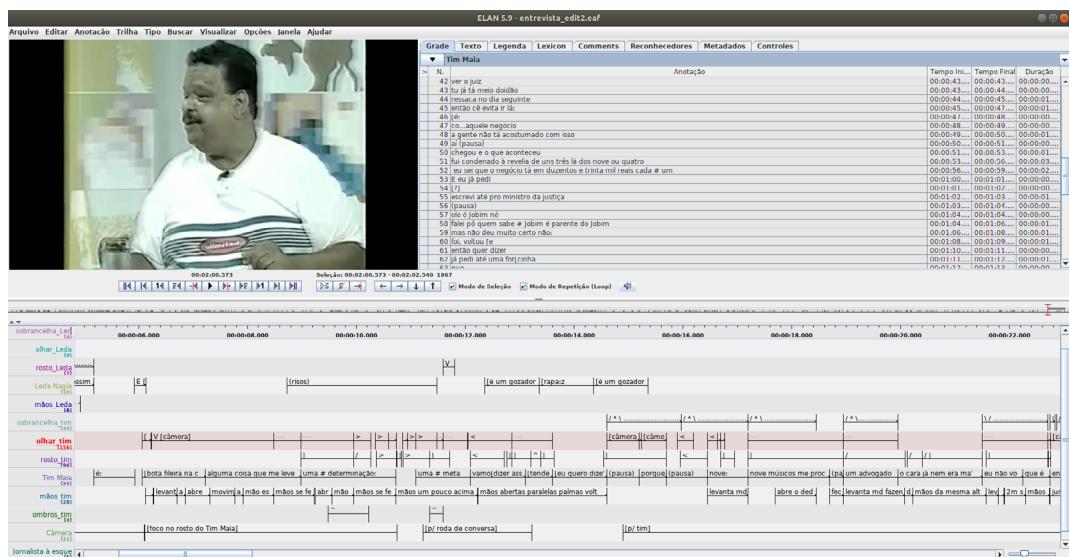
Essa seleção passou por uma descrição minuciosa no *software* ELAN (Wittenburg *et al.*, 2006). ELAN é um programa de código aberto de anotação de áudio e gravação de vídeo, criado e desenvolvido pelo *Max Planck Institute for Psycholinguistics* especificamente para descrição e análise linguísticas detalhadas. Ele permite a criação de trilhas ilimitadas para anotação multimodal, e a sua escolha se deu pelo fato de que nele é possível separar as trilhas de descrição concomitantemente às imagens do vídeo.

Embasando-se no trabalho de McCleary e Viotti (2007), portanto, foram destrinchadas 13 trilhas, cada qual destinada a um participante e/ou modalidade da interação. Elas foram criadas com o objetivo de capturar o maior número possível de detalhes da corporeidade dos participantes em interação, como forma de serem “lentes poderosas para identificar várias características das línguas que, sem elas, poderiam passar despercebidas” (McCleary; Viotti, 2007, p. 25). O sistema de transcrição, por sua vez, seguiu o modelo sugerido por Goodwin (2018, p. 18); mas, diferentemente desse modelo, aqui serão usadas setas (↑) e (↓) para indicar entoação ascendente e descendente, respectivamente. A ortografia segue a convenção tradicional, e vírgulas não foram utilizadas. A referência da notação pode ser verificada ao final do texto.

É preciso notar que, além dos itens da pauta sugeridos em McCleary e Viotti (2007), foram acrescentadas outras duas modalidades: movimento do rosto para os participantes Leda Nagle e Tim Maia, e movimento dos ombros para o participante Tim Maia. Essa trilha destinada somente para a descrição do movimento dos ombros de Tim Maia, e não de Leda ou dos outros jornalistas, se justifica já que em boa parte do tempo as câmeras se focam apenas no cantor. Além disso, como se trata de um programa jornalístico e televisivo, também foi acrescentada uma trilha para o movimento da câmera, uma vez que ela ora se foca na jornalista Leda Nagle, ora em Tim Maia, ou ainda, se movimenta de modo panorâmico na roda de conversa<sup>4</sup>. Ainda, foi destinada uma trilha para um jornalista não identificado à esquerda de Tim Maia, e, dado que a câmera em nenhum momento o enfoca isoladamente, esta trilha foi usada somente para a anotação de sua voz, que aparece poucas vezes. Desse modo, as treze trilhas usadas para a anotação da interação foram: (i) sobrancelha da jornalista Leda Nagle, (ii) movimento dos olhos de Leda Nagle, (iii) movimento do rosto de Leda Nagle, (iv) voz de Leda Nagle, (v) movimento das mãos de Leda Nagle, (vi) sobrancelha de Tim Maia, (vii) movimento dos olhos de Tim Maia, (viii) movimento do rosto de Tim Maia, (ix) voz de Tim Maia, (x) movimento das mãos de Tim Maia, (xi) movimento dos ombros de Tim Maia, (xii) movimento da câmera, e, por fim, (xiii) voz do jornalista à esquerda, como pode ser observado na captura de tela na imagem (1):

<sup>4</sup> É preciso fazer uma ressalva sobre este ponto: esse formato jornalístico limitou a visão de todos os participantes da interação simultaneamente, e, por isso, a anotação feita se deu, majoritariamente, sobre ações de Tim Maia. Por outro lado, a não-presença de um pesquisador na cena pode ter contribuído para a espontaneidade da conversa.

Imagen 1 – Captura de tela do processo de anotação da interação multimodal no programa ELAN



Fonte: Elaborado pela autora.

Ainda, é preciso lembrar que a descrição apresentada, em texto, no papel, é um projeto de captura das performances que estão sendo experienciadas e realizadas pelos falantes em tempo real, naquela interação e naquele lugar, que, por sua vez, são dinâmicas, simultâneas, multimodais, corporeadas e que se desenrolam contingencialmente momento a momento; sendo assim, essa captura representa um estado, um “congelamento” em dimensões físicas limitadas (espaço bidimensional do papel) e uma modalidade (a escrita) das ações performadas pelos participantes da interação.

### 3.1 A história contada por Tim Maia

A interação que se segue é organizada em forma de semicírculo; a jornalista anfitriã do programa, Leda Nagle, se situa em uma posição no centro desse semicírculo, possuindo uma visão panorâmica dos participantes do programa sentados em roda. À sua frente, no meio do semicírculo está Tim Maia, cercado de outros jornalistas, e à frente dele há uma bancada que delimita a estrutura desta meia-rosa.

Imagen 2 – Captura de tela da organização dos interlocutores da conversa



Fonte: Canal Leda Nagle, YouTube.

Tim Maia conta a história de que está sendo processado por nove músicos. Conta que havia dias em que os julgamentos ocorriam no mesmo dia, um de dia e dois de tarde e que, às vezes não comparecia a eles. Justifica dizendo que eram muitos processos e que não estava acostumado com isso, com o fato de ter que “ir ver o juiz”; ele acrescenta que, na época, ainda bebia uísque e ficava meio “doidão”, de ressaca, e que “então, você evita ir lá”. Consequentemente, foi condenado à revelia, ou seja, condenado novamente pela recusa ou não comparência em juízo, “de três ou quatro”, como diz, dos nove processos em andamento e que agora está sendo autuado em duzentos e trinta mil reais por cada processo. Ele diz que tentou recorrer, escrevendo até para o Ministro da Justiça, mas que até então nada foi feito. Tim Maia repete os valores, “duzentos e trinta mil cada processo”, e conta que, por ser um valor alto, quando chega ao tribunal, todos ficam tumultuados, como abelha: nesse momento, faz um som, zzzz, alterando a qualidade da voz, enquanto simultaneamente movimenta a cabeça para baixo, abaixa a frequência da voz, levanta os braços e balança as mãos repetidamente na horizontal concomitantemente ao som produzido.

IMAGEM 3 – Quatro capturas de tela de Tim Maia contando um trecho da história



Fonte: Canal Leda Nagle, *YouTube*.

Na primeira imagem, acima à esquerda, Tim olha para a frente, para a jornalista Leda Nagle; na segunda, acima à direita, ele abaixa os olhos e começa a narrar o modo como todos ficam no Tribunal, tumultuados, ao mesmo tempo em que levanta as mãos, a mão esquerda na altura do ombro esquerdo e a mão direita mais acima do ombro direito; na terceira imagem abaixou à esquerda, o cantor inicia uma comparação da percepção de tumulto com a abelha, e, na quarta imagem, à direita abaixou, ele performa a cena, fazendo uma depicção do som da abelha enquanto levanta os braços e balança as mãos concomitantemente ao som produzido.

Então, em seguida, a jornalista Leda Nagle pergunta a Tim Maia se ele está pensando em trabalhar mais para poder pagar esse valor, em que ele responde que sim, já que, de acordo com ele, a mídia não valoriza os cantores. Diz que músicos que trabalham com ele há bastante tempo, há dezesseis anos, nunca o processaram antes e que são músicos novos que estão o processando; por fim, repete que são muitos processos contra ele e logo em seguida muda de assunto.

## 4 Discussão

O trecho escolhido começa com a jornalista Leda perguntando diretamente para Tim se ele está mais empenhado na sua carreira. A jornalista faz essa pergunta com o corpo e o olhar também direcionado para Tim, acompanhado de um movimento gestual com as mãos fechadas em frente ao seu corpo, e de uma expressão facial, em que as suas sobrancelhas se levantam, como pode ser visto em (1):

(1)

- 1 Leda: ↑ ô Tim você está levando↓(0.3) você ,está mais empenhado na sua ^carrei:ra  
digamos assim↓
- 2 Mão direita de L: (para cima e para baixo)
- 3 Mão direita de L: (para cima)
- 4 Tim: (0.3) é: (0.36)
- 5 Leda: ↑ é isso ↓
- 6 Tim: [fileira na carreira alguma coisa que me leve a algum lugar
- 7 I: (h)
- 8 Mão de T: (palmas abertas paralelas à frente de seu corpo)
- 9 Tim: (0.44) uma [determinação
- 10 I: (h)
- 11 Mão de T: (se movimentam para frente)
- 12 Cabeça de T: (direita)
- 13 Tim: (0.68) ou uma meta vamos [dizer assim
- 14 Cabeça de T: (esquerda)
- 15 Leda: [é um gozador
- 16 Tim: [esse rapaz↑  
entendeu
- 17 Leda: [é um gozad-
- 18 Tim: [mas eu quero dizer o seguinte
- 19 Tim: (0.75) *porque* (0.73) nove nove músicos me processaram

Imagen 4 – Captura de tela da jornalista Leda Nagle no momento em que pergunta a Tim Maia se ele está mais empanhado na sua carreira



Fonte: Canal Leda Nagle, *YouTube*.

No momento em que o gesto e a fala são simultaneamente finalizados por Leda, abre-se uma oportunidade para que o turno de fala seja continuado por Tim. Não apenas pelo vocativo pronunciado, mas também pelo fato de que Leda direciona o seu corpo e seu olhar para ele, realizando uma entoação descendente após o seu turno. A jornalista também usa recursos prosódicos enfáticos (1.1) e gestos manuais (1.2 e 1.3) concomitantes ao arqueamento da sobrancelha (identificados por “<sup>”</sup> e “<sub>v</sub>”), dando ênfase na performance e orientando a atenção dos co-participantes da arena coletiva em que foram dispostas essas ações.

É possível observar, quando Tim começa a falar em resposta ao fim do turno de Leda logo em seguida, duas questões centrais: i) como a coordenação da fala, do olhar, dos gestos, e dos movimentos corporais, aliada à disponibilidade mútua dos falantes possuem efeitos significativos na construção do turno e da sequência de fala; e ii) o fato de que a pergunta de Leda foi percebida e interpretada como tal, obtendo uma resposta, como esperado, o que mostra que a ação realizada pela jornalista foi bem-sucedida.

Em seu turno, Tim Maia faz uma brincadeira com a palavra “carreira”: se valendo da polissemia da palavra, e de gestos manuais co-ocorrentes, faz um trocadilho. Esse item lexical não apenas se refere ao percurso profissional, mas também pode estar relacionado, dentre outras coisas, a um caminho ou mesmo à disposição de substância psicoativa em pó. Os interlocutores logo em seguida riem (1.7 e 1.10), e, ao mesmo tempo, Tim direciona o olhar ao seu redor (1.12 e 1.14), checando o sucesso da sua piada; então, continua “uma meta vamos dizer assim” (1.13) enquanto as risadas continuam e a entrevistadora diz: “é um gozador esse rapaz”, e repete “é um gozador” (1.15 e 1.17).

Imagen 5 – Capturas de telas: à esquerda, imagem de Tim Maia enquanto pronuncia a palavra “carreira” e, à direita, a reação dos interlocutores da interação



Fonte: Canal Leda Nagle, *YouTube*.

A reação dos co-participantes, que compartilham o *common ground* dessa interação, demonstra que a construção da ação de Tim também foi bem sucedida. Os interlocutores operam com uma sequência de ações que é construída através das atividades corporeadas dos seus interlocutores e o falante age e, com o todo o seu corpo, percebe os outros corpos envolvidos. Ao perceber o corpo dos outros no seu próprio, o falante encontra um prolongamento de suas ações. Seu corpo e dos participantes da interação se tornam um único todo, que se torna público na arena. É esse “esquema corporal” que faz com que o falante aja, é a partir da percepção de que tem da ação (falas, gestos, movimentos corporais, risadas) que os outros co-participantes fazem a seu redor que a interação é organizada.

Ainda nesse trecho, é possível observar a utilização com transformação dos materiais providenciados pelos participantes durante a interação. Retomando a transcrição de (1), aqui resumida e ilustrada em (2):

(2)

- |   |                                              |                            |                                        |
|---|----------------------------------------------|----------------------------|----------------------------------------|
| 1 | Leda: ↑ ô Tim você está levando↓(0.3) você ↓ | está mais empenhado na sua |                                        |
|   |                                              | ^carrei:ra                 | digamos assim↓                         |
| 2 | Tim: (0.3) é: (0.36)                         |                            |                                        |
| 3 | Leda: ↑ é isso ↓                             |                            |                                        |
| 4 | Tim: fileira na                              | carreira                   | alguma coisa que me leve a algum lugar |

É possível observar, tanto nos pares 2.1 e 2.4 quanto em 2.2 e 2.3, a reutilização de itens lexicais previamente usados em turnos anteriores por diferentes falantes. No par 2.1 e 2.4, Tim retoma a palavra *carreira*, antes introduzida pela jornalista, para, em seguida, fazer a o trocadilho. Em 2.3, por sua vez, vemos que a jornalista recupera a resposta prolongada dada por Tim em 2.2, para introduzir novamente seu turno com o mesmo item “é”, agora transformando-o em uma pergunta com uma prosódia bem marcada.

O que antes foi dito, portanto, serve para dar forma ao presente que agora os co-participantes habitam, enquanto o passado organizado, quando apreendido, ao mesmo tempo que restringe, também faz ser possível inúmeras possibilidades futuras relevantes e visíveis (Goodwin, 2018), como o tipo de ação que pode ocorrer a partir desse momento particular. Nesse caso, a resposta de Tim para a pergunta de Leda, a utilização com transformação dos materiais previamente usados para construir o trocadilho, e a resposta em forma de risada dos co-participantes. Então, vemos que a ação co-operativa e cumulativa ocorre não somente quando os interlocutores fornecem uma resposta a uma ação anterior, mas também por meio da emergência de novas ações e possibilidades.

Tim Maia prossegue dizendo que nove músicos e um advogado o processaram; repete – nove músicos. Continua dizendo que eram muitos julgamentos a que ele precisava comparecer, e que ele não gostava de ir nos julgamentos.

(3)

- 1 Tim: tinha  $\uparrow$ dias (0.54) viu Leda $\downarrow$  (1.18) que tinha um julgamento assim
- 2 Mãos de T: (mão esquerda em cima da bancada) (indicador esticado)
- 3 Cabeça de T: (esquerda)
- 4 Tim: um de dia e **dois** de tarde
- 5 Mãos de T: (indicador e dedo médio esticados)
- 6 Tim: então eram muito- foram **nove**:
- 7 Cabeça de T: (direita)
- 8 Mãos de T: (à frente de seu corpo, se movimentando)
- 9 Tim: (0.39) processos eu não gosto daquele negócio
- 10 Cabeça de T: (frente)
- 11 Tim: que a gente sabe com' é que é (0.45)
- 12 Cabeça de T: (baixo)
- 13 Tim: naquela época ainda tomava um uisquizi:nho
- 14 Mãos de T: (depicta pegar um copo e tomar algo)
- 15 Cabeça de T: (cima)
- 16 Tim: vai ficar indo lá: ver o juiz tu já tá meio doidão de   ressaca no  $\wedge$ dia seguinte
- 17 I:   é:
- 18 Tim: então você evita ir lá: $\uparrow$
- 19 Mãos de T: (palmas da mão fechadas paralelas ao corpo)

Nesse momento, podemos perceber a sua inquietação, já que não apenas enfatiza a quantidade de julgamentos por dia (3.1 e 3.4) e o total (3.6), tanto com uma prosódia marcada (3.4, 3.6) quanto com gestos co-ocorrentes (3.2, 3.6), como também movimenta rapidamente o seu corpo e cabeça tanto para a esquerda (3.3), para a direita (3.7), em direção aos outros jornalistas à sua volta, e para frente (3.10), enquanto faz movimentos repetitivos com as mãos (3.8) e franze as sobrancelhas para cima (3.16), para, finalmente, dizer, com ênfase, “então você evita ir lá”.

Imagen 6 – Captura de tela de sequência de fala no momento em que Tim Maia conta o número de processos em que estava envolvido



Fonte: Canal Leda Nagle, YouTube.

Depois de uma pausa, diz que foi condenado à revelia de três ou quatro dentre todos os processos e, com a testa franzida e olhando diretamente para algum jornalista mais à sua direita, diz que cada processo está em 230 mil reais. Ao fundo, podemos ouvir a jornalista Leda dizer “caramba”. Concomitantemente ao uso de gestos manuais, Tim continua dizendo que já recorreu, mas que não deu certo (novamente, franzindo as sobrancelhas). E repete: “são duzentos e trinta mil cada processo”. É após esse momento que uma expressão linguística motivada aparece: Tim usa uma depicção<sup>5</sup> do voo da abelha para performar aos co-participantes da interação o modo como as pessoas se comportam no tribunal, considerando o alto valor do dinheiro em disputa.

(4)

- 1 Tim: são duzentos e trinta mil ↑*cada* processo
- 2 Cabeça de T: (frente)
- 3 Mão de T: (mãos paralelas sobre a mesa, dedos indicadores esticados)
- 4 Tim: (0.44) eu já [ pedi:  
5 I: eu falo para o tribunal que é muita [ coi:sa  
6 Cabeça de T: (esquerda) [ muita coi:sa  
7 Tim:  
8 Mão de T: (palmas abertas para cima paralelas ao corpo)  
9 Tim: recorri tudo ma- [ ninguém (h)  
10 I: eu  
11 Tim: porque quando chega lá assim↑  
12 Mão de T: (mão direita para cima)  
13 I: acho que essas [ coisas devem ser simbólicas  
14 Tim: fica todo mundo tumultuado *que nem abelha sabe com é que é*  
15 Cabeça de T: (direita)  
16 Tim: *^duzentos e trintinha cada um fica todo mundo zzzzzz*  
17 Mão de T: (paralelas acima da ombro, se movimentando)  
18 Cabeça de T: (baixo) (frente)

Observa-se, nesse trecho, novamente o reúso com transformações, não apenas em 4.5 e 4.7, mas especialmente também em 4.14 e 4.16. Esse uso, produzido com uma alteração na qualidade da voz dando ênfase ao trecho, não é meramente ilustrativo: ele depicts um evento

<sup>5</sup> Do inglês, *depiction*, segundo o dicionário de Oxford, se refere à “ação de depictar; uma representação figurada, uma pintura; descrição gráfica”. Do original: *The action of depicting; painted representation, picture; graphic description* (disponível em: <https://www.oed.com/search/dictionary/?scope=Entries&q=depiction>. Último acesso em 20 de janeiro de 2024). Na falta de uma palavra em português que mais se aproxime do termo em inglês, será usada a tradução literal, ‘depicção’, e termos derivados.

previamente experienciado por Tim Maia (o som da abelha e o modo como as abelhas ficam ao redor de algo que querem muito, insistente) e a sensação de tumulto, incômoda de ir ao tribunal. Tim reusa esses materiais, performando para os seus interlocutores a sensação particular pela qual experiencia toda vez que vai ao tribunal, a sensação de tumulto e cobrança em relação à ambição do dinheiro requisitado.

Essa performance, todavia, não ocorre *a priori*: é somente pela percepção dos corpos dos participantes da interação que está se desenrolando momento a momento, da tensão que está sendo paulatinamente acumulada na arena, e de todos os componentes da interação, como o próprio ambiente que está à volta dos interlocutores, que é possível que a sua emergência ocorra.

IMAGEM 7 – Captura de tela do momento em que Tim Maia produz o som com gesto co-ocorrente zzzzz



Fonte: Canal Leda Nagle, *YouTube*.

Essa depicção sonora, zzzzz, é simbólica sonora, iconicamente motivada uma vez que veicula intrinsecamente uma semelhança entre forma e sentido – nesse caso, a forma do som, zzzzz, estabelece uma relação de semelhança às propriedades percebidas pelo falante daquilo a que ela se refere, ao barulho e ao movimento do bater das asas que a abelha produz. Essa forma também pode ser entendida como uma onomatopeia ou ainda, um ideofone onomatopeico (Akita *et. al.*, 2019), uma vez que é uma forma marcada que faz a depicção de uma imagem sensorial (Dingemanse, 2012), ou seja, uma palavra que *mostra* aos interlocutores da conversa, que *performa*, ao ser produzida, a percepção sonora tanto do barulho quanto do movimento da abelha para ilustrar a sensação inconveniente, para o cantor, de ir até o tribunal.

É curioso notar que simultaneamente a essa produção sonora icônica e ideofônica, há a co-ocorrência de um gesto manual também icônico (ou ainda, *co-speech gesture*, cf. Özyürek *et al.* 2020, p. 2), em que as mãos de Tim são colocadas na altura dos ombros, com os dedos abertos, e são movimentadas rapidamente para cima e para baixo, na horizontal, como se também simulasse o voo da abelha. Além disso, observamos a direção do olhar de Tim em busca de se alinhar com os outros participantes, o levantamento das sobrancelhas e a alteração na qualidade da voz, evidenciando, de certo modo, um clímax da organização da história, uma vez que a partir desse momento o turno é ocupado pela jornalista Leda que pergunta se Tim resolveu encarar, trabalhar mais para poder pagar esse valor.

Essa pergunta, portanto, revela que as motivações de Tim Maia foram satisfeitas e suas ações bem-sucedidas, uma vez que a fala subsequente de Leda assegura que ela, enquanto participante da interação, aceita (e não rejeita) as ações performadas por Tim. Posteriormente,

ele responde à pergunta de Leda afirmativamente e repete que são muitos processos, novamente reusando, com transformação, a expressão já utilizada. Em seguida, muda de assunto. Percebemos assim, que a atividade foi finalizada e que a ação pode, enfim, seguir adiante.

Assim, o evento é construído pela justaposição de recursos corporais que constituem a cena em progresso. Vemos, então, que à medida que a interação se desenrola *in situ*, momento a momento no tempo, todos esses recursos semióticos são usados para produzir e fazer emergir atividades significativas inteligíveis, sempre em relação à percepção corpórea dos participantes da interação.

## 5 Considerações finais

A Linguística enquanto disciplina, ao entender língua enquanto um construto abstrato, um sistema fechado e monolítico, não apenas se limita como também coloca tudo a perder quando não olha mais de perto para o que está de fato ocorrendo no processo interacional, uma vez que cada aspecto que se desenrola na interação é significativo e demonstra ter efeito estratégico proposital. Nesse sentido, este trabalho fez um recorte para demonstrar o modo como isso ocorre. Tendo como norte a iniciativa de aproximar da análise linguística fenômenos que geralmente são deixados às margens (Ameka, 2020; Dingemanse, 2017; Negrão; Viotti, 2020), a investigação colocou em foco não apenas o tratamento da interação como central na análise, como também ilustrou a emergência de uma expressão onomatopeica, ou ainda, de um ideofone onomatopeico, acompanhada de gestos e dos movimentos icônicos corporais, que, muitas vezes, são ignorados na descrição linguística de línguas oralizadas. Aqui vemos que esse é um componente fundamental para a completude da ação. A análise detalhada também possibilitou ilustrar o caráter intrinsecamente multimodal da comunicação humana (Özyürek *et al.*, 2020).

O trecho selecionado para análise ilustra o modo como os co-participantes em uma interação co-presencial se organizam coletivamente – utilizando todos recursos multimodais (fala, gestos das mãos, direção do olhar, movimento da cabeça, ombros, expressões faciais, posturas corporais etc.) a seu dispor, necessários para o desenvolvimento da ação em que estão envolvidos. O modo como observamos que as ações dos interlocutores incluem o uso de experiências anteriores com transformação, e que esse uso, no momento em que é realizado, serve como fonte para restringir e, ao mesmo tempo, possibilitar ações futuras que se darão a partir deste momento, nos fornece uma importante perspectiva para a descrição do trabalho *in situ* em que os participantes se engajam. Vemos que eles não apenas apreendem o que o outro está falando, mas também participam ativamente do processo semiótico em que a prática linguística acontece, e, a partir de então, criam novas modalidades de ação. A interpretação das faixas de consciência, antes notada por Schütz, é então visualizada de forma prática, com relação aos projetos específicos desses participantes em particular, como um processo acumulativo de ação co-operativa que se desenrola no tempo. Assim, por meio do trabalho co-operativo de construção da ação, ao usar propriedades combinatórias linguísticas e corporais distintivas como uma prática pública, os participantes habitam, juntos, o tempo que se desenrola naquele momento e naquele lugar específico (Goodwin, 2018, p. 52).

Ao nos darmos aos detalhes, portanto, podemos ver a preciosidade da prática linguística em ação. Eles não apenas descontinuam características fundamentais para a organização e desenvolvimento da ação, como também nos revelam um movimento organizado em

uma ecologia mais ampla de práticas de construção de fazer-sentido. Isso possibilita, assim, novos caminhos de investigação e de entendimento da natureza da linguagem humana pela qual navegamos e construímos o mundo ao nosso redor.

## Agradecimentos

Agradeço à professora Dra. Evani Viotti pelo oferecimento da disciplina “Bases epistemológicas da Semiótica das Interações” inserida no programa de Pós-Graduação de Linguística da Universidade de São Paulo (USP), que fundamentou a realização deste trabalho. Agradeço também às leituras criteriosas da professora Dra. Esméralda Vaillati Negrão e dos revisores deste artigo. Quaisquer falhas remanescentes são de minha inteira responsabilidade. Agradeço também ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da Bolsa de Produtividade em Pesquisa, processo 140847/2021-9.

## Notas da transcrição

Negrito em itálico: ênfase

Traço (-): interrupção da produção sonora

Colchetes: fala sobreposta

Números em parênteses: silêncio em segundos

Risada: (h)

“:”: prolongamento

“↓”: contorno entoacional descendente

“↑”: contorno ascendente

“^”: sobrancelhas arqueadas para cima

“\_” sobrancelhas arqueadas para baixo

Mão de L: mão (direita/esquerda) de Leda

Mão de T: mão (direita/esquerda) de Tim Maia

Cabeça de L: direção da cabeça de Leda

Cabeça de T: direção da cabeça de Tim Maia

I: interactantes

## Referências

AKITA, K.; DINDEMANSE, M. Ideophones (Mimetics, Expressives). In: ARONOFF, M. (ed.). *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. 2019. DOI:10.1093/acrefore/9780199384655.013.477.

AMEKA, F. K. Ideophones and the nature of the adjective word class in Ewe. In: VOELTZ, F. K. E.; KILIAN-HATZ, C. (eds.). *Ideophones*. Amsterdam: Benjamins, 2001. p. 25–48.

- AMEKA, F. K. *The gem is in the details*. ILARA Online – 12 de novembro de 2020. Disponível em: <<https://youtu.be/1x6Q2joOhsc>>. Acesso em 11 jan. 2025.
- AMEKA, F. K.; Terkourafi, M. What if...? Imagining non-Western perspectives on pragmatic theory and practice. *Journal of Pragmatics*, Amsterdam, v. 145, n. 7, p. 72–82. 2019. DOI: doi:10.1016/j.pragma.2019.04.001.
- AMHA, A. Directives to humans and to domestic animals: the imperative and some interjections in Zargulla. In: SIMEONE-SENELLE, M. C.; VANHOVE, M. (eds.). *Proceedings of the 5th International Conference on Cushitic and Omotic Languages*. Paris: Rüdiger Köppe, p. 211–229, 2013.
- ANCHIMBE, E. A.; JANNEY, R. W. Postcolonial pragmatics: An introduction. *Journal of Pragmatics*, Amsterdam, v. 43, n. 6, p. 1451–1459, 2011. DOI: 10.1016/j.pragma.2010.10.027.
- DINGEMANSE, M. Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, v.6, n.10, p. 654–672. 2012. DOI:<https://doi.org/10.1002/lnc3.361>.
- DINGEMANSE, M. On the margins of language: Ideophones, interjections and dependencies in linguistic theory. In: ENFIELD, N. J. (ed.). *Dependencies in language*. Berlin: Language Science Press, 2017. p. 195–203. DOI:10.5281/zenodo.573781.
- DINGEMANSE, M. Redrawing the margins of language: Lessons from research on ideophones. *Glossa: a journal of general linguistics*, Londres, v. 3, n. 1, 1–30, DOI: <https://doi.org/10.5334/gjgl.444>. 2018.
- DINGEMANSE, M.; LIESENFELD, A.; RASENBERG, M.; ALBERT, S.; AMEKA, F. K.; BIRHANE, A.; BOLIS, D.; CASSELL, J.; CLIFT, R.; CUFFARI, E.; DE JAEGHER, H.; DUTILH NOVAES, C.; ENFIELD, N. J.; FUSAROLI, R.; GREGOROMICHELAKI, E.; HUTCHINS, E.; KONVALINKA, I.; MILTON, D.; RĄCZASZEK-LEONARDI, J.; REDDY, V.; ROSSANO, F.; SCHLANGEN, D.; SEIBT, J.; STOKOE, E.; SUCHMAN, L. A.; VESPER, C.; WHEATLEY, T.; WILTSCHKO, M. Beyond single-mindedness: A figure-ground reversal for the cognitive sciences. *Cognitive Science*, v. 47, n. 1, e13230, 2023. DOI:10.1111/cogs.13230.
- ENFIELD, N. J. *Relationship Thinking*: Agency, enchrony, and human sociality. Oxford: Oxford University Press, 2013.
- FAVAREAU, D. The Theory of Meaning. In: FAVAREAU, D. (ed.). *Essential readings in biosemiotics*. Anthology and commentary. Vol. 5. Berlin: Springer Science, 2010. p. 81–114.
- GOODWIN, C. *Co-Operative Action*. Cambridge: Cambridge University Press, 2018.
- HANKS, W. F.; IDE, S.; KARAGIRI, Y. Towards an emancipatory pragmatics. *Journal of Pragmatics*, Amsterdam, v. 41, n. 1, p. 1–9. 2009. DOI:<https://doi.org/10.1016/j.PRAGMA.2008.02.014>.
- JOSEPH, B. D. On the linguistics of marginality: the centrality of the periphery. In: ANDERSON, G. et al. (eds.). *Papers from the 33rd Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, p. 197–213. 1997.
- LEVINSON, S. C. *The Interaction Engine*: Language in Social Life and Human Evolution. Cambridge University Press. 2025.
- MCCLEARY, L. E.; VIOTTI, E. C. Transcrição de dados de uma língua sinalizada: um estudo piloto da transcrição de narrativas na língua de sinais brasileira. In: LIMA-SALES, H. M. M. (ed.). *Bilingüismo dos surdos: questões lingüísticas e educacionais*. Goiânia: Cânone Editorial, 2007. p. 73–96.
- MCCLEARY, L. E.; VIOTTI, E. C. Linguistics in search of a semiotics of interaction. In: 10th BRAZILIAN INTERNATIONAL MEETING ON COGNITIVE SCIENCE (EBICC), 2015, São Paulo. *Annals of EBICC*

2015. São Paulo: Sociedade Brasileira de Ciência Cognitiva – SBCC. , 2017. p. 1–9. DOI: 10.13140/RG.2.2.26235.00800.
- MCCLEARY, L.; VIOTTI, E. Fundamentos para uma semiótica de corpos em ação. In: FIORIN, J. L. (Org.). *Novos caminhos da linguística*. São Paulo, SP: Editora Contexto, 2017. p. 171–193.
- MCNEILL, D. *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press. 1992.
- MONDADA, L. Conversation analysis: Talk and bodily resources for the organization of social interaction. In: MÜLLER, C; CIENKI, A.; FRICKE, E.; LADEWIG, S.; MCNEILL, D.; TEBENDORF, S. (eds.), *Body – Language – Communication*, Berlin, Boston: De Gruyter Mouton, 2013. p. 218–226. DOI: <https://doi.org/10.1515/978310261318.218>.
- MONDADA, L. Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*. v. 20, n. 3, p. 336–166, 2016. DOI: [https://doi.org/10.1111/josl.1\\_12177](https://doi.org/10.1111/josl.1_12177).
- NEGRÃO, E. V.; VIOTTI, E. C. Abordando a emergência das peculiaridades do sujeito do português brasileiro a partir das margens da ciência linguística. *Cuadernos de la ALFAL*, Santiago de Chile, v. 12, n. 2, p. 174–198. 2020.
- NGUÉ UM, E. Had Ferdinand de Saussure spoken Wolof or Basaa..., the discipline of linguistics would have fared differently. *Language, Culture and Society*, Amsterdam, v. 2, n. 1, p. 107–115. 2020. DOI: <https://doi.org/10.1075/LCS.00022.NGU>
- ÖZYÜREK, A.; RASENBERG, M.; DINGEMANSE, M. Alignment in multimodal interaction: An integrative framework. *Cognitive Science*, New Jersey, v. 44, n. 11, p. 1–29, 2020. DOI: <https://doi.org/10.1111/cogs.12911>.
- SACKS, H.; SCHEGLOFF, E.; JEFFERSON, G. A Simple Systematic for the Organisation of Turn Taking in Conversation. *Language*, New York, v. 50, n. 4. p. 696–735. 1974. DOI: 696–735. 10.2307/412243.
- SAUSSURE, F. *Curso de Linguística Geral*. Organização Charles Bally e Albert Sechehaye; com a colaboração de Albert Riedlinger. Tradução Antônio Chelini, José Paulo Paes e Izidoro Blikstein. 28ª ed. São Paulo: Cultrix, 2012.
- SCHÜTZ, A. *The phenomenology of the social world*. Evanston: Northwestern University Press. 1967.
- TOULMIN, S. *Cosmopolis: The Hidden Agenda of Modernity*. New York: The Free Press, 1990.
- VARELA, F. J.; THOMPSON, E.; ROSCH, E. *The embodied mind: Cognitive science and human experience*. The MIT Press. 1991.
- RASENBERG, M.; AMHA, A.; COLER, M.; VAN KOPPEN, M.; VAN MILtenburg, E.; DE RIJK, L.; STOMMEL, W.; DINGEMANSE, M. Reimagining language Towards a better understanding of language by including our interactions with non-humans. *Linguistics in the Netherlands*, v. 40, n. 1, p. 309–317, 2023. DOI: <https://doi.org/10.1075/avt.00095.ras>.
- WITTENBURG, P.; BRUGMAN, P.; RUSSEL, A.; SLOETJES, H. ELAN: a Professional Framework for Multimodality Research. In: FIFTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION (LREC), 2006, Gênova. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, Gênova: European Language Resources Association (ELRA), 2006. p. 1556–1559.