UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE PSICOLOGIA

Charles Desena Pereira

**GESTOS CONVERSACIONAIS:
UMA VISÃO INCORPORADA DA COGNIÇÃO**

Um estudo experimental sobre a produção e o processamento de gestos manuais
acompanhados da fala no aprendizado de línguas e no ensino de ciências

São Paulo

2021

Charles Desena Pereira

**GESTOS CONVERSACIONAIS:
UMA VISÃO INCORPORADA DA COGNIÇÃO**

Um estudo experimental sobre a produção e o processamento de gestos manuais
acompanhados da fala no aprendizado de línguas e no ensino de ciências

Tese apresentada ao Instituto de Psicologia da
Universidade de São Paulo, como parte dos
requisitos para obtenção de Título de Doutor
em Ciências.

Programa: Neurociências e Comportamento.

Orientadores: Hamilton Haddad Júnior
                          André Frazão Helene

Nome: Pereira, Charles Desena

Título: Gestos conversacionais: uma visão incorporada da cognição

Tese apresentada à banca avaliadora

Nível: Doutorado

Universidade de São Paulo

Instituto de Psicologia

Resultado:

Em: __ / __ / __

Banca Examinadora

Prof(a). Dr(a).: _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a).: _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a).: _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a).: _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a).: _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a).: _____

Instituição: _____ Assinatura: _____

Agradeço ao Professor Hamilton Haddad Júnior, que há 10 anos me recebeu em seu laboratório e me presenteou com um objeto de estudo desafiador e fascinante. Agradeço ao Professor Hamilton por seu trabalho intelectual e simplicidade na amizade e pessoa, exemplos que me inspiram e orientam desde então.

Agradeço ao Professor André Frazão Helene, que participou de todas as etapas de desenvolvimento deste trabalho. Agradeço por me receber em seu laboratório e grupo de pesquisa. Agradeço por sua generosa orientação, sempre com contagiante entusiasmo, incontáveis boas ideias, alegria e arte.

Agradeço à Professora Ivana Brito por, mesmo em tempos de grandes desafios, aceitar prontamente revisar o trabalho, conduzir o exame de qualificação e se colocar permanentemente à disposição minha e de meus colegas para orientação e encaminhamento de nossos trabalhos.

Agradeço à Professora Ivana Brito e aos Professores André Cravo e Marcus Vinícius Baldo pela leitura do trabalho e valiosas contribuições no exame de qualificação.

Que o Professor Marcus Vinícius Baldo, em nome de todos os docentes que contribuíram com meu percurso formativo e realização deste estudo, receba minha mais sincera gratidão.

Agradeço às Professoras Ivana Brito e Leila Miguel Stávale por toda colaboração diante dos inúmeros desafios na elaboração de nosso estudo experimental 2.

Agradeço ao colega Jimmy Nikaido, parceiro de laboratório, de estudo e de experimento, cuja amizade e parceria chegaram-me como mais um presente através do Professor Hamilton. Com incommuns compromisso e generosidade, para muito além de sua iniciação científica, Jimmy trabalhou lado-a-lado ativa e incansavelmente em todas as etapas do estudo, razão pela qual, sem hesitar, divido com ele a autoria deste trabalho. Sua admirável tranquilidade trouxe serenidade e equilíbrio ao nosso estudo.

Agradeço a toda a equipe INCO-CEPEL (Centro de Estudos e Pesquisas em Ensino de Línguas) da Faculdade de Educação da USP (FEUSP): Professoras Adriana Weigel, Lélia Souza, Sílvia Santos e Tatiana Reschke pela indispensável colaboração com a realização de nosso estudo experimental 1. Um agradecimento especial à Professora Livia de Araújo Donnini Rodrigues, coordenadora do programa, por seu tempo dispendido em nos receber, esforço em mediar o uso do espaço e tempo para a realização do experimento, por suas valiosas contribuições em nossas análises e divulgação dos resultados.

Um especial agradecimento a todos os participantes dos estudos experimentais.

Agradeço ainda a Arcelino Bezerra da Silva Neto do Laboratório de Pesquisa em Ensino de Química e Tecnologias Educativas (Lapeq) da Faculdade de Educação da USP (FEUSP) pelo auxílio no uso de tecnologias de análise de gestos.

Agradeço ao Professor Carlos Hitoshi Morimoto (IME-USP) e a todos seus orientandos e colaboradores do grupo LaTIn (Laboratory of Technologies for Interaction) por generosamente me receberem em diversas ocasiões em suas reuniões de estudo e por todo aprendizado que levo da interação com o grupo e com seu trabalho de pesquisa.

Agradeço a Rafael Barros de Oliveira da seção de Audiovisual e à equipe da Seção Técnica de Informática do Instituto de Biociências (IBUSP), aos estatísticos Luiz Silva dos Santos e Vinícius Frayze David do Instituto de Psicologia (IPUSP) e a todos os funcionários técnicos, administrativos e dos mais diversos setores e serviços da Universidade de São Paulo, que sempre prontamente nos atendem, permitindo que nosso trabalho de pesquisa aconteça e que nossa passagem formativa pela universidade seja plena.

Agradeço a meus alunos e seus pais por todos esses anos de apoio e confiança. Com pretexto de lhes ensinar, tive eu na companhia deles meus maiores aprendizados!

Agradeço a Sílvia por seu exemplo, incentivo e todo aprendizado ao seu lado.

Agradeço à CAPES pelo apoio financeiro à realização desta pesquisa.

A todos que de alguma forma deram rumo a este barco.

“[Gesture and speech] are a way of cognitively existing at the moment of speaking.”

McNeill & Duncan (2000)

RESUMO

Nas duas últimas décadas, produziu-se um enorme corpus de evidências da contribuição da gesticulação manual para o ensino-aprendizagem de diversas áreas do conhecimento. Abordagens incorporadas da cognição humana têm procurado elaborar modelos cognitivos que expliquem o fenômeno. Com a tecnologia hoje disponível, é possível acompanhar em tempo real processamentos cognitivos ocorrendo no sistema nervoso central. No entanto, pouco se sabe ainda sobre os processos neurocognitivos de integração linguística dos gestos.

Para situarmos nosso objeto de estudo e encaminharmos nossa problemática, abordamos na introdução deste trabalho as principais contribuições da literatura para a compreensão atual da origem e do processamento visual, linguístico e cognitivo da gesticulação manual conversacional. Trazemos também as principais tipologias e modelos cognitivos de sua produção.

Dados experimentais de trabalhos anteriores de nosso grupo de pesquisa mostraram que gestos manuais beneficiam a compreensão de narrativas por estudantes de língua estrangeira sem, contudo, demandarem alocação aberta da atenção. Como estudo complementar, exploramos experimentalmente no presente trabalho o impacto da gesticulação manual na *produção* de narrativas em língua estrangeira. Avaliamos alterações em diferentes parâmetros linguístico-discursivos das produções narrativas sob controle da gesticulação manual durante a realização da tarefa experimental de produção narrativa.

Também a partir de dados experimentais de trabalhos anteriores de nosso grupo de pesquisa, hipotetizamos que gestos manuais acompanhados da fala sejam processados majoritariamente por visão periférica. Buscamos testar nossa hipótese avaliando experimentalmente durante a apresentação de uma aula de ciências o impacto da gesticulação manual da professora em ambos o comportamento ocular de aprendizes e o seu desempenho em tarefa de aprendizagem.

Esperamos contribuir para modelos teóricos, bem como para modelos experimentais de investigação de processos neurocognitivos da gesticulação manual.

PALAVRAS-CHAVE

Gestos, cognição incorporada, aprendizagem, atenção, movimento ocular.

ABSTRACT

In the last two decades, a great body of evidence has been produced of the contribution of manual gesticulation to the teaching-learning process of different areas of knowledge. Embodied approaches to human cognition have sought to develop cognitive models to explain the phenomenon. Recent technology has made it possible to track real-time cognitive processing in the central nervous system. However, little is known about the neurocognitive processes of linguistic integration of gestures.

In the introduction, we address the main contributions of the literature to the current understanding of the origin and of the visual, linguistic and cognitive processing of conversational manual gesticulation. We also address the main typologies and cognitive models of its production.

Experimental data from previous works by our research group have shown that manual gestures benefit narrative comprehension by foreign language students, without requiring visual attention allocation to the gestural field. As a complementary study, we experimentally explored in this work the impact of hand gestures on the production of narratives in a foreign language. We evaluated changes in different linguistic-discursive parameters of narrative productions under manual gesticulation control.

Also based on experimental data from previous works by our research group, we hypothesize that manual gestures accompanied by speech are processed mainly by peripheral vision. We tried to test our hypothesis by experimentally evaluating during the presentation of a science class the impact of the teacher's manual gesticulation on both the learners' eye movement and their performance in a learning task.

We hope to contribute to theoretical and experimental models of investigation of neurocognitive processes of manual gesticulation.

KEYWORDS

Gestures, embodied cognition, learning, attention, eye movement.

SUMÁRIO

INTRODUÇÃO

1. Por que os gestos?	12
1.1 Mas para onde foram os gestos?	13
2. Gestos como origem evolutiva da linguagem	15
2.1 O valor social da percepção de movimento: o Sulco Temporal Superior (STS) no processamento de pistas sociais	15
2.2 Ações intencionais na origem de uma sintaxe gestual simples: neurônios-espelho e a atenção social	17
2.3 Do gesto à fala: o controle motor fino do trato vocal, uma alça fonológica e uma memória operacional auditiva	18
3. A percepção visual dos gestos	22
3.1 Movimentos oculares e a atenção visual explícita	22
3.2 Especializações na retina	26
3.3 Da retina ao córtex: duas vias do sistema visual	29
4. A tipologia dos gestos	32
4.1 Paul Ekman e Wallace V. Friesen e a sistematização do comportamento não-verbal	32
4.2 Adam Kendon e a unidade sincrônica gesto-fala	36
4.3 David McNeill e a tipologia dos gestos ilustradores	38
5. Modelos cognitivos de produção integrada gesto-fala	42
5.1 Sketch Model (De Ruiter, 2008)	42
5.2 O modelo de Krauss, Chen & Chawla (1996)	45
5.3 Por um modelo integrado de processamento de informação	48
6. Gestos na cognição: gestos para quem?	50
6.1 A função dos gestos na produção da fala: uma abordagem incorporada da cognição	50
6.1.1 A carga cognitiva e a memória operacional	51
6.1.2 O acesso lexical	53

6.2 A função comunicativa dos gestos	55
6.2.1 O empacotamento da informação	56
6.2.2 O número e o posicionamento de ouvintes-expectadores	57
6.2.3 Efeitos de visibilidade dos gestos na compreensão de histórias	58
6.2.4 Perspectivas divergentes	59
6.2.5 A intersubjetividade	61
7. Gestos em diferentes línguas	63
7.1 Diferenças étnico-culturais e linguístico-estruturais	63
7.2 Gestos na aprendizagem de línguas estrangeiras	65
8. Gestos na produção narrativa	70
8.1 Gestos no desenvolvimento linguístico-narrativo etário	70
8.2 Elementos visuo-espaciais e a criatividade na gesticulação narrativa	74
8.3 O planejamento discursivo	77
8.4 O uso compensatório de gestos	80
8.5 O uso compensatório de gestos e a competência em língua estrangeira	83
8.6 Aumento da carga cognitiva e a disfluência em gesto e fala	85
8.7 Gestos e a hiperexplicitação de referentes discursivos	88
9. A gesticulação como elemento inerente à comunicação e ao aprendizado	91
 ESTUDOS EXPERIMENTAIS	
Objetivo geral	94
Experimento 1	95
Experimento 2	126
Discussão geral	188
REFERÊNCIAS	191
ANEXOS	207

CAPÍTULO 1

POR QUE OS GESTOS?

Quando consideramos tradicionalmente a comunicação verbal humana, talvez em virtude do prestígio cultural que desfruta o registro escrito, observamos com frequência o tratamento da fala isoladamente. É frequente que se faça a análise apenas do conteúdo linguístico como componente da mensagem veiculada, negligenciando uma série de outros componentes comunicativos não-verbais, como a postura corporal, o olhar e a expressão facial, a voz e os gestos, compreendidos na maior parte do tempo como enquadramentos secundários da mensagem propriamente dita.

Entretanto, alguns desses componentes não-verbais da comunicação, como a gesticulação, desempenham função importante na elaboração e na compreensão da fala. A gesticulação caracteriza-se por movimentos corporais que emergem no ato comunicativo perfazendo, com outros componentes que possam se fazer presentes, um único esforço de elocução. Gestos organizam turnos de fala entre interlocutores; dão andamento e acentuação prosódica conversacional; desambigam, analisam e marcam unidades semânticas; introduzem e organizam espacialmente referentes discursivos, focalizando sobre eles a atenção compartilhada entre os interlocutores.

Mas para além da intencionalidade comunicativa no ato de elocução, gestos atuam diretamente em processos cognitivos, como na facilitação do acesso lexical, no alívio de carga cognitiva e no incremento de representações mentais por redes neurais. McNeill (2005) propõe a integração de fala e gestos operando dialogicamente em uma mesma economia cognitiva: a fala, agenciando unidades analíticas hierárquicas formais e sequenciadas no tempo, e os gestos, assumindo formas livres no esforço de acomodação espacial de significados mais globais e imagéticos.

Gesto e fala potencializam-se, portanto, ao elaborarem conteúdo linguístico e imagético em uma mesma unidade significativa; mais que isso, operando por diferentes modalidades em uma mesma economia cognitiva, abastecem-se semioticamente propulsionando a elaboração de sentido e do próprio pensamento (McNeill, 2005).

Estas são apenas algumas das funções desempenhadas pela gesticulação, dentre outras importantes que desenvolveremos no corpo deste trabalho, que nos permitem desde já, não obstante a brevidade desta introdução, vislumbrarmos a íntima relação entre a produção de gestos e a produção da fala; bem como anteciparmos a problemática central do estudo dos gestos e de

nosso presente trabalho: a razão e a finalidade da emergência dos gestos, expressas no custo-benefício cognitivo, linguístico e social de sua produção.

1.1 MAS PARA ONDE FORAM OS GESTOS?

Nas últimas três décadas, a tecnologia em telecomunicações avançou exponencialmente. A telefonia móvel e a popularização da internet permitiram a comunicação a distância de imagem e som em alta qualidade em tempo real. Atualmente, familiares e amigos em diferentes localidades no mundo se encontram com frequência sem necessidade de deslocamento; reuniões e conferências a distância são cada vez mais frequentes; pessoas atendem a eventos sociais em tempo real a partir de suas residências.

Mas, se a globalização econômica, social e cultural já vinha como uma realidade inevitável, pressionando-nos a uma necessidade cada vez maior de permanente conectividade, se a modalidade virtual de nossas atividades sociais já se fazia presente em nosso cotidiano, hoje, por circunstâncias ainda mais prementes, encontramos-nos dela definitivamente dependentes. Condições sanitárias em virtude de uma pandemia global (a COVID-19) impuseram-nos de forma contundente, sem aviso prévio nem data de validade, o confinamento e o isolamento social da população mundial. Essa atual realidade, ou *nova realidade* como convencionamos chamá-la, fez com que, sem atraso, até mesmo os mais resilientes mergulhassem de cabeça nas plataformas, redes e comunidades sociais virtuais.

Hoje, atividades físicas são conduzidas a distância, aulas de yoga e artes marciais adaptaram suas práticas para exercícios individuais no espaço da sala dos próprios praticantes; reuniões profissionais são agendadas e realizadas virtualmente por pessoas em suas casas; aulas do ensino infantil ao ensino universitário passaram a ocorrer regularmente a distância através de plataformas educativas; até eventos sociais celebrativos como nascimentos, aniversários e casamentos, ou mesmo um simples encontro social entre amigos em um *happy hour* após o horário de trabalho também já se adaptaram e passaram a ser parte dessa nova rotina a distância.

Interessantemente, contudo, à exceção de exposições artísticas, atividades físicas e outras que tradicionalmente envolvem trabalho e exibição do corpo, observamos na grande maioria das transmissões *online* o enquadramento apenas do rosto dos participantes – quando não transmitindo apenas a sua voz: “costumava dar aulas para pessoas”, disse recentemente uma professora, “hoje dou aulas para bolinhas. Preciso me adaptar.”

A anedota não é caso raro, e nos serve para ilustrar a pouca conta em que se têm o papel e a contribuição de movimentos corporais para a realização de encontros sociais, presenciais ou virtuais. Gestos manuais beneficiam não apenas o falante na organização do pensamento e elaboração de ideias durante a produção da fala, como também o ouvinte-expectador em seu

engajamento atencional e esforço de compreensão de uma cognição social que se pretende compartilhada.

Ainda que tenha colocado algum constrangimento, a modalidade virtual não impediu as pessoas de gesticularem durante seus encontros virtuais, tanto que observamos vez ou outra no enquadramento da câmera o vazamento do vulto veloz da mão ou antebraço de nossos interlocutores – aliás, é notório o fato de que pessoas gesticulam ao falarem ao telefone, no escuro ou mesmo na ausência de interlocutores (Goldin-Meadow, 2003).

No entanto, pela própria configuração física e posicionamento dos equipamentos, geralmente uma câmera integrada ou posicionada na porção inferior ou superior de uma tela de computador ou de celular a poucos centímetros de distância, mas também por privilegiarem os internautas a exibição da imagem de seu rosto negligenciando no enquadramento a importância da exibição de outras articulações do corpo (aliás, como dissemos ao início ser o senso comum em uma perspectiva tradicional da comunicação humana), tem a modalidade *online* dessas atividades sociais privado por tais razões as pessoas de acessarem a gesticulação corporal de seus interlocutores. Alheios ou pouco atentos para o fato, observamos recorrentemente no enquadramento de suas transmissões a exibição apenas de seus rostos, cabelos e penteados, flores, quadros e estantes de livros.

Ao prepararem suas transmissões, as pessoas de fato têm presente a importância do contato visual social através do olhar, da expressão facial, até mesmo da articulação motora fonológica – todos temos a percepção do prejuízo para a compreensão e boa comunicação quando ocorrem, com certa frequência infelizmente, congelamentos de imagem ou dissincronias sonoras nas transmissões. Negligenciam elas, não obstante, os gestos.

Apesar do custo cognitivo de planejamento e execução motora, que se expressa também em custo linguístico-discursivo (tese que aqui defenderemos, como colaboração à área de estudo), gestos, dentre outras funções das quais algumas já antecipamos, veiculam informação, participam na elaboração do sentido, na manutenção do interesse e da atenção visual do expectador, na construção do espaço cognitivo conversacional comum aos interlocutores.

Mais do que nunca, portanto, em tempos de aulas e encontros virtuais em uma nova realidade imposta por inesperada pandemia, é hora de olharmos para nossos gestos.

Quais gestos, porém? Quando passamos a percebê-los? Olhamo-os de fato? Como integramos gestos sociais em nosso repertório individual? Quanto carregam de significado universal e de idiossincrasia? Gesticulamos intencional ou espontaneamente? Quão conscientes são a produção e a percepção de gestos? Quais benefícios cognitivos e sociais trazem ao falante e a seu interlocutor? Essas são questões centrais no estudo dos gestos, questões que desenvolveremos nos capítulos a seguir.

CAPÍTULO 2

GESTOS COMO ORIGEM EVOLUTIVA DA LINGUAGEM

2.1 O valor social da percepção de movimentos: o Sulco Temporal Superior (STS) no processamento de pistas sociais

A atuação dos gestos em nossa economia cognitiva e, mais especificamente, em nosso sistema linguístico, na produção e compreensão de significados elaborados com a fala, remonta na linha evolutiva de nossa espécie à percepção social de movimentos corporais. Circuitos neurais que se ativam na presença de determinados movimentos, tais como o caminhar, o agarrar, o olhar e o apontar, foram selecionados ao longo de nossa evolução, de modo que passássemos a percebê-los como relevantes e significativos. A percepção desses movimentos teria caminhado para a atribuição de intencionalidades a outros indivíduos de mesma espécie, antecipando a construção social de uma alteridade subjetiva (Allison, Puce, Spencer & McCarthy, 1999).

Alguns estudos têm recebido crescente destaque ao investigarem a percepção social de gestos manuais na origem evolutiva da linguagem. Perrett et al. (1989) estudaram a responsividade a movimentos manuais por células no Sulco Temporal Superior (STS) em macacos. Os autores observaram variações no padrão de disparo dessas células conforme variaram qualitativamente os movimentos: encontraram, por exemplo, maior ativação para movimentos de agarre do que para movimentos de retração da mão; bem como maior padrão de ativação das células para movimento direcionado a um objeto, independentemente da identidade ou tamanho deste (ver também Grèzes, Costes & Decety, 1989).

Estudos com humanos também mostraram ativação de regiões do STS por estímulo de movimentos funcionais, mesmo na ausência de pistas corporais formais; o andar, o dançar, o arremessar e outros movimentos são percebidos como movimentos significativos por regiões do STS mesmo quando representados apenas por pontos luminosos localizados nas extremidades corporais e vistos no escuro (Johansson, 1973; Bonda et al., 1996; ver também Howard et al, 1996).

Gallagher et al. (2000) sugerem, contudo, que tais movimentos sejam apenas um dos aspectos sociais processados pelo STS. Em estudo com ressonância magnética por imageamento funcional (fMRI) em humanos, os autores verificaram ativação do STS e regiões adjacentes estimulados não só por imagens em movimento como também por outras pistas sociais de imagens estáticas de boca, mãos e corpo, sugerindo processamento de intencionalidades e de movimentos implícitos. Também foi observada semelhante atividade neural tanto na porção rostral do Lobo Parietal Inferior (IPL), associado ao processamento de informações de observação e execução de ações (Perret et al., 1990), como na área pré-motora F5, adjacente à área B44, esta última compreendida mais recentemente como um local de integração de sinais sensório-motores parietais com informação pré-motora e pré-frontal (Gerbella et al., 2011).

Figura 1. Ilustração do Sulco Temporal Superior (STS), Lobo Parietal Inferior (IPL), Área de Broadman 44 e F5.

Apesar de extensa e complexa, a localização anatômica do STS parece de fato favorecer a integração de informações de ambos sistemas ventral, de reconhecimento de objeto (“o quê”) e dorsal, de localização e movimento espacial (“onde”; Ungerleider & Haxby, 1994; Cusick, 1997). Nessa mesma linha, Allison, Puce & McCarthy (2000; ver também Brothers, 1990) propõem que pistas sociais sejam inicialmente processadas pelo STS e suas imediações. Análises subsequentes seriam realizadas com projeções de *feedforward* do STS à amígdala, que por sua vez adicionaria saliência emocional ao estímulo sensorial com projeções de *feedback* ao córtex visual e ao STS (Allison et al., 2000; Rockland, K., 1997; Adolph, R., 1999; Oram & Richmond, 1999; Sugase et al.,

1999), como uma indução à amplificação atencional da atividade evocada por estímulos socialmente salientes (Posner & Dehaene 1994).

2.2 Ações intencionais na origem de uma sintaxe gestual simples: neurônios-espelho e a atenção social

Alguns autores se apoiaram em tais evidências experimentais para proporem um quadro conceitual no qual circuitos neurais dessa natureza estariam na origem do comportamento gestual social precedente à linguagem (Rizzolatti & Arbib, 1998; Arbib 2005, 2012). Na atividade de agarre e controle manual voluntário, por exemplo, estaria um comportamento simples – mas direcionado – de alcance de objeto; uma ação motora simples, mas com potencial social quando na presença de um observador.

Em um cenário evolutivo, a observação do gesto intencional direcionado, que em princípio teria um valor imitativo favorecendo o aprendizado da ação por observação, como se verifica em outros animais, teria contribuído, no curso da evolução, para o fenômeno social de compartilhamento da atenção. A imitação pantomímica dessa ação por diferentes indivíduos teria ritualizado seu significado, permitindo a emergência de uma primeira semântica de referência externa (Aboitiz, 2012). Posteriormente no cenário evolutivo, o uso de ferramentas teria enriquecido esse repertório semântico convencionalizado (Stout & Chaminade, 2012). Nessa linha, corroboram com a proposta de Rizzolatti & Arbib (1998), sobre a origem gestual da linguagem, a difusão do gesto de apontamento significativo em primatas não-humanos (Leavens, 2004), bem como a maior preferência pela mão direita por bebês humanos e babuínos no uso de apontamentos em tarefas comunicativas do que em ações de agarre de objetos, evidenciando a dominância do hemisfério esquerdo sobre gestos comunicativos (Meunier, Vauclair & Fagard, 2012).

Mas mesmo gestos intencionais estando presentes em outros primatas, o aspecto abstrato de comunicação simbólica da representação pantomímica só se verifica em humanos (Cartmill, Beilock & Goldin-Meadow, 2012). Para tanto, esquemas neurais com representações semânticas ativados pela observação de ações, os chamados neurônios-espelho (di Pellegrino et al., 1992; Rizzolatti & Luppino, 2001; Rizzolatti & Craighero, 2004), teriam favorecido a origem de uma sintaxe simples ao permitirem a compreensão significativa e intencional dessas ações (Arbib, 2005). Em macacos, neurônios-espelho foram relatados presentes em F5, sendo ativados por gestos faciais como o bater dos

lábios e outros comportamentos relacionados a hábitos alimentares. Em humanos, verificou-se ativação na área 44 na presença de movimentos de boca relacionados a objetos, bem como na imitação de gestos (di Pellogrino et al., 1992; Buccino et al., 2001).

Outros estudos (Izumi & Kojima, 2004) mostraram ainda que chimpanzés podem associar vocalizações com gestos faciais, bem como (Taglialetela et al., 2008) que a área homóloga a área de Broca no chimpanzé atinge máxima ativação durante ações comunicativas gestuais e vocais, particularmente quando gestos e vocalizações são orientadas para chamada da atenção de outros indivíduos.

Aboitiz (2012) tem sugerido fortemente, nesse sentido, uma integração de gestos manuais e faciais com padrões vocais, estes tendo se desenvolvido nos humanos pela evolução de uma circuitaria de controle neural de aprendizagem vocal comunicativa, como a encontrada em aves canoras (Tokarev et al., 2011). Alguns autores sugerem que o circuito associado à alça fonológica em humanos contenha elementos semelhantes a neurônios-espelho que participem na geração de interface sensório-auditivo-motora (Aboitiz et al., 2006; Arbib, 2012; Aboitiz, 2012). Corroboram essa tese estudos com imageamento cerebral em humanos, que mostram ativação das áreas BA44, BA45 e BA47 pela cooperação gesto-fala na comunicação (Willems, Özyürek & Hagoort, 2007; Gentilucci & Dalla Volta, 2008). Dessa forma, em ambos humanos e macacos, um sistema de comunicação multimodal faria uso de circuitos neurais sobrepostos subservindo ambos gestos vocais e corporais (Aboitiz & Garcia, 2009).

2.3 Do gesto à fala: o controle motor fino do trato vocal, uma alça fonológica e uma memória operacional auditiva

Como visto até aqui, habilidades cognitivas conceitualmente tratadas na literatura como teoria da mente (Baron-Cohen, 1995; Permack & Woodruff, 1978; Ellis & Gunter, 1999), tais como a percepção de movimentos, a atribuição de intencionalidades por ativação de esquemas neurais de representação interna de ações, bem como a capacidade de compartilhar socialmente a atenção teriam sido aquisições fundamentais na evolução para a emergência da linguagem em nossa espécie. No contexto evolutivo, a gesticulação manual teria desempenhado papel fundamental no incremento cognitivo, funcionando como índices na percepção de intencionalidades e disposições com informações relevantes para a atenção e cognição sociais (Langton et al., 2000; Perrett et al., 1992).

Para a aprendizagem vocal comunicativa e processamento complexo de sequências fonológicas, contudo, teriam sido necessárias ao menos outras duas etapas evolutivas, a saber: a evolução de um controle voluntário fino sobre o trato laríngeo e supralaríngeo via projeção cortical direta do tronco encefálico a neurônios motores vocais, e a expansão da capacidade de memória operacional auditiva, possibilitando ensaios fonológicos de itens recentes na memória de curto prazo (Baddeley, 2003). Contribuiu para tanto o desenvolvimento de um circuito apoiado na via dorsal conectando as áreas de Wernicke e de Broca, uma alça fonológica. Tal circuitaria ampliou significativamente o leque de vocalizações, tornando possíveis novas combinações de fonemas, a elaboração de mensagens complexas e a habilidade da fala interna (Aboitiz, 2012).

Baddeley & Hitch (1974) propuseram um modelo cognitivo de armazenamento temporário de informação facilitador de uma série de atividades cognitivas como o raciocínio e a aprendizagem: a memória operacional. Com recursos atencionais e capacidade de armazenamento limitados, a memória operacional permite que informações perceptuais sejam ativamente reensaiadas antes de seu esvanecimento, controlando em certa medida a alimentação de informação em uma memória de maior longo prazo.

Em seu modelo original, Baddeley & Hitch (1974) admitem como componentes desse sistema uma central executiva e dois componentes de retenção temporária, um fonológico (a alça fonológica) e outro visuo-espacial (*visual sketchpad*), como veremos a seguir (**Figura 2**).

A.

B.

Figura 2. Memória operacional (Helene, 2006).

(A) Modelo de três componentes proposto inicialmente por Baddeley e Hitch (1974), a área central se refere ao componente de gerenciamento atencional (a central executiva), enquanto as áreas laterais da figura representam as alças de manutenção de informações por curto período de tempo (adaptado de Baddeley, 1999);

(B) Em 2000, Baddeley inseriu um quarto componente, o retentor episódico, em seu modelo original. As áreas em branco representam os componentes atencionais e de retenção temporária de informações e as áreas em cinza os sistemas de retenção de longa duração (adaptado de Baddeley, 2000).

O primeiro componente, a alça fonológica, mais estudado por sua maior tratabilidade, é capaz de reter traços de memória por ensaios rearticulatórios análogos à fala subvocal. Pelo fato de a articulação da fala ocorrer na linha do tempo, este sistema é limitado temporalmente em seu volume de unidades perceptuais potencialmente rearticuláveis. Estudos de pacientes com lesão cerebral e de neuroimagem em sujeitos normais sugerem o envolvimento da região temporoparietal esquerda, identificando a área de Broadman BA 40 como o local do componente de armazenamento da alça, e a área de Broca (BA 6/44) como envolvida no componente de ensaio articulatório (Paulesu et al., 1993; Awh et al., 1996; Jonides et al., 1996; Jonides & Smith, 1997). Para alguns autores, a evolução da alça fonológica teria facilitado a emergência da linguagem (Baddeley, Gathercole & Papagno, 1998); de forma análoga, nessa mesma linha de raciocínio, a capacidade da alça fonológica seria um bom preditor da habilidade de crianças e adultos de aprenderem uma língua estrangeira (Baddeley, 2003).

Quanto ao *visual sketchpad*, ou esboço esquemático visual, também ele possui recursos limitados de armazenamento, mas por razões ecológicas de economia cognitiva, uma vez que, considerando que o mundo visual tipicamente persiste no tempo, seria redundante e um desperdício de recursos o armazenamento detalhado de informações

visuais do ambiente. Ademais, o armazenamento de informações visuais é dependente da composição de informações constituintes do objeto visualizado, que podem concorrer entre si por recursos atencionais se de mesma categoria (cor, forma, localização, etc; Baddeley, 2003). Smith et al. (1996) identificaram a memória operacional visuo-espacial como localizada primariamente no hemisfério direito, com ativação também do córtex parietal inferior direito (BA 40), córtex pré-motor (BA 6) e córtex frontal inferior (BA 47), bem como do córtex occipital extra-extriado anterior (BA 19).

Já a central executiva estaria associada ao controle atencional em ocasiões de insuficiência do controle automático dos processos de rotina, comportamentos baseados em hábitos ou esquemas aprendidos, frequentemente alheios à consciência do sujeito.

Ligado à central executiva, propõe-se ainda um quarto componente ou subsistema, o *buffer episódico*, compreendido como um espaço de trabalho global responsável por agrupar informações de diferentes sistemas em episódios integrados, tornando-os acessíveis ao conhecimento consciente e à posterior evocação pela memória.

Damasio (1994) chama a atenção para a importância do controle conativo, emocional e motivacional na memória operacional e, por consequência, no direcionamento da ação e dos processos mentais como a aprendizagem. Estudos experimentais e modelos de processamento cognitivo têm sugerido, como veremos mais adiante, a atuação direta de gestos manuais em processos envolvendo a memória operacional.

Retornando ao contexto evolutivo, e contemplando agora em grande salto retrospectivo, o que observamos em nossa espécie é, portanto, a passagem de um estágio social comunicativo dependente do direcionamento explícito da atenção para ações motoras simples e intencionais de outros indivíduos, para um estágio atual de total e natural integração de movimentos corporais gesticulatórios à ação de interlocução e construção de uma cognição socialmente compartilhada.

Observamos, assim, a passagem evolutiva de ações motoras sobre circunstâncias reais, ou seja, sobre o mundo sensorialmente perceptível, para ações virtuais de manipulação conceitual, dentre elas os gestos. Ademais, igualmente interessante e desafiador à ciência, com o advento da linguagem, observamos a dispensa do direcionamento aberto da atenção para tais ações motoras gesticulatórias; dito em outras palavras, não mais olhamos para os gestos. Teria, portanto, o advento da linguagem nos deslocado não apenas as ações como também a atenção e a consciência de um plano real objetivo para um plano virtual subjetivo.

No estágio atual evolutivo, contudo, como percebemos e integramos os gestos em interação social? Vejamos no capítulo a seguir estudos que investigaram experimentalmente a questão.

CAPÍTULO 3

A PERCEPÇÃO VISUAL DOS GESTOS

Gestos manuais emergem bastante cedo na interação entre movimentos corporais intencionais e a produção da fala: crianças entre 14 e 23 meses de vida passam a combinar e sincronizar palavra e gesto, até então produzidos isoladamente, antes mesmo de combinarem palavra com palavra (Butcher & Goldin-Meadow, 2000). Uma vez sincronizados, fala e gesto passam a operar sincrônica e dialogicamente, abastecendo-se um ao outro semioticamente (McNeill, 2005). Mas para além de atuarem na economia cognitiva do próprio falante, gestos atuam também em uma cognição socialmente compartilhada; a gesticulação está presente em interlocuções por falantes de diferentes idades, línguas e culturas.

Curiosamente, no entanto, em interação social, pouco nos damos conta de sua presença. Este é um fenômeno que surpreende e desafia pesquisadores cognitivistas e neurocientistas ainda hoje. De fato, não olhamos – ou olhamos muito pouco – para os gestos; ou seja, não lhe dirigimos explicitamente nossa atenção. Outras regiões corporais são sabidamente alvos preferenciais de nosso olhar e de nossa atenção: a face humana talvez seja desde cedo o alvo mais frequente de nosso olhar e de maior interesse atencional. Da região dos olhos extraímos informações sobre identidade, idade, sexo, aspectos de atratividade e de estado emocional; a partir do olhar, somos também capazes de identificar direcionamento atencional sobre o ambiente, e de inferir intenções de ações, importantes pistas sociais em humanos – estudos mostram capacidade de percepção e direcionamento atencional por influência do olhar em crianças já em seus 3 primeiros meses de vida (para uma revisão sobre olhar e contato visual, ver Klein, 1986).

Mas como percebemos e integramos os gestos sem lhes dirigirmos o olhar?

3.1 Movimentos oculares e a atenção visual

O estudo de movimentos oculares tem fornecido dados para a compreensão de processos perceptuais e cognitivos há quase um século (Burswell, 1935; Yarbus, 1964;

Carpenter, 1977; Duchowski, 2007; Van Gompel et al., 2007). Os sistemas que organizam atividades oculomotoras, como o colículo superior e o campo ocular frontal, estão estreitamente vinculados a mecanismos atencionais (Rizzolatti et al., 1987; Awh et al., 2006). O direcionamento do olhar coincide frequentemente com o direcionamento atencional, fenômeno chamado de orientação aberta da atenção (Posner et al., 1980). Contudo, um estímulo ou região do campo visual pode vir a ser objeto de atenção dissociado de movimento dos olhos (Wright & Ward, 1994); neste caso, dizemos de um fenômeno de orientação encoberta da atenção.

A análise de fixações oculares tem permitido, assim, rastrear a mobilização e direcionamento de recursos atencionais de sujeitos experimentais, fornecendo pistas comportamentais de processos cognitivos subjacentes ao cumprimento de uma determinada tarefa.

Pereira & Haddad (2015) estudaram movimentações oculares durante a compreensão de narrativas por aprendizes de língua estrangeira. Os pesquisadores controlaram a presença dos gestos do narrador durante a apresentação em vídeo das narrativas, e obtiveram como resultado melhora de aproximadamente 10% na tarefa de compreensão na condição de visualização dos gestos. Não obstante, a análise da movimentação ocular dos participantes mostrou que os sujeitos experimentais olharam para os gestos do narrador (na condição com gestos) apenas 7% do tempo total das narrativas; no restante do tempo, o olhar foi direcionado ao rosto do narrador, principalmente na região dos olhos.

Rimé, Boulanger & d'Ydewalle (1988, citado em Rimé & Schiaratura, 1991) estudaram o comportamento visual de sujeitos assistindo a apresentação em vídeo da narração de uma história contada em três línguas diferentes (na língua nativa dos sujeitos experimentais, o flamengo, em uma língua parcialmente conhecida, o francês, e em uma língua totalmente desconhecida por eles, o russo), e obtiveram como resultado que os sujeitos olharam para apenas 6,7% dos gestos na condição língua nativa, 11,2% na condição língua parcialmente conhecida, e 19,6% na língua totalmente desconhecida.

Esses e outros estudos mostram que, apesar de sua contribuição para a comunicação face-a-face, gestos são muito poucas das vezes e muito pouco do tempo olhados por interlocutores (ver também Argyle & Cook, 1976; Kendon, 1990); o olhar mantendo-se, na maior parte do tempo, direcionado aos olhos do falante (**Figura 3**). Igualmente interessante, mesmo interlocutores em língua de sinais apresentam

comportamento semelhante de manutenção do contato visual como norma social (Gullberg & Holmqvist, 1999).

Figura 3. Ilustração do foco atencional visual e da visão periférica em interação face-a-face.

Tal fenômeno nos conduz a uma série de questões sobre o processamento dos gestos pelo sistema visual e sua integração em processos cognitivos altos com a fala, a começar pela concorrência de recursos atencionais entre as duas modalidades, uma com informação visual, a outra com informação visuo-sonora: a alocação da atenção aberta (fixação) nos gestos seria ela condicionada por mecanismos dirigidos por aspectos relacionados a estímulos no campo visual (*bottom-up*) ou fatores sociais comunicativo-pragmáticos (*top-down*)? Quando não atendidos por fixações oculares (o que efetivamente ocorre na maior parte do tempo de uma interlocução), seriam os gestos processados majoritariamente por visão periférica? Seriam eles fonte de informação complementar exploradas durante ruído no canal sonoro da fala? Quais seriam os aspectos relevantes contidos nos gestos (e em quais tipos de gestos) capazes de desafiar a norma social de manutenção do contato visual entre os interlocutores e forçarem o direcionamento do olhar para o campo gestual? Qual a região do campo gestual (central ou periférica) receberia preferência durante a fixação? Experimentos interessantes utilizando metodologia com rastreamento de fixações oculares buscaram trazer algumas respostas.

Gullberg & Kita (2009) investigaram a influência de fatores social e psico-físicos no direcionamento do olhar de ouvintes-expectadores para os gestos de um falante-

narrador. Os autores analisaram como fator social o direcionamento do olhar pelo próprio falante; e como fatores psico-físicos, a localização dos gestos no campo gestual e as pausas gestuais. Analisaram também os autores o impacto do atendimento atencional aberto aos gestos (fixação ocular dos gestos) na extração de informação e compreensão da história. Resultados mostraram que: i) há uma tendência de fixação do gesto quando falantes olham eles mesmos para seu próprio gesto; ii) falantes tendem a olhar gestos seguidos de pausa (pausas pós-golpe; ver McNeill 2005 para anatomia dos gestos); iii) a localização do gesto no campo gestual (central ou periférica) não interfere nas fixações; iv) o tempo de latência para as fixações difere quanto ao tipo de gesto: fixações de gestos com pausa pós-golpe apresentam menor tempo de latência do que fixações em gestos dirigidas pelo olhar do falante.

Gullberg & Kita (2009) concluíram de seu estudo que olhamos para gestos por razões distintas: por engajamento e atendimento social atencional (no caso de fixações dirigidas pelo olhar do falante – fator *top-down*), e por aspectos físicos intrínsecos aos movimentos gesticulatórios, que desafiam a visão periférica (fator *bottom-up*). Uma vez o gesto momentaneamente congelado, e o sistema visual periférico não ideal para processamento de imagem estática, os sujeitos tendem a olhá-lo buscando aumento de resolução espacial no processamento visual com sua projeção recaindo agora sobre a região central da retina, a fóvea. Diferentemente do que poderíamos pensar, contudo, a fixação dos gestos não impactou significativamente o desempenho dos participantes na extração de informações ou compreensão da história. A fixação ocular dos gestos seria, portanto, multifuncional, mas não poderia ser tomada como índice de alocação atencional, nem como índice de extração ou compreensão de informação veiculada na modalidade gestual.

Pereira & Haddad (2015) chegaram a conclusões convergentes com os resultados de Gullberg & Kita (2009). Estudando o papel da gesticulação para a compreensão de narrativas em língua estrangeira, os autores não encontraram diferenças significativas de desempenho na tarefa de compreensão por sujeitos com estratégias distintas de fixação (entre aqueles que mais e aqueles que menos fixaram os gestos). No entanto, resultados sugerem maior tempo de fixação dos gestos por sujeitos com nível linguístico mais avançado (7% e 8% para sujeitos com nível médio e acima, contra 4% para sujeitos com nível regular e inferior), com preferência por fixações em gesticulação icônica (11% do tempo, contra 7% a 8% para demais gestos ilustradores; ver McNeill, 1992, para classificação dos gestos ilustradores).

Recuperemos brevemente nas próximas duas seções deste capítulo alguns aspectos relevantes do sistema visual primário. Uma breve compreensão da especialização e distribuição de neurônios na retina, bem como da convergência de sinais pelas células ganglionares, pode nos ajudar a melhor compreendermos comportamentos de movimentação ocular, estratégias de alocação atencional e, por conseguinte, a percepção e integração visual de gestos.

3.2 Especializações na retina

Nossa retina é constituída por cinco tipos de neurônios: as células fotorreceptoras, as células bipolares, as células horizontais, as células amácrinas e as células ganglionares. As células fotorreceptoras, conhecidas como cones e bastonetes, são as responsáveis pela absorção de energia luminosa devido ao fotopigmento contido em seu segmento externo. As células horizontais e amácrinas trabalham para dar conta de contrastes de luminância e intensidade de luz. As células ganglionares são aquelas cujos axônios formarão um feixe, o nervo ótico, com projeções em direção a outras estruturas internas do sistema nervoso central. Para compreendermos mecanismos atencionais e de movimentação ocular, vejamos mais em detalhe a seguir a constituição e distribuição dos fotorreceptores na retina.

Cones são células fotorreceptoras com sensibilidade especializada em diferentes comprimentos de onda eletromagnética (ondas curtas, médias e longas); são células especializadas em alta resolução espacial e altamente densas na região da fóvea, área de apenas pouco mais de 1 milímetro de diâmetro da região central da retina. Bastonetes são células fotorreceptoras altamente sensíveis a variações na intensidade luminosa; estão presentes na retina em quantidade muito superior à de cones (um total aproximado de 91 milhões de bastonetes para apenas 4,5 milhões de cones na retina; Purves et al., 2004), com maior densidade na região parafoveal e decaimento progressivo em direção à periferia (**Figura 4; Gráfico 1**).

Figura 4. Diâmetro e excentricidade das regiões da fóvea, parafóvea e perifóvea na retina (adaptado de Lin et al., 2013).

Gráfico 1. Densidade de células fotorreceptoras (cones e bastonetes) na retina a partir da fóvea (excentricidade 0; adaptado de Purves et al., 2004), e decaimento progressivo da acuidade visual.

Apesar de presentes em quantidade muito inferior à de bastonetes na retina como um todo, cones são as células fotorreceptoras responsáveis pela alta acuidade visual. Por tal razão, como ilustra a **figura 4**, a acuidade visual da fóvea decai drasticamente (75% já no 6° excêntrico da linha visual; Purves et al., 2004), à medida em que, saindo-se de seu centro em direção à periferia, decai drasticamente a densidade de cones. Tal restrição da alta acuidade visual a uma área tão restrita da retina nos obriga a movermos os olhos constantemente, de forma que a projeção de estímulos de interesse no campo visual venha recair na região da fóvea, para melhor percepção e identificação (**Figura 5**).

Figura 5. Campo visual aproximado correspondente a diferentes regiões na retina (adaptado de Valenko et al., 2019).

Outro aspecto estrutural da retina que contribui para diferenças de resolução espacial do campo visual é o grau de convergência de células ganglionares. Células ganglionares respondem à ativação de ambos cones e bastonetes. Os primeiros estágios das vias (até atingirem as células amácrinas) permanecem, contudo, bastante independentes, com diferentes graus de convergência: cada célula bipolar de bastonetes recebe contatos de um grupo destes, estabelecendo contato com uma única célula amácrina; cada célula bipolar de cone, porém, recebe sinal de apenas um único receptor, estabelecendo contato com uma única célula ganglionar dominante na visão central (células ganglionares anãs) – uma resolução de um-para-um, portanto, no sistema de cones (**Figura 6**).

Figura 6. Ilustração da convergência de sinais dos fotorreceptores para as células ganglionares nas regiões periférica e central da retina. Para simplificação esquemática, não estão representadas as células horizontais nem as amácrinas. Adaptado de Bear et al. (2016)

3.2 Da retina ao córtex: duas vias do sistema visual

A via do sistema visual responsável por movimentos sacádicos, são projeções filogeneticamente mais antigas que deixam a retina em direção ao colículo superior no mesencéfalo, interconectadas com núcleos motores e pré-motores do tronco cerebral: a via retino-colicular. Já a via de processamento visual responsável pelo que chamamos de percepção consciente é filogeneticamente mais recente e mais proeminente em primatas (Milner & Goodale, 1995); são projeções que deixam a retina em direção à porção dorsal do núcleo geniculado lateral do tálamo, terminando no córtex primário (V1): via retino-geniculado.

A via retino-geniculado, por sua vez, se serve de processamentos paralelos de diferentes frequências espaciais por duas vias visuais primárias: a via parvocelular, composta por população de células ganglionares de menor porte e especializadas em processamento de frequências espaciais altas (alta resolução espacial), e a via magnocelular, composta por população de células ganglionares de maior porte, com maior campo receptor, maior velocidade de condução de potenciais de ação e mais

sensíveis a estímulos de baixo contraste. Ambas as vias parvo- e magnocelular conduzem sinais de populações de células ganglionares distintas a populações neurais localizadas em camadas específicas do córtex estriado (Bear et al., 2016).

Entretanto, além de suas especializações no processamento de sinais de diferentes frequências de e para diferentes populações neurais, evidências mostram também diferentes velocidades entre as vias: sinais de baixa frequência enviados pela via magnocelular atingiriam o córtex de 10 a 20 ms mais rápido do que sinais enviados pela via parvocelular (Cavanaugh, Bear & Movshon, 2002; Bullier, 2001; Nowak et al., 1995; Schmolesky et al., 1998; ver também Breitmeyer, 1975, Harwerth & Levi, 1978, e Parker & Dutch S., 1987 para resultados semelhantes observados no tempo de reação em humanos para estímulos visuais simples de baixa frequência). Além disso, o sistema de processamento de baixa frequência (via magnocelular) processaria com maiores precisão e velocidade a discriminação de faces humanas e o reconhecimento de emoções negativas (Goffaux & Rossion, 2006; Goffaux, 2009; Pilz, 2009). Já estímulos multi-modais complexos como gesto-fala seriam melhor processados pela via parvocelular especializada em processamento de alta frequência (Kelly et al., 2012). Há também evidências neuroanatômicas de que informação de alta frequência espacial seja principalmente processada ao longo da via ventral, que conecta o córtex estriado diretamente a, entre outras áreas, o sulco temporal superior, STS, uma região envolvida na compreensão da linguagem e na integração multimodal; Calvert, 2001, e Sekiyama et al., 2003 citados em Kelly et al., 2012; **Figura 7**).

Figura 7. Vias visuais ventral e dorsal.

Estímulos estáticos e dinâmicos contendo sinais na Língua de Sinais Americana (ASL), por exemplo, são mais facilmente identificados e mais informativos quando filtrados para alta do que para baixa frequência (Bosworth et al., 2006, e Riedl & Sperling, 1988, citados em Kelly et al., 2012); Munhall et al. (2004) relatam que pessoas identificam palavras melhor quando a apresentação de uma locução em vídeo é filtrada para banda de média a alta frequência do que para banda de baixa frequência; por técnica fMRI, Callan et al. (2004) identificaram o STS como possível local de processamento responsável pelo efeito observado. Parece que regiões multimodais no cérebro são destinos de informação visual de média a alta frequência quando estímulos consistem de faces falantes. De fato, em adição ao fato de estar implicado em processamento fonêmico de baixo nível, o STS é também conhecido por processar o significado da linguagem (Hickok & Poeppel, 2007) e de gestos manuais (Hubbard et al., 2008; Montgomery et al., 2007).

Mas quanto e que tipo de informação visual (frequência) seriam necessários ao processamento otimizado de gestos? Foi o que buscaram responder Kelly, Hansen & Clark (2012). Os autores estimularam o processamento das vias parvo e magnocelular por filtragem de frequências espaciais baixas e altas (com degradação moderada e severa da imagem) na exibição de imagens de gestos a sujeitos experimentais. Resultados mostraram que seu desempenho foi pior e mais lento para o processamento de gesto e fala na condição de degradação severa comparada com a condição moderada; o desempenho foi também mais lento no processamento de frequências baixas comparado com o processamento de frequências altas, com maior diferença na condição de degradação severa. No entanto, análises mostraram que a vantagem do processamento de altas frequências foi modulada pela amplitude e quantidade de energia (velocidade) do movimento. A pesquisa concluiu que a frequência espacial e a quantidade de energia contida nos gestos são fatores importantes para a velocidade e a acurácia da integração sensorial multimodal.

Mas para compreendermos a natureza das fixações oculares dedicadas aos gestos (bem como a sua dispensa), é necessário conhecermos mais em detalhe os diferentes aspectos que compõem movimentos gesticulatórios. Vejamos no capítulo a seguir recortes teóricos que permitem o estudo metodológico e o encaminhamento das questões centrais no estudo dos gestos.

CAPÍTULO 4

A TIPOLOGIA DOS GESTOS

Neste capítulo, faremos uma breve introdução à sistematização do comportamento não-verbal por Ekman & Friesen (1969), apontando questões e critérios que fundamentam até os dias de hoje tipologias empregadas no estudo dos gestos. Focalizaremos, contudo, neste trabalho, a gesticulação manual acompanhada da fala, os chamados gestos ilustradores (Ekman & Friesen, 1969), trazendo as principais contribuições de Adan Kendon (1980) e David McNeill (1992) para melhor pensarmos a abordagem de nosso objeto de estudo. Para a análise de nossos dados experimentais, utilizaremos a tipologia de David McNeill (1992) dos gestos ilustradores, descrita mais ao final deste capítulo.

4.1 Paul Ekman e Wallace V. Friesen e a sistematização do comportamento não-verbal

O comportamento não-verbal é consciente? Qual seu grau de deliberação e de intencionalidade? Sua natureza é expressiva, comunicativa ou simplesmente adaptativa? Como ocorre sua decodificação ou atribuição de significado, de forma universal ou culturalmente determinada? De forma sistemática ou global e intuitiva? Como o observador aprende a decodificar esses comportamentos através de pistas corporais? Como um comportamento passa a integrar o repertório pessoal?

Essas foram algumas das questões levantadas por Ekman & Friesen (1969) ao investigarem naturezas distintas de comportamentos não-verbais. Para darem conta da complexidade com que se depararam, os pesquisadores se viram na necessidade de estabelecerem múltiplas categorias desses comportamentos, distinguindo-os fundamentalmente quanto a suas circunstâncias de uso, origem, e regras de codificação e de transmissão de informação. Levando em conta esses aspectos, Ekman & Friesen (1969), inspirados nos trabalhos de Efron (1941) e de Mahl (1968), sistematizaram cinco categorias de comportamentos não-verbais que pautam diferentes tipologias e orientam

linhas de trabalhos da área ainda hoje: os *emblemas*, os *ilustradores*, as *mostras faciais de emoções*, os *reguladores* e os *adaptadores* (**Figura 8**).

Figura 8. Classificação do comportamento não-verbal por Ekman & Friesen (1969).

Os *emblemas*, chamados inicialmente pelos autores de “gesto”, são comportamentos conscientes e intencionais, deliberadamente comunicativos mas não dependentes da fala; são considerados lexicalizados por apresentarem um significado claro e bem definido compartilhado por todos os membros de um determinado grupo ou cultura. Os emblemas compõem a categoria de gestos a que associam as pessoas com maior frequência ao mencionarmos nosso objeto de estudo; exemplos bastante comuns são: o gesto de pedido de silêncio com dedo indicador perpendicular aos lábios, o sinal de “Ok” em alguns países pela união do dedo indicador com o dedo polegar, o sinal “positivo” com o punho fechado e o dedo polegar estirado ao alto, dentre outros. Gestos emblemáticos têm origem por aprendizagem em contexto cultural, sua codificação podendo ser icônica ou arbitrária, muitas vezes estilizada por conveniência do uso, facilidade de realização e discriminação como sinal comunicativo. Trazem muito mais de compartilhamento de significado do que de idiosincrasias.

Os *ilustradores* de fato ilustram o conteúdo da fala que acompanham; ditam o andamento do movimento mental com acentuações lexicais, frasais ou discursivas (*batuta*, ou gestos de pulso), esboçam um caminho ou direção lógica de pensamento

(*ideográficos*), apontam para referentes discursivos (*dêiticos*), ilustram relações espaciais (*movimentos espaciais*) e ações corporais (*cinetográficos*). São aprendidos por imitação desde cedo através do contato social, variando em qualidade em culturas distintas.

Já as *mostras* ou *expressões faciais de emoções*, apresentadas pelos autores como a consequência comportamental da emoção (ao invés da emoção ela mesma), procuram dar conta de configurações dos músculos faciais específicas a determinados estados emocionais, mais frequentemente estados emocionais primários, facilmente identificáveis em e por qualquer indivíduo de nossa espécie: a alegria, a surpresa, o medo, a tristeza, a raiva, a aversão ou nojo, e o interesse. Seriam resultantes da evolução filo e ontogenética da musculatura facial associada a atividades funcionais mais rudimentares. Com ou sem intenção comunicativa, trazem em si informações idiossincráticas importantes em uma mistura ou sequência de emoções; por tal motivo, recebem grande parte da atenção visual do interlocutor, mais do que qualquer outro comportamento ou parte do corpo, trazendo em resposta e com certa constância para a consciência do emissor seu comportamento de expressão afetiva, que, por sua vez, é encorajado a monitorá-lo, inibi-lo ou dissimulá-lo, mediante regras sociais atuantes de local, posição social e hierárquica entre os interlocutores, de idade, sexo, e outros fatores.

Os chamados gestos *reguladores*, por sua vez, procuram tratar de comportamentos frequentes pelo hábito e produzidos com baixo nível de consciência que regulam o fluxo conversacional e os turnos de fala entre interlocutores, como pedidos de atenção, de espera ou pressa, de continuação, repetição, ou elaboração, manifestações posturais de interesse ou de discrição. Não demandam grande elaboração nem atenção à sua produção; sua eliminação do contexto conversacional, entretanto, pode ser imediatamente estranhada.

Por fim, os *adaptadores* abrangem comportamentos que procuram atender ou administrar necessidades do próprio corpo, do contato interpessoal e do contato com instrumentos ou objetos. São normalmente aprendidos no período da infância como padrão adaptativo e conformados socialmente, sofrendo evoluções ao longo do tempo, com modificações graduais e reduções, sua codificação adquirindo aspectos icônicos, de forma que apenas fragmentos são mantidos por hábito ao longo da vida adulta do sujeito. Alguns exemplos de adaptadores são de toque do próprio corpo, muitas vezes do rosto ou cabelos, de objetos pessoais como canetas (estes aprendidos em atividades instrumentais como o escrever), ou comportamentos que permeiam a relação corporal com outros indivíduos, como movimentos de ataque e proteção, gestos de afeição, convite sexual,

dentre outros. Trazem em si significado psicológico e informações idiossincráticas importantes. Apesar de em sua maioria não serem deliberadamente comunicativos, são informativos e interativos. Não têm relação intrínseca com a fala, mas podem ser disparados por ela. De modo geral, não recebem retorno explícito do interlocutor. Interessantemente, Ekman & Friesen (1969) especulam que movimentos de mãos e pés que sugerem ansiedade em situação de estresse poderiam se tratar de ações residuais de movimentos mais primários em sua origem, como chutes de agressão e movimentos de fuga.

Diante do fato de que informações sobre emoções, atitudes e papéis interpessoais podem ser inferidas com razoável acurácia por observadores sem treinamento especial, os autores hipotetizaram a presença tanto de elementos culturalmente específicos quanto de elementos panculturais na expressão facial das emoções (Ekman, 1964; Ekman & Friesen, 1965, 1968), corroborando com os trabalhos de seus predecessores (Tomkins & McCarter, 1964), que encontraram consistência em aspectos universais de estados emocionais primários em nossa espécie. Descobriram, por exemplo, que o rosto é informativo da natureza das emoções, mas que gestos corporais informam sobre sua intensidade (Ekman & Friesen, 1967a); e que alguns comportamentos não-verbais podem escapar ao controle e censura de seu emissor permitindo o vazamento de informação ou dando pistas de fraude (Ekman & Friesen, 1967b).

Sempre interessados no grau de universalidade dos comportamentos não-verbais, os autores indagaram-se também sobre a natureza de informação e variação do tamanho do repertório de vocabulário emblemático entre diferentes culturas. Hipotetizaram, nessa direção, que emblemas panculturais tendem à referência ou mostra de funções corporais mais simples como o andar, sentar-se, comer, defecar, dormir, copular, enquanto que atividades humanas mais complexas relacionadas a eventos cognitivos, com referência a ferramentas ou características da ecologia local tendem a ser codificadas por emblemas específicos restritos culturalmente (Ekman & Friesen, 1969). A partir de achados de Efron (1941) da variabilidade da natureza dos gestos ilustradores com a variação da origem étnico-cultural do indivíduo, por exemplo com maior gesticulação de batuta e ideográfica para imigrantes judeus e cinetográfica e pictórica para imigrantes italianos nos EUA, Ekman & Friesen (1969) perguntaram-se sobre a constância da natureza da gesticulação no indivíduo ao longo de sua vida, apontando trabalhos que ajudaram a pensar sobre a questão: Rosenfeld (1966) teria, por exemplo, observado maior frequência de gesticulação para indivíduos que procuram aprovação de seus interlocutores.

Por estes e outros de seus trabalhos, Ekman & Friesen trouxeram à luz observações e questionamentos importantes que despertam interesse e promovem pesquisas na área ainda hoje. Apesar de bastante complexos, a tipologia proposta e os critérios utilizados pelos autores para abordarem o amplo espectro dos gestos (para a conveniência de nosso estudo, apresentamos acima apenas um resumo simplificado), núcleos de pesquisa nas áreas da psicologia, da educação, e das neurociências, para citar apenas algumas, utilizam-se largamente dos grandes cortes metodológicos propostos pelos autores.

4.2 Adam Kendon e a unidade sincrônica gesto-fala

Adam Kendon (1980), já inserido em uma longa tradição de trabalhos anteriores sobre comportamentos não-verbais (Efron, 1941; Ekman & Friesen, 1969; dentre os mais estudados), reafirmou a distinção entre a gesticulação associada à fala, desde leves balanços da cabeça, movimentos dos olhos e sobrancelhas, a movimentos amplos e complexos de mãos e braços, de outros tipos de movimentos corporais observados em interação social não necessária e diretamente relacionados ao conteúdo da fala, como por exemplo os chamados “adaptadores” (Ekman & Friesen, 1969), movimentos posturais e de toque no próprio corpo ou manipulação de objetos acessórios, como cabelos, óculos e roupas.

Diante de observações como a de Trevarthen (1977) de movimentos das mãos coordenados com movimentos de órgãos da fala presentes desde o nascimento, de Ingram (1975) de que crianças já aos três anos de idade engajam em movimentos gestuais com a mão dominante quando falando em uma conversa, bem como evidências contundentes por Kimura (1976) da alta correlação entre preferência manual durante gesticulação e lateralização linguística ¹, Kendon (1980) compreendeu tratar-se o fenômeno gesticulatório acompanhado da fala como definitivamente distinto de outros movimentos corporais não-verbais. Resolveu, portanto, explorar a intimidade da relação da gesticulação corporal com a atividade motora e a organização semântico-discursiva da

¹ Movimentos acompanhados da fala tidos como gesticulatórios são produzidos predominantemente pela mão direita em sujeitos destros, enquanto que movimentos de toque do próprio corpo são produzidos igualmente por ambas as mãos; sujeitos canhotos tendem a gesticular com ambas as mãos – a linguagem é compreendida como mais bilateralizada em sujeitos canhotos do que em sujeitos destros (Kimura, 1976).

fala, compreendendo-a, não como de dependência, mas como de manifestações do mesmo processo de elocução.

Emprestando da metodologia de trabalho de estudos anteriores (Condon & Ogston, 1966; Kingdon, 1958), Kendon (1980) se debruçou sobre os detalhes da sincronia entre as estruturas frasais expressas no movimento corporal gesticulatório e as estruturas frasais expressas nos níveis de organização da fala. Para a análise da estrutura da gesticulação, usa o conceito de *frase gesticular*, marcada por um *pico de esforço* (Dell, 1970) ou golpe (*stroke*) durante o movimento de mãos e braços entre duas *posições de descanso* consecutivas. O golpe, por sua vez, é precedido de uma fase preparatória, em que o membro deixa a posição de repouso, e sucedido por uma pausa (pausa pós-golpe), antes de retornar à posição de repouso já em fase de recuperação (ou retração), ou antes de se preparar para novo golpe (**Figura 9**).

Figura 9. Anatomia do gesto conversacional (a partir de Kendon, 1980).

Kendon (1980) observou que as unidades melódicas de entonação da fala, segmentadas em quatro níveis (locuções, grupos locucionais, aglomerados locucionais e discurso) e marcadas por pausas, mudanças na qualidade, intensidade ou altura da voz, bem como na postura do falante, são acompanhadas em cada um dos quatro níveis por padrões distintos de movimento corporal de braços, mãos e pulsos, cabeça e postura em geral. A coocorrência entre as unidades melódicas de proeminência vocal na fala e os núcleos de ênfase cinésica nas frases gesticulares levou Kendon a pensar a gesticulação, não como ação disjunta, mas pertencente a um único e mesmo processo de elocução manifesto simultaneamente nos órgãos vocais e nos movimentos corporais. Um dos achados importantes nas análises de Kendon (1980) que sustentam com contundência sua hipótese foi o de que frases gesticulatórias tendem a emergir pouco antes do segmento da fala no qual a mesma ideia é codificada (também chamado na literatura por *afiliado lexical*); a fase preparatória do gesto ocorrendo ainda mais cedo.

Mas, não bastasse o paralelismo melódico-enfático estrutural das duas modalidades, observou Kendon (1980) que os movimentos gesticulares também se relacionam claramente com o conteúdo da fala, embora assumam aspectos de codificação distintos dos dela; enquanto que a fala se serve de formas discretas altamente convencionalizadas e hierarquizadas gramaticalmente, gestos codificam por uma sucessão relativamente livre de atos regidos pela ordem de apresentação das ideias no discurso.

Os resultados de seus e de outros estudos (como o trabalho de Butterworth & Beattie, 1978) sugerem fortemente a Kendon que os processos que originam a produção da fala e da gesticulação comecem a um único e mesmo tempo, no processo de elocução, pela manipulação semântica de representações mentais de imagens e ações sobre objetos reais, e não na organização de formas derivadas apenas de linguagem verbal; mais ainda, que o fato da gesticulação tender a antecipar a fala e por vezes avançar continuamente mesmo quando esta é interrompida possa ser uma pista de que o canal gestual seja mais fácil e prontamente recrutado por a representação de ideias em movimentos gesticulatórios ser mais primitiva, enquanto que a fala, por outro lado, demandaria tratamentos hierárquicos mais complexos e mais distendidos no tempo.

Em suma, a gesticulação seria, portanto, assim como a fala, a manifestação aberta resultante do aumento do nível de energia engajado em uma elocução.

4.3 David McNeill e a tipologia dos gestos ilustradores

A partir de mais de dez anos de estudos experimentais com resultados empíricos de análises minuciosas de discursos narrativos induzidos experimentalmente em sua maior parte, mas também de emissões de televisão e conversas filmadas, McNeill (1992) elaborou em níveis semânticos e pragmáticos as contribuições de Kendon (1972, 1980) sobre a relação sincrônica dos sons da fala com os movimentos gesticulatórios. McNeill ilustrou essa relação apresentando o que chamou de “*continuum* de Kendon” – originalmente quatro *continua* que podemos aqui sem prejuízo resumir para nossa conveniência em um único apenas – em que dispõe quatro qualidades marcadamente distintas de uso dos gestos: a *gesticulação*, a *pantomima*, os *emblemas* e a *língua de sinais*, nessa ordem dispostos para explicitar o caráter unitário e dialógico do sistema gesto-fala (**Figura 10**).

Figura 10. Continuum dialógico do sistema gesto-fala. Adaptado de McNeill (2005).

Ao avançarmos entre as categorias, observamos o declínio progressivo do acompanhamento pela fala e o aumento proporcional de propriedades linguísticas presentes nos gestos. Evidenciam-se, portanto, as extremidades do contínuo: a *gesticulação* acompanhada da fala, com movimentos idiossincráticos espontâneos de representação informal de imagens mentais globais, e, do outro, a *língua de sinais*, com representações linguísticas analíticas hierárquicas formais convencionalizadas como de qualquer língua falada ou sinalizada. McNeill delimitará assim seu objeto de investigação no primeiro dos extremos, a gesticulação acompanhada da fala.

Inspirado dos trabalhos de Vygotsky (1962, 1978) sobre modelos de mecanismos de processamento psicolinguístico da informação, McNeill emprestou do conceito vygotskyano de “predicado psicológico” para construir em sua teoria o conceito de “ponto de crescimento” (*growth point*). Compreendido como o nó inicial de uma elocução (*utterance*), o ponto de crescimento McNeill (1992) define como uma saliência mínima dentre os fragmentos desordenados de imagens e categorias linguísticas, de resíduos de pensamentos precedentes, a partir dos quais, motivada por forças disruptivas extracognitivas de necessidades afetivo-volitivas movimenta a composição verbo-gesticular da elocução (*utterance*) com direcionamento global e sequencial. Gestos emergiriam como um estágio primitivo da elocução, com representações globais idiossincráticas simbólicas do imageamento mental; a fala, como um estágio final público e socialmente regulado por unidades analíticas mínimas, lineares e hierárquicas.

No entanto, por conta de a fase preparatória do gesto anteceder a fala (por vezes mesmo o golpe gestual precede o pico fonológico silábico da fala), McNeill (1992) especula que a unidade mínima de ideia, ou nó inicial gerador, tome forma global antes de emergir em qualquer uma das duas modalidades. Já o contraste entre as duas modalidades semióticas ativas na experiência mental do falante geraria uma instabilidade cognitiva que, por sua vez, propulsionaria o sistema de significação. Nesse contexto, os momentos de convergência e estabilidade cognitivo-estrutural momentânea entre as duas

modalidades seriam marcados pelo pulso rítmico a elas subjacente. A materialização do sentido viria não em representação, mas na atualização do estado cognitivo momentâneo do falante (e do ouvinte-expectador, como consequência); a comunicação se dando não como troca de sinais, mas como coabitação de mesma morada cognitiva.

Apresentamos a seguir a classificação proposta por McNeill (1992) dos gestos que acompanham a fala mais recorrentes no material por ele analisado, tipologia mais frequentemente estudada na literatura desde então, e que adotaremos nós também para análise de nossos dados experimentais; são eles: os gestos *icônicos*, os *metafóricos*, os *de pulso* (*beat* ou *batuta*) e os *dêiticos* (**Figura 11**); todos os quatro relacionados com aspectos semânticos e/ou pragmáticos da fala que acompanham.

Figura 11. Tipologia dos gestos ilustradores por McNeill (1992).

Gestos *icônicos* são representações materiais mais transparentes de referentes concretos, tais como objetos, formas, personagens narrativos, ações, espaços, etc. Gestos *metafóricos*, também pictóricos, são representações materiais de conteúdo abstrato, tais como estados emocionais ou relações afetivas, processos e conceitos. McNeill (1992) observa que gestos icônicos tendem a predominar no gênero narrativo; enquanto que gestos do abstrato (metafóricos, de pulso e dêiticos) tendem a predominar em conversações e palestras.

Gestos de pulso e dêiticos, por sua vez, não são enquadrados por McNeill como pictóricos. Gestos *de pulso* são descritos por ele como movimentos bifásicos simples (dentro/fora, cima/baixo) em que uma ou ambas mãos se movem ritmicamente com o pulso prosódico da fala, acentuando junto com esta palavras, orações, encerramento de ideias e argumentos, desempenhando, portanto, função pragmática no discurso.

Gestos *dêiticos*, por fim, são gestos de apontamento, apontamento para referentes concretos ou abstratos, presentes ou ausentes, no próprio ambiente ou no espaço narrativo ficcional construído no espaço gestual logo à frente do falante e entre ele e seu interlocutor. Para McNeill (1992), analisando-se os gestos que acompanham a fala, os gestos ilustradores, é possível contemplar a estrutura discursiva mais abrangente.

É importante ressaltarmos, contudo, que a classificação dos gestos ilustradores em incônicos, metafóricos, dêiticos e de pulso leva em consideração sua dimensão preponderante (mas não única) e se justifica por fins analíticos. Empiricamente, entretanto, um único gesto pode – e frequentemente apresenta – aspectos que nos remetem a mais de uma das categorias (Goldin-Meadow, 2003), fenômeno que explica dentre outras razões a complexidade na codificação da gesticulação.

Vejamos a seguir, no próximo capítulo, modelos cognitivos que buscaram dar conta do fenômeno de sincronia acomodando a produção de gestos de forma integrada à produção da fala.

CAPÍTULO 5

MODELOS COGNITIVOS DE PRODUÇÃO INTEGRADA GESTO-FALA

Neste capítulo abordaremos dois modelos cognitivos bastante referidos na literatura que procuram explicar o fenômeno da sincronia da produção de gestos com a produção da fala. Os modelos procuram integrar unidades de armazenamento e de processamento de informação conceitual, linguística e gestual. Além de nos ajudarem a visualizarmos esquematicamente aspectos da produção de gestos ilustradores que abordamos até aqui, ajudam-nos também a localizarmos alguns dos principais dilemas no tratamento teórico de gestos produzidos com a fala.

5.1 Sketch Model (De Ruiter, 2008)

Procurando explicar o fenômeno de coocorrência gesto-fala (Hadar, 1989; Morrel-Samuels & Krauss, 1992; Krauss, 1998), De Ruiter (2008) propôs um modelo cognitivo de produção de gestos a partir de adaptações no modelo de arquitetura de processamento da fala por Levelt (1989). Do modelo de Levelt, De Ruiter conservou os três módulos de produção da fala: o *conceitualizador*, o *formulador* e o *articulador*. O *conceitualizador* recorre a conhecimentos declarativos e procedurais para construir uma estrutura conceitual com uma série de especificações semânticas em uma intenção comunicativa; o *formulador* recebe a mensagem pré-verbal do conceitualizador, mapeia no léxico mental o conceito a ser lexicalizado e formula um plano fonético. O conjunto de instruções segue para o *articulador*, que lhe atribui, então, um plano prosódico adequado à elocução. O produto do estágio articulatório é a explicitação da fala.

Em um dos módulos, porém, no *conceitualizador*, originalmente no modelo de Levelt (1989), responsável pela geração de uma mensagem pré-verbal, De Ruiter (2008) incluiu uma segunda tarefa: a geração de um esquema geral (*sketch*) a ser enviado à *unidade planejadora de gesto*. Assim, o *conceitualizador* dividiria a intenção comunicativa em duas partes: a parte proposicional a ser transformada em mensagem pré-

verbal e enviada ao *formulador*, e a parte imagética a ser transformada em um esboço geral (*sketch*) a ser enviada ao planejador de gesto (**Figura 12**).

Figura 12. *Sketch model*, adaptado de De Ruiter (2008). Conforme o modelo de De Ruiter (2008), caixas representam processos e elipses representam armazenamento de conhecimento. Setas representam representações enviadas de processo a outro, e linhas pontilhadas representam acesso a um armazenamento de conhecimento específico.

O esquema geral (*sketch*) seria então enviado do *conceitualizador* ao *planejador de gesto* para elaboração de um programa motor e envio subsequente deste a unidades de controle motor mais baixas. O *planejador de gesto* teria acesso ao *gestuário* (*gestuary*), a esquemas motores procedurais (*schemata*), e a informações sobre o ambiente, e seria responsável dentre outras tarefas por: atribuir o movimento a uma determinada parte do corpo (mão esquerda, direita, braços, ombros, pés ou pernas, ou ainda a cabeça), adequar

o movimento a circunstâncias ambientais, tais a amplitude do movimento em função do espaço disponível, e – esta mais complexa – fundir diferentes gestos.

No modelo de De Ruiter (*sketch model*, 2008), o *gestuário* seria um repositório de informações gestuais tais formas e orientação de gestos, culturalmente convencionalizadas, indexadas pelo conceito que representam. Tanto o *conceitualizador* quanto o *planejador de gesto* recorreriam ao *gestuário* na geração de gestos convencionalizados, tais gestos emblemáticos e dêiticos. No caso de gestos representacionais espontâneos, tais gestos icônicos e metafóricos, sua forma seria em grande parte determinada por um conteúdo imagético com origem na memória operacional. Seria tarefa do *conceitualizador* extrair do conteúdo imagético a informação relevante de representação espaço-temporal, e criar uma representação esquemática (*sketch*) que possa ser transformada mais adiante pelo *planejador de gesto* em um programa motor. Uma vez finalizado o programa motor, este seria enviado pelo *planejador de gesto* a unidades de controle motor mais baixas, resultando na explicitação do movimento.

O modelo de De Ruiter (2008) explica, portanto, a sincronia semântica gesto-fala propondo ambas realizações derivarem de uma mesma intenção comunicativa. A sincronização temporal global entre elas seria consequência da mensagem pré-verbal e do modelo geral (*sketch*) serem gerados aproximadamente a um mesmo momento. Já a sincronização da fala com gestos dêiticos, notoriamente mais precisa, viria, conforme o modelo de De Ruiter (*ibid.*), de um sinal da unidade motora ao codificador fonológico.

Contudo, o modelo de De Ruiter (2008) não explica determinados aspectos da produção de gestos acompanhando a fala: gestos de pulso (ou batuta), por exemplo, um padrão gesticulatório que se observa como de maior ocorrência, não foram contemplados em seu modelo. De Ruiter (*ibid.*) admite ele próprio a sincronização gesto-fala se tratar de uma questão bastante problemática, explicando a ausência de gestos de pulso em seu modelo por insuficiência de conhecimento sobre a previsibilidade de sua ocorrência – McNeill (1992) explica tratar-se de gestos servindo a funções metanarrativas, cuja ocorrência não pode ser indexada a nenhum índice convencional, como acentuação tônica silábica ou frasal, unidade lexical ou classe gramatical específica.

Como alternativa à problemática, De Ruiter (2008) faz uma proposição interessante, uma substituição conceitual que nos parece bastante produtiva. Considerando-se que tanto unidades gestuais quanto unidades da fala têm duração temporal, propõe De Ruiter (*ibid.*), no lugar de *afiliação lexical* ao gesto, uma *afiliação*

frasal. Tal mudança na abordagem permitiria melhor encaminhamento da questão da sincronização gesto-fala por se tratar da sincronização, não de duas unidades, mas de dois intervalos de tempo cujos limites podem, por vezes, serem difíceis de se precisar com exatidão.

De qualquer forma, a antecipação do início da fala pelo início do gesto (*gesture onset*) se explicaria pelo fato do *conceitualizador* enviar a mensagem pré-verbal ao *formulador* apenas após o *planejador de gesto* ter finalizado o programa motor; bem como pela geração do gesto levar menos tempo que a geração da fala, esta última demandando adequações propositivas com tempo de processamento adicional ao tempo de processamento dos gestos.

5.2 O modelo de Krauss, Chen & Chawla (1996)

Já outro modelo teórico interessante de produção integrada gesto-fala, proposto por Krauss et al. (1996), conseguiu acomodar o aspecto importante da natureza de gestos conversacionais que o modelo de De Ruiter (2008) não foi capaz: o fenômeno dos gestos de pulso. Krauss et al. (ibid.) pensaram a adequação do modelo de processamento de gestos a partir de duas naturezas claramente distintas dentre os gestos conversacionais quanto ao seu grau representacional ou de lexicalização: os *movimentos lexicais* e os *movimentos motores* (**Figura 13**).

Figura 13. Contínuo do grau de lexicalização dos gestos. Adaptado de Krauss et al. (1996)

Por *movimentos motores*, os autores definiram um tipo de gesto conversacional simples com forma fixa, repetitivo e em movimentos ritmados coordenados. Por *movimentos lexicais*, os autores definiram movimentos das mãos que variam consideravelmente em comprimento, não repetitivos, complexos e que mudam de forma, aparentemente relacionados com o conteúdo semântico da fala que os acompanha.

A pressuposição central dos autores é a de que movimentos lexicais seriam gerados a partir de representações do conceito originário na memória de curto-prazo, sobretudo de características dinâmico-espaciais, enquanto que movimentos motores parecem ser antes produto da produção da fala, pois sua coordenação com a prosódia da fala sugere que não poderiam ter sido planejados antes do plano fonético de elocução. Movimentos motores se originariam, assim, em um estágio relativamente tardio do processo de produção da fala (**Figura 14**).

Figura 14. Modelo de processamento fala-gesto adaptado de Krauss et al. (1996). Caixas representam processos e elipses representam armazenamento de conhecimento. Setas representam representações enviadas de processo a outro, e linhas pontilhadas representam acesso a um armazenamento de conhecimento específico.

Krauss et al. (1996) defendem que ambos movimentos motores e lexicais tenham um papel na produção da fala e que esta seja de fato sua função primeira. Argumentam que gestos lexicais teriam efeito sobre a codificação gramatical, por facilitação da busca

no léxicon mental por lemmas (radicais semânticos) que satisfaçam especificações da mensagem pré-verbal, seja por pré-ativação cross-modal (*priming*), seja por “congelarem” (*hold*) propriedades conceituais na memória durante a busca lexical. Voltaremos ao acesso lexical por gestos no capítulo 5 ao abordarmos em maior detalhe a atuação da gesticulação na cognição.

5.3 Por um modelo integrado de processamento de informação

Apesar das contribuições de modelos de processamento, como os acima descritos de De Ruiter (2008) e de Krauss et al. (1996), ainda nos é de fato desconhecido o processo através do qual gestos se originam. Ademais, abordagens por modelos de processamento de informação são passíveis de enganos, levando a possíveis equívocos na produção de novos conhecimentos.

Não é desejável, no entanto, que avanços na produção de conhecimentos neurocientíficos falsem descrições funcionais ultrapassadas e guiem reformulações mais refinadas ou mesmo a construção de novos modelos alternativos? Modelos de processamento de informação pressupõem a redução de incerteza sobre um número de alternativas (Shannon & Weaver, 2008) em uma construção teórica formal; além disso, têm a vantagem de serem colocados à prova de coerência interna ou inconsistências por simulações computacionais. O modelo geral de processamento de informação de De Ruiter (2008) e de Krauss et al. (1996) acomodam uma gama de conhecimentos acumulados sobre a produção de gestos, e podem ser usados para explicarem ou predizerem de forma consistente alguns dos fenômenos dentro de um enquadramento conceitual teórico; podem, portanto, serem compreendidos como ferramenta importante enquanto ponto de partida na investigação de processos e representações envolvidas na produção de gestos acompanhando a fala.

Nessa direção, é válido pensarmos em um modelo teórico expandido que integre, não apenas a produção de gestos com a produção da fala no ato da elocução, mas também os módulos e processos analíticos e integrativos das informações das duas modalidades no ato da compreensão e da interlocução. Lançaremos novamente mão dos modelos cognitivos de produção integrada gesto-fala mais ao final do trabalho, nos capítulos em

que discutiremos nossos estudos experimentais, para melhor compreendermos os fenômenos observados e especularmos sobre os resultados de nossas análises.

Vejamos a seguir implicações dos modelos teóricos e de resultados experimentais que têm alimentado na literatura grandes discussões sobre a funcionalidade dos gestos.

CAPÍTULO 6

GESTOS NA COGNIÇÃO: GESTOS PARA QUEM?

Neste capítulo, abordaremos a problemática central do estudo dos gestos: a função da gesticulação. Traremos em revisão alguns dos trabalhos com contribuições das mais significativas para o atual estágio do debate. A apresentação e discussão desses trabalhos nos ajudará a melhor circunscrever nossa problemática de investigação e a situar as contribuições de nossos estudos experimentais para o debate na literatura.

6.1 A função dos gestos na produção da fala: uma abordagem incorporada da cognição

Como vimos até aqui, abordagens teóricas têm buscado explicar como padrões sensório-motores de movimentos espontâneos integram a cognição e a comunicação humanas. Uma abordagem que vem ganhando importância dentre pesquisadores psicólogos, filósofos e neurocientistas é a abordagem incorporada da cognição humana, que enfatiza a capacidade do corpo em moldar nossa percepção e nossa consciência (Gibson, 1979; Clark, 1998, 2008; Hurley, 1998; Nöe, 2004, 2009; Gallagher, 2005; Rowlands, 2010; Shapiro, 2011; Zadravko, 2013). Nessa abordagem, gestos manuais têm atraído especial interesse dos pesquisadores.

Como propõem os modelos cognitivos de produção integrada gesto-fala de De Ruiter (2008) e Krauss et al. (1996) descritos no capítulo anterior, a linguagem verbal não é o único suporte de construção de sentido na elaboração de representações mentais. Quando falamos sobre objetos espaciais ou eventos de movimento, usamos gestos representacionais para ativar, manter e manipular informação espacial motora (Hostetter & Alibali, 2019; Kendon, 2017; Smithson & Nicoladis, 2014; Kita et al., 2017; Chu & Kita, 2011; McNeill 2005).

Mas para além de uma modalidade visuo-espacial de manipulação de representações mentais e elaboração de sentido, gestos atuam diretamente em processos mentais desempenhando papel importante na economia cognitiva.

6.1.1 A carga cognitiva e a memória operacional

Beilock & Goldin-Meadow (2010) observaram alterações cognitivas em representações mentais promovidas por gesticulação durante explicações de resolução de problema. Em uma série de 2 experimentos, as pesquisadoras observaram que o modo como sujeitos representam em gestos explicações sobre suas ações em resolução a uma tarefa-problema motora (Torre de Hanoi) tem impacto direto em desempenho subsequente da mesma tarefa. As autoras propõem que gestos não apenas simulem ações motoras como também sejam ações eles próprios capazes de incrementar ou modificar representações mentais.

Cooperrider, Wakefield & Goldin-Meadow (2015) replicaram o experimento de Beilock & Goldin-Meadow (2010) para investigarem, não mais o benefício ou alterações cognitivas promovidas pela gesticulação, mas sim qual tipo de feedback (visual, proprioceptivo ou ambos) seria responsável por promover os efeitos observados. Para tanto, controlaram o feedback visual dos gestos durante a fase experimental de explicação de resolução da tarefa motora; os sujeitos do grupo experimental foram impedidos de visualizarem seus próprios gestos por uma barreira física instalada à altura de seus ombros. Observaram-se os mesmos resultados, semelhantes ao do experimento de Beilock & Goldin-Meadow (ibid.), apesar da eliminação do retorno visual dos gestos. Interpretaram seus resultados Cooperrider et al. (ibid.) sugerindo que, conforme argumentam abordagens incorporadas da cognição, gestos modulam processos cognitivos por retorno proprioceptivo.

Outro trabalho bastante importante nessa linha de investigação foi o estudo conduzido por Goldin-Meadow, Nusbaum, Kelly, & Wagner (2001), que explorou o impacto da gesticulação manual na memória operacional de crianças e adultos durante tarefa de resolução de expressões matemáticas. Após a resolução de cada expressão, sujeitos experimentais foram apresentados a uma lista de palavras (no caso das crianças) ou sequência de letras (para os adultos) para lembrança posterior. Em seguida, os sujeitos deveriam explicar a solução utilizada por eles para a resolução da expressão matemática e, logo depois, evocar a lista de palavras ou sequência de letras. As experimentadoras controlaram a gesticulação durante a fase de explicação, pedindo aos participantes que mantivessem as mãos paradas sobre a mesa na condição *sem gestos*; controlaram ainda

em dois níveis (curta e longa) a extensão dos itens apresentados para lembrança posterior. Resultados mostraram que ambos adultos e crianças lembraram-se de significativamente mais itens na condição *com gestos* quando a memória operacional era mais exigida pela apresentação de lista longa de itens. Interessantemente, em mesma direção apontada por outros trabalhos referidos mais adiante, as pesquisadoras se depararam com tempo ligeiramente maior gasto durante a fase de explicação da resolução matemática na condição experimental *com gestos*, para ambos os grupos de crianças e adultos. As pesquisadoras concluíram de seus resultados que gestos – apesar de se servirem de maior tempo para a realização da tarefa – diminuíram a demanda cognitiva durante a explicação da resolução ao problema matemático, resultando na liberação de recursos cognitivos disponíveis para a tarefa de memória simultânea.

Resultados como o de Goldin-Meadow et al. (2001) de melhora de desempenho em tarefas de resolução de problema promovida pela gesticulação são explicados na literatura de abordagem incorporada da cognição pela transferência de alocação de recursos de um armazenamento verbal para um armazenamento espacial; ao liberar recursos lineares para a produção da fala, o falante teria seu esforço cognitivo diminuído e sua economia cognitiva potencializada. Fenômeno semelhante foi reportado por Pouw et al. (2016) ao investigarem relações entre benefícios de desempenho promovidos pela gesticulação e memória operacional do sistema visual. Resultados do estudo mostraram que, ao produzirem gestos durante tarefa de resolução mental de problemas, sujeitos produzem menor quantidade de movimentos oculares. Gestos seriam, portanto, uma maneira de indexar no espaço simulações mentais, reduzindo a necessidade de projeções visuais.

Depreendemos, portanto, dos trabalhos acima mencionados que gestos, para além simularem ações, permitiriam, por retorno proprioceptivo, a minupalação e alteração de representações mentais; ademais, atuariam na memória operacional, promovendo alívio da carga cognitiva durante a produção da fala por realocação de recursos para um armazenamento espacial, bem como durante tarefa de resolução de problemas por indexação espacial de simulações mentais. Notemos, contudo, que o paradigma de controle experimental utilizado por Goldin-Meadow et al. (2001) coloca aos sujeitos experimentais uma condição de dupla tarefa: sendo o uso dos gestos o padrão natural da comunicação, a instrução à não-gesticulação impõe um desvio atencional e uma sobrecarga cognitiva, obrigando os sujeitos experimentais a um controle atencional de

inibição motora dos gestos. Seria, portanto, no estudo de Goldin-Meadow (ibid.), de fato esperada uma diminuição de desempenho na tarefa de memória.

Em nosso estudo experimental de número 1 reportado na seção experimental deste trabalho, propomos um controle do fator gesticulação que dispensa a instrução à não-gesticulação, contornando o problema do desvio atencional e da sobrecarga cognitiva. Esperamos trazer a partir de nosso estudo alternativas mais ecológicas para o estudo dos gestos. Mais detalhes serão apresentados e discutidos no momento oportuno na discussão do referido experimento.

6.1.2 O acesso lexical

Outro aspecto relevante da atuação da gesticulação na cognição que merece menção e reflexão é sua contribuição para o acesso lexical durante a produção da fala.

Rauscher, Krauss & Chen (1996) analisaram o papel de movimentos corporais relacionados à produção de conteúdo lexical. Os pesquisadores controlaram a produção de gestos em tarefa de descrição de desenhos animados; manipularam também a dificuldade de acesso lexical colocando restrições ao uso de léxico específico. Hipotetizaram que, se gestos manuais que acompanham espontaneamente a fala emergem para facilitação do acesso lexical, sua inibição deveria produzir efeitos semelhantes aos observados por restrições lexicais. Conforme hipotetizado pelos autores, os efeitos observados na fala pela eliminação dos gestos, ao menos quanto a conteúdos tratando de relações espaciais, assemelharam-se àqueles observados pelo aumento da dificuldade de acesso lexical nas produções verbais. Os autores concluíram seu estudo corroborando a hipótese de que gestos que acompanham espontaneamente a fala tenham implicações em sua produção por facilitarem acesso ao léxico mental. Movimentos lexicais derivariam, portanto, de conhecimento codificado espacialmente, cujas características semânticas participariam da busca lexical durante a codificação gramatical.

Morsella & Krauss (2004), outro trabalho bastante referido na literatura, também procuraram colocar à prova a hipótese de acesso lexical dos gestos. Investigaram a quantidade de gestos produzidos durante a descrição de imagens de objetos, com estímulos consistindo em imagens de objetos de fácil identificação, portanto nomeáveis (casa, flor, sorvete, entre outros), e de objetos não-identificáveis, portanto não-nomeáveis (objetos sem sentido). Controlaram também a presença da imagem do objeto no campo

visual e a complexidade dos objetos não-identificáveis. O estudo revelou que pessoas gesticulam mais quando descrevendo objetos não disponíveis no campo visual (de memória), objetos de difícil codificabilidade verbal, e objetos complexos (de difícil lembrança). Quando impedidos de gesticularem, os participantes no experimento de Morsella & Krauss (ibid.) apresentaram disfluência na fala, mesmo com o estímulo disponível no campo visual. Diante dos resultados, os autores do trabalho concluíram que gestos afetam ambas memória espacial e evocação lexical.

Estudos como o de Rauscher et al. (1996) e Morsella & Krauss (2004) nessa linha de pesquisa têm proposto a facilitação do acesso lexical durante a produção da fala por pré-ativação motora (*priming*) a partir de movimentos corporais gestuais. Com efeito, assim como observado por Kendon (1980), Hadar (1989) também notara que gestos icônicos precedem o afiliado lexical da fala; Morrel-Samuels & Krauss (1992) observaram, ademais, que essa assincronia se correlaciona negativamente com a familiaridade do termo lexical evocado. Diante de tais evidências, Krauss (1998) sustenta a facilitação do acesso lexical como explicação principal da emergência dos gestos. Gestos, portanto, seriam movimentos corporais espontâneos produzidos no ato da fala diretamente atuantes na memória operacional e na facilitação do acesso lexical. Ademais, enriqueceriam representações neurais lexicais por torná-las multimodais pelo acréscimo de traços motores (Engelkamp & Zimmer, 1985; Nilsson, 2000; Masumoto et al., 2006; Tellier, 2008; Macedonia et al., 2011).

Entretanto, apesar da hipótese de facilitação do acesso lexical por gestos encontrar respaldo em outros trabalhos na literatura, nem sempre os resultados foram convergentes (verificar Graham & Heywood, 1975, e Lickiss & Wellens, 1978). Parte da explicação da divergência entre resultados relatados na literatura quanto ao efeito da restrição de gestos sobre a produção da fala pode estar em observações mais pontuais dos trabalhos experimentais, desde a caracterização e tamanho da população amostral, ao paradigma utilizado no desenho e controle experimentais, até em resultados observados por análises qualitativas. Como exemplo, voltemos rapidamente aos trabalhos de Morsella & Krauss (2004) e Rauscher, Krauss & Chen (1996).

Os resultados reportados por Morsella & Krauss (2004) são passíveis da mesma crítica apontada por nós em seção anterior sobre o estudo de Goldin-Meadow et al. (2001) quanto ao paradigma de controle experimental do fator gesticulação. No trabalho de Morsella & Krauss, para a condição experimental, foram instalados falsos eletrodos sobre ambos antebraços dos participantes, explicando-lhes que qualquer movimento dos

membros poderia prejudicar a qualidade dos sinais de condutância galvânica da pele, que seriam coletados como índice de resposta emocional à tarefa. Nitidamente, observamos também no controle experimental de Morsella & Krauss (ibid.) a adição de uma segunda tarefa com sobrecarga cognitiva atencional de controle inibitório dos gestos.

Quanto ao estudo de Rauscher et al. (1996), ademais os resultados anteriormente mencionados, algumas observações merecem ponderação. Os autores reportam que falantes gesticularam em média 51% do tempo de fala nas passagens com conteúdo espacial, e apenas 17% nas demais. Encontraram ainda os pesquisadores efeito de disfluência na fala na condição sem gestos, mas apenas para passagens com conteúdo espacial. A eliminação dos gestos também aumentou a frequência relativa de pausas preenchidas, novamente, apenas em falas com conteúdo espacial². Chamamos, assim, a atenção para as limitações dos resultados experimentais reportadas no estudo, limitações que encontrarão paralelo também nos estudos descritos na seção a seguir investigando a efetividade comunicativa dos gestos e seu impacto na compreensão da fala.

O uso de gesticulação pode, portanto, ter papel especialmente importante na representação de aspectos visuais relacionados a tamanhos e formas, sobretudo quando não usuais, assim como de posição e comparação; gestos podem efetivamente auxiliar o tratamento de informação e simplificar a tarefa verbal explicativa. Não obstante, as limitações dos resultados apontadas nos trabalhos não nos permitem assumir de forma generalizada que sempre haverá uma facilitação do acesso e do tratamento lexical por gestos.

Voltaremos à discussão do acesso lexical por gestos mais adiante quando tratarmos de produções narrativas.

6.2 A função comunicativa dos gestos

Não obstante todo corpo de evidências trazidas à luz pela linha de pesquisa que investiga a hipótese da atuação dos gestos na produção da fala, a questão fundamental sobre a função da gesticulação, se serviria primeiramente ao falante na manipulação de

² O número de pausas preenchidas não-gramaticais (*nonjuncture filled pauses*, chamadas *hesitações* em nosso estudo experimental de número 1 descrito na seção experimental deste trabalho) são pausas com vocalização (“hm...”, “é...”, “ah...”, etc) distribuídas na fala em localizações outras que não entre unidades discursivas como orações e períodos, onde são normalmente esperadas. Pausas preenchidas são uma medida sensível à dificuldade de evocação lexical.

representações mentais imagéticas para elaboração de significados ou ao ouvinte-expectador como uma modalidade semiótica visuo-espacial na construção de sentido no ato comunicativo, suscita ainda hoje controversos debates e a produção de novos conhecimentos. Algumas questões como os limites da efetividade comunicativa de gestos conversacionais, bem como alterações em sua produção por influência da presença/ausência do interlocutor-expectador permanecem abertas. Vejamos alguns aspectos comunicativos da produção de gestos.

6.2.1 O empacotamento da informação

Em acompanhamento ao fluxo verbal, naturalmente segmentado em unidades semântico-gramaticais (Goodwin, 2001), observa-se o desenvolvimento de movimentos corporais cadenciados (Kendon, 1980); sujeitos usam também suas mãos para delimitar e manipular unidades de significado, marcar e agrupar informações no esforço de construção do sentido que desejam compartilhar (Gregersen, Olivares-Cuhat & Storm, 2009). Gestos marcam, por exemplo, status de informações discursivas, distinguindo ou salientando a introdução de informação nova em contraste com informação previamente introduzida ou de conhecimento comum entre os interlocutores (*common ground*; Goodwin, 2001).

Nessa linha, Alibali, Kita & Young (2000) sugerem como a principal função da produção de gestos o empacotamento de informação. A hipótese do empacotamento de informação propõe que a produção de um gesto representacional ajudaria falantes a organizarem conteúdo rico em aspectos motores espaciais em pacotes informacionais adequados à fala; gestos auxiliariam, portanto, o falante em sua produção, mas sua função ou motivação principal viria da intencionalidade comunicativa, de sua contribuição para a compreensão e construção de sentido pelo sujeito ouvinte-expectador.

Para testarem sua hipótese, Alibali, Heath & Meyers (2001) investigaram possíveis efeitos de visibilidade do ouvinte-expectador no número de gestos representacionais (icônicos, metafóricos e dêiticos) e de pulso (ou batuta) produzidos pelo falante. Resultados mostraram que gestos de pulso não sofreram alterações entre as condições de visibilidade e invisibilidade do ouvinte-expectador. Já gestos representacionais, apesar de sua alta taxa de ocorrência para ambas condições experimentais, esta foi significativamente maior na condição de visibilidade do ouvinte-

expectador. Os pesquisadores reconhecem, portanto, a permanência de gestos mesmo na invisibilidade do interlocutor; contudo, admitem que a finalidade da gesticulação seria em última análise a comunicação. Gestos se manteriam mesmo na ausência de interlocutores por razões de hábito, ou pela construção de um ouvinte-expectador imaginário receptor da informação.

6.2.2 O número e o posicionamento de ouvintes-expectadores

Também Özyürek (2002) buscou evidências experimentais da função comunicativa dos gestos; investigou a influência, não da visibilidade do ouvinte-expectador como Alibali, Heath & Myers (2001), mas do número e posicionamento de interlocutores-expectadores sobre a qualidade da produção de gestos. A pesquisadora utilizou em seu experimento a tarefa de narração de história de desenho animado, controlando o número e o posicionamento (face-a-face ou lateral) dos interlocutores-expectadores. Resultados mostraram que falantes mudam o direcionamento e o eixo (de sagital, ou frente-trás, para lateral, direita-esquerda) de elaboração de gestos específicos a conteúdos da fala relacionados a preposições espaciais que codificam uma trajetória com limites inicial e final (*in*, *out*, *into*, *out of*) em função de mudanças no posicionamento de face-a-face para lateral de seus interlocutores-expectadores. Özyürek concluiu de seu estudo que falantes, assim como levam em conta o conhecimento compartilhado com seus interlocutores na elaboração da fala, levam também em conta o campo espacial comum (ou elaboram um espaço comum na intersecção entre o seu espaço e o deles) para a elaboração dos gestos. A pesquisadora rejeita, assim, a hipótese de ocorrência incidental de gestos representacionais e reafirma a hipótese de sua função comunicativa, sem, contudo, descartar sua contribuição para a produção da fala.

Observemos, contudo, com mais atenção algumas ponderações feitas pela pesquisadora em seu estudo.

As alterações verificadas nos gestos relacionadas a preposições espaciais codificando trajetória não se verificaram nem para outras preposições espaciais sem limites inicial e final determinados (como, por exemplo, a preposição *através*), nem para verbos de movimento, como era a hipótese inicial da autora. Também não se verificaram alterações no conteúdo da fala com as alterações observadas nos gestos. Özyürek (2002) especula, de maneira interessante, que os resultados por ela observados podem se limitar

a referências espaciais próprias à língua dos falantes (o inglês), uma vez que usam enquadramento espacial relativo nos gestos³. O experimento de Özyürek (ibid.) merece, portanto, replicação, e sua hipótese, ser colocada a prova com falantes de diferentes culturas e comunidades linguísticas.

Beattie & Coughlan (1998), por um desenho experimental bastante interessante, também colocaram a prova a hipótese comunicativa dos gestos. Hipotetizaram que se a produção de gestos icônicos estivesse exclusivamente a serviço da facilitação do acesso lexical durante a produção da fala, uma vez previamente realizadas as buscas lexicais, seria esperada uma diminuição no número de gestos icônicos em repetições subsequentes da mesma fala. Pedindo a sujeitos experimentais para narrarem sucessivamente uma mesma história a ouvintes-espectadores diferentes, esperaram os pesquisadores uma mudança gradual de uma fala espontânea hesitante com alta proporção de tempo de pausas (relacionadas a acesso lexical) para uma fala mais fluente, melhor formulada e com progressivamente menor número de gestos icônicos por os itens lexicais já terem sido previamente acessados. Como resultado principal, os pesquisadores observaram uma diminuição na frequência de gestos apenas da primeira para a segunda tentativa, mas reportaram que a frequência de gestos icônicos claramente não diminuiu significativamente entre as tentativas como um todo, os sujeitos gesticulando na sexta tentativa 91% da quantidade de gestos da primeira. Concluíram os autores que gestos podem contribuir para o processo de busca lexical, mas que qualquer teoria ou modelo de processamento que encerre a função da gesticulação icônica no processo de acesso lexical não estará completa nem adequada.

6.2.3 Efeitos de visibilidade dos gestos na compreensão de histórias

Em outro trabalho conduzido pelo mesmo grupo de pesquisa, mas em desenho experimental mais clássico, o estudo de Beattie & Shovelton (1999) explorou a compreensão da narração de uma história de desenho animado em condições de visibilidade e invisibilidade dos gestos do narrador. Resultados mostraram que mais informação da história original é obtida quando sujeitos ouvintes-espectadores têm

³ A autora observa que falantes de língua autóctone na Austrália usam enquadramentos absolutos de referência espacial para representar direção em ambos fala e gestos, ou seja, orientam seus gestos em relação a coordenadas geográficas fixas (Haviland, 1993).

acesso à imagem dos gestos icônicos produzidos pelos narradores do que quando apenas ao áudio da narração. Os resultados mostraram melhora de aproximadamente 18% no desempenho dos sujeitos ouvintes-expectadores na tarefa de compreensão da história narrada na condição com gestos. Entretanto, a melhora de desempenho foi observada apenas para as perguntas de compreensão relacionadas à posição e ao tamanho relativos de objetos, e não, como esperavam os pesquisadores, para perguntas relacionadas às propriedades das ações propriamente ditas das personagens.

Ademais as limitações dos resultados, também caberia crítica ao controle experimental do estudo de Beattie & Shovelton (1999) por não ter isolado o fator gesticulação de outros fatores relacionados à exibição do narrador como um todo, de forma que o resultado observado pudesse estar associado não necessária e exclusivamente à gesticulação mas a aspectos atencionais sociais (como de empatia, por exemplo) pela presença do rosto do narrador, uma vez que na condição sem gestos exibiram apenas o áudio das narrações. Não obstante, atestando a comunicabilidade de gestos, também chegaram a resultados semelhantes Pereira & Haddad (2015) ao investigarem a compreensão de narrativas por aprendizes de língua estrangeira. Em seu estudo, apresentaram aos sujeitos, aqui sim em ambas condições experimentais, a imagem do narrador, controlando apenas a exibição de seus gestos. Os pesquisadores também reportaram melhora de desempenho (aumento de 10%) na tarefa de compreensão das narrativas pelos sujeitos experimentais quando na condição de visualização dos gestos do narrador.

6.2.4 Perspectivas divergentes

Passemos agora em rápida revista estudos que, apesar de em menor número na literatura, apresentaram resultados destoantes dos trabalhos até aqui mencionados, alguns deles chegando a colocar mesmo dúvidas sobre as contribuições e efetividade comunicativa dos gestos. Mais do que seu posicionamento teórico, contudo, interessam-nos as ponderações feitas e as hipóteses levantadas pelos pesquisadores em seus estudos a partir de seus resultados experimentais; estas nos abrem possibilidades apontando direcionamento na condução de novos trabalhos. Entendemos que, dados o estágio atual de compreensão e a complexidade do fenômeno estudado, o avanço no conhecimento se assente por ora na dinâmica do debate, na replicação de experimentos, na contestação de

resultados, de posicionamentos e de modelos teóricos. Vejamos a seguir alguns desses trabalhos.

Krauss, Morrel-Samuels & Colasante (1991) investigaram por uma série de cinco experimentos i) o quanto gestos e fala veiculam o mesmo significado e informação complementar entre si, e ii) as contribuições relativas de informação visual e verbal no julgamento de categorias semânticas dos gestos. Resultados de suas análises mostraram que sujeitos são capazes de discriminar o afiliado lexical de gestos com 69% de acertos para ações e localizações e 55% (valor praticamente de aleatoriedade no acerto) para nomeações e descrições de objetos. Os autores do trabalho admitem, assim, que não é difícil de se ver como gestos se relacionam semanticamente com a fala quando esta está presente. No entanto, desafiando evidências experimentais de trabalhos como os de Pereira & Haddad (2015) e de Beattie & Shovelton (1999), especulam os pesquisadores que o reconhecimento semântico do gesto seja um efeito ilusório; que a relação aparente entre seu significado e o significado da fala seja uma construção *post-hoc* derivada da compreensão da fala, sem relação sistemática com os movimentos observados. Para os autores, portanto, ainda que a gesticulação (icônica, eminentemente) veicule algum sentido, este seria pouco informativo e largamente redundante com o da fala; na ausência desta, seu sentido se tornaria opaco, comunicando muito pouco ou mesmo nada.

Em posição teórica semelhante, Lickiss & Wellens (1978) reportaram resultados contrários à hipótese da efetividade comunicativa dos gestos. Os autores investigaram experimentalmente o papel de gestos na comunicação durante tarefa de identificação de retratos fotográficos. Os pesquisadores controlaram tanto a gesticulação do falante quanto a visibilidade de seu interlocutor. Resultados mostraram que: i) o número total de gestos produzidos pelo falante não variou significativamente entre as condições de visibilidade do interlocutor; ii) nem a visibilidade do interlocutor nem a restrição de gestos afetou significativamente o número total de palavras usadas na descrição das fotos; iii) a taxa de erros na fala (repetições e interjeições) aumentou significativamente na condição de visibilidade do interlocutor mas não foi afetada pela restrição dos gestos; iv) a visibilidade e a restrição de gestos não afetaram significativamente o número de identificações corretas de fotos por outros sujeitos experimentais, ou seja, não se verificaram efeitos na efetividade comunicativa dos gestos. Concluíram os autores de seu estudo que o papel dos gestos na comunicação deve ser revisto, uma vez que não encontraram efeitos da restrição dos gestos sobre a produção da fala, disfluência verbal, ou desempenho na tarefa;

mais do que a restrição dos gestos, teve maior impacto sobre a disfluência na fala a presença do interlocutor⁴.

Conforme podemos depreender, os resultados reportados por Lickiss & Wellens (1978) divergem de resultados de outros trabalhos importantes na literatura, alguns deles mencionados anteriormente: a invariância no número de gestos produzidos pelo falante entre as condições de visibilidade do interlocutor difere dos resultados reportados por Alibali, Heath & Meyers (2001; ver seção 6.2.1 sobre o empacotamento da informação); a invariância no número de palavras entre as condições de gesticulação e de restrição de gestos contrasta com resultados de Graham & Heywood (1975); o fato da restrição dos gestos não ter impactado de forma significativa a taxa de erros na fala contraria resultados apontados por Morsella & Krauss (2004; ver seção 6.1.2 sobre acesso lexical); por fim, o fator de visibilidade dos gestos não ter impactado significativamente o desempenho da tarefa de identificação de fotos destoa dos resultados obtidos por Beattie & Shovelton (1999) e daqueles por Pereira & Haddad (2015; ver seção 6.2.3 sobre efeitos de visibilidade dos gestos na compreensão de histórias).

6.2.5 A intersubjetividade

Em uma perspectiva alternativa, poderíamos especular ainda que a contribuição comunicativa dos gestos esteja antes relacionada a aspectos sociais de ordem subjetiva. Gregersen, Olivares-Cuhat & Storm (2009) chamam a atenção para o aspecto de espontaneidade no ato comunicativo, diferente entre as modalidades verbal e gestual. Observam os autores que enquanto que uma mensagem verbal é aprendida, mostras emocionais não-verbais são inatas. Por ser a vasta maioria de comportamentos não-verbais espontânea, estes são mais livres de distorções ou enganos, e por isso socialmente mais fiáveis. Por comportamentos não-verbais se apresentarem sob menor controle consciente, indivíduos os percebem como manifestações autênticas, com significados

⁴ Rimé & Schiaratura (1991) também atestam contra a função comunicativa dos gestos a partir de evidências experimentais de que ouvintes-expectadores não se recordam do conteúdo de gestos visualizados, além de não encontrarem diferença na compreensão de ouvintes-expectadores quando na eliminação dos gestos do falante. Também observam os autores que participantes não são capazes de adivinhar o conteúdo da fala afiliado ao gesto.

mais arraigados e legítimos (ver também Nathan & Alibali, 2011, para o papel de gestos na construção da intersubjetividade).

Diante de tamanha contenda e das limitações ponderadas pelos trabalhos, como olharmos, então, os gestos conversacionais? Poderíamos admitir que, apesar de intimamente relacionados com a fala, mas por serem insuficientemente informativos, gestos conversacionais permaneçam no processo de elocução como vestígio do processo evolutivo da linguagem (Hewes, 1973)? Ou ainda seriam movimentos aleatórios para dissipação da tensão durante a busca lexical (Dittmann & Llewelyn, 1969)? Jacobs & Garnham (2007) ponderam ao observarem que as hipóteses de acesso lexical e da comunicação por gestos não são necessariamente mutuamente excludentes e que, como sugerem os estudos, diferentes gestos podem servir a objetivos diferentes em diferentes contextos.

Portanto, antes de respondermos qual a função primeira dos gestos quando emergem espontaneamente com a fala, talvez seja mais oportuno e produtivo, perguntarmo-nos quais gestos e em que circunstâncias emergem. Voltaremos a abordar as funções de acesso lexical e comunicativa de gestos ilustradores (dentre outras funções) ao tratarmos mais adiante de sua ocorrência em produções narrativas. Vejamos por ora, de maneira breve, curiosas e importantes diferenças étnico-culturais e interlinguísticas na produção de gestos.

CAPÍTULO 7

GESTOS EM DIFERENTES LÍNGUAS

É fato que falantes de diferentes línguas de diferentes partes do mundo gesticulam. Também é do senso comum que determinados grupos étnico-culturais são mais ou menos gesticuladores, como nós brasileiros e italianos por um lado e ingleses e japoneses por outro. Há diferenças no modo como gesticulamos, entretanto? Se sim, quais e como explicá-las? Investigações de recursos e estratégias linguísticas de falantes de diferentes línguas apontam tanto para diferenças culturais quanto linguístico-estruturais na integração semiótica entre as duas modalidades. Vejamos algumas delas a seguir.

7.1 Diferenças étnico-culturais e linguístico-estruturais na produção de gestos

Efron (1972), já referido anteriormente, encontrou diferenças culturais na maneira como italianos e judeus imigrantes nos Estados Unidos gesticulam. Tais diferenças, entretanto, tendem a desaparecer nas gerações seguintes assimiladas culturalmente.

Graham & Argyle (1975) estudaram a comunicação de formas de baixa e alta codificabilidade (dificuldade de descrição) por falantes nativos de inglês e italiano a falantes de mesma língua respectivamente, controlando-se a gesticulação manual durante uma tarefa de descrição de formas. Resultados mostraram melhor desempenho em ambas as línguas quando a descrição das formas foi realizada acompanhada de gestos, com maior efeito para formas de baixa codificabilidade (de difícil descrição). A análise do material de áudio revelou que mudanças de conteúdo indicaram maior dependência de recursos verbais para comunicação de informação espacial na condição sem gestos. Interessantemente, a melhora observada pela presença de gestos foi maior para falantes de italiano do que para falantes de inglês, para ambos níveis de codificabilidade. Os autores concluíram do estudo que gestos comunicam informação espacial mais eficientemente do que a fala, e que a gesticulação manual de falantes nativos de italiano

é mais eficiente na comunicação dessas informações do que a gesticulação de falantes nativos de língua inglesa.

Em uma tradição teórica de relativização linguística (Slobin, 1987), McNeill & Duncan (2000) argumentam que falantes de diferentes línguas operam por modos distintos de pensamento, induzidos pela estrutura linguística de sua língua nativa; a seleção de características conceituais de um determinado evento ocorreria em atendimento a exigências do código linguístico. Os autores trazem como exemplo aspectos de agentividade e trajetória de uma determinada ação disponíveis separadamente em inglês (verbo + preposição) e apenas juntos em georgiano (verbo apenas); em espanhol, haveria uma tendência de omissão do aspecto de modo da ação pela fala, mas uma abundância de sua representação por gestos, coocorrendo com unidades lexicais (verbos) que descrevem trajetória. Inglês, espanhol e georgiano, portanto, apesar de se assemelharem como línguas predicativas, difeririam entre si no modo como focalizam aspectos semânticos da ação. Em mandarim, por outro lado, por conta da topicalização de sua estrutura linguística, golpes gesticulatórios tenderiam a ocorrer cedo no fluxo da fala, recaindo em unidades proposicionais ao invés de predicativas como no espanhol, no inglês e no georgiano (McNeill & Duncan, *ibid.*).

Goldin-Meadow (2003), em mesma direção, argumenta que a disponibilidade de recursos lexicais e estruturais em uma determinada língua resulta em diferentes formas de gesticulação. Diferenças estruturais de proeminência entre o inglês e o japonês, bem como diferenças na maneira como codificam eventos de movimento, resultam em diferenças no padrão de movimentos gesticulatórios que as acompanham. Falantes de turco, por exemplo, demonstram preferência por uso de gestos separados para representarem modo e trajetória, enquanto que falantes de inglês condensam em um mesmo movimento ambos aspectos (Goldin-Meadow, *ibid.*).

Em seu estudo sobre o efeito do posicionamento do interlocutor sobre a produção de gestos (estudo referido anteriormente na seção 6.2.2), Özyürek (2002) pondera que os resultados encontrados de restrição da comunicabilidade de gestos a determinadas classes de preposições espaciais poderiam se limitar ao sistema de referências espaciais próprio à língua dos falantes participantes do experimento (o inglês), uma vez que usam enquadramento espacial relativo nos gestos. Özyürek propõe investigações semelhantes com falantes de diferentes culturas e comunidades linguísticas (falantes de língua autóctone na Austrália, por exemplo, usam enquadramentos absolutos de referência

espacial para representarem direção em ambos fala e gestos, ou seja, orientam-se espacialmente a partir de coordenadas geográficas; Haviland, 1993).

Laurent & Nicoladis (2015; estudo que será discutido em maior detalhe mais adiante no item 8.4), compararam produções narrativas em francês e inglês por falantes bilíngues residentes em uma região de língua inglesa no Canadá com uma comunidade ativa falante de francês. Verificaram maior produção de gesticulação icônica nas produções narrativas em língua francesa. Esse e outros resultados levaram as pesquisadoras a interpretar o fator de práticas culturais como determinante da funcionalidade do uso de gestos.

Estudos como os acima referidos nos mostram que o tratamento dado em gesto e fala a uma ideia dependem, portanto, de ambos fatores culturais e aspectos estruturais gramaticais próprios à estrutura da língua, como a relação de papel temático (de agentividade) e aspectos verbais (de modo ou trajetória em verbos de deslocamento, por exemplo), preferência de uso de estruturas sintáticas predicativas ou tópicas, entre outros. Como consequência, movimentos gesticulatórios diferirão qualitativamente entre falantes de diferentes culturas, bem como entre falantes de diferentes línguas, com diferenças também na eleição do afiliado lexical (com qual unidade lexical coocorrerão) e aspecto imagético que elaborarão.

7.2 Gestos na aprendizagem de línguas estrangeiras

Manuela Macedonia tem se debruçado há mais de duas décadas sobre a função de movimentos corporais na aprendizagem de línguas estrangeiras. A autora aponta (Macedonia, 2014) que, por forte influência da tradição filosófica em nossa cultura linguística ocidental (Descartes, 1637; Saussure, 1916; Chomsky, 1965, 1975; Fodor, 1976, 1987), o ensino-aprendizagem de línguas tem tradicionalmente se apoiado na modalidade audio-visual quase que exclusivamente, tratando a linguagem como um fenômeno abstrato da mente. Macedonia (ibid.) adverte, porém, que já é conhecimento notório, com corroborações mais recentes das ciências cognitivas, que a aquisição linguística se trata de processo multi-sensorial e motor; em outras palavras, que o corpo participa da cognição e do processo de ensino-aprendizagem (ver também Piaget, 1976; Tomasello, 2005; Goldin-Meadow & Alibali, 2013).

Macedonia (2014) esclarece que, durante a aquisição da língua materna, falantes percebem e interagem com o mundo antes de representarem-no simbolicamente. Diferentemente, na aprendizagem de uma língua estrangeira, normalmente em ambiente de sala de aula, um item lexical é muitas vezes apresentado em uma primeira abordagem esvaziado de significado ou de representações sensoriais. No entanto, se uma nova palavra em língua estrangeira é aprendida apenas visual e acusticamente, sua representação se limitará em grande medida a regiões visuais e auditivas no cérebro.

Por outro lado, estudos neurocientíficos mostram que a aprendizagem de um item lexical corresponde a uma experiência impressa de forma funcional em nossa rede neural (Pulvermüller, 1999; Kiefer & Pulvermüller, 2012; Moseley & Pulvermüller, 2014); ao ativarmos um componente da rede, outros componentes a ele conectados também são ativados em resposta ao estímulo. Assim como a leitura de uma palavra relacionada à culinária pode ativar redes neurais de processamento de odor e paladar (González et al., 2006; Barros-Loscertales et al., 2012), a descrição de ações pode promover a ativação de redes responsáveis por controle motor (Fischer & Swaan, 2008; Watson et al., 2013), inclusive para conceitos abstratos relacionados, por exemplo, a emoções (Straube, Sauer & Miltner, 2011; Citron, 2012).

Portanto, assim como a tradução para a língua materna de um item lexical durante aprendizagem em língua estrangeira pode ativar representações mentais já estabelecidas, a realização de um gesto pode acionar representações sensório-motoras pré-existentes. O espalhamento da atividade elétrica pelo cérebro estabelece, assim, uma rede multimodal de ativação, aumentando a ancoragem semântica e o grau de tratamento do item lexical aprendido (**Figura 15**). O tratamento complexo multimodal aumentaria, por sua vez, a retenção do item lexical na memória, bem como a qualidade e velocidade de seu processamento.

Figura 15. Espraçamento de atividade neural por rede multimodal estimulado por tratamento semântico lexical (adaptado de Macedonia, 2014).

De forma interessante, Dudschig, DeLaVega & Kaup (2014) hipotetizaram que a percepção de itens lexicais em língua estrangeira poderia eventualmente disparar respostas sensório-motoras semelhantes (com mesma velocidade) ao observado em língua materna. Para testar sua hipótese, as pesquisadoras investigaram o tempo de reação motora para representações lexicais em língua estrangeira (inglês) comparada com língua materna (alemão). Resultados mostraram não haver diferença entre o tempo de reação motora entre as duas línguas. As autoras tomaram os resultados como evidência para a existência de representação sensório-motora também para itens lexicais codificados em língua estrangeira.

Em um estudo experimental longitudinal (14 meses) para obtenção de sua titulação de doutoramento (Macedonia, 2003; ver também Macedonia & Klimesch, 2014), Manuela Macedonia ensinou a 29 estudantes universitários falantes de alemão 36 palavras artificiais a partir de regras fonotáticas do italiano. Os itens lexicais eram pertencentes às categorias gramaticais de substantivos, adjetivos, verbos e preposições, e foram criados e significados arbitrariamente a partir de um programa linguístico. Por razões ecológicas de interesse pedagógico, o experimento foi conduzido em ambiente de sala de aula e sem controle experimental de: extensão de palavra, frequência de

ocorrência, familiaridade, ou valência emocional. O ensino foi realizado a partir de projeção da ortografia da palavra e de sua tradução em alemão sobre a parede da sala, acompanhada de reprodução verbal pela experimentadora e repetição pelos aprendizes. Ora os itens lexicais eram acompanhados de gesto, ora, não. Na condição experimental com gesto, a apresentação do item lexical era acompanhada pela produção de um gesto ilustrador de algum aspecto semântico elegido arbitrariamente e desempenhado pela experimentadora, e imitado pelos sujeitos experimentais. Também por razões ecológicas de interesse pedagógico, não se eliminou da etapa de ensino a expressão facial natural à gesticulação corporal, como é frequentemente observado no controle experimental de estudos exploratórios sobre o papel de gestos na codificação de itens lexicais. O desempenho na tarefa de evocação foi significativamente melhor para os itens codificados com gesto, depois de 1, 8, 15, 73 dias e mesmo 14 meses.

Outros trabalhos que avaliaram o impacto de gestos na aprendizagem de novos vocábulos por aprendizes de língua estrangeira obtiveram resultados semelhantes. Tellier (2008) ensinou itens lexicais em inglês a crianças francesas em idade pré-escolar, metade delas por exposição a imagens, e metade pela realização de gestos pelas próprias crianças (por imitação de gestos da experimentadora). O desempenho na memória foi significativamente melhor na condição com gestos. Engelkamp & Zimmer (1985) e Zimmer, Helstrup & Engelkamp (2000) mostraram também que itens codificados na presença de gestos aparecem com maior facilidade e frequência na produção de aprendizes; ademais, Kelly, McDevitt & Esch (2009) e Macedonia, Müller & Friederici (2011) mostram melhora de desempenho na memória (tanto para período curto quanto longo de tempo) para itens lexicais codificados em língua estrangeira acompanhados de movimentos gesticulatórios semanticamente congruentes com o conteúdo de seu afiliado lexical verbal, quando comparados com itens codificados acompanhados de movimentos gesticulatórios aleatórios.

Por tais evidências, Macedonia (2014) critica abordagens exclusivamente simbólicas da linguagem. Para a pesquisadora, gestos desempenhados durante aprendizagem de itens lexicais em língua estrangeira podem criar representações incorporadas novas bem como conectarem-se a representações pré-existentes na língua materna do falante. A gesticulação enriqueceria, dessa forma, a representação mental da palavra aprendida pelo acréscimo de um componente motor em sua representação neural (ver também Nilsson, 2000; Masumoto et al., 2006; Eschen et al., 2007; Macedonia & Knösche 2011).

Admite Macedonia (2014), contudo, que a aprendizagem de vocabulário é apenas uma faceta da aprendizagem de línguas e mais pesquisa será necessária para esclarecer como gestos auxiliam na aprendizagem de palavras de classe gramatical funcional, como preposições, conjunções, dentre outros, bem como em outros domínios da linguagem, tais a morfologia e a sintaxe. Abordaremos no próximo capítulo alguns desses e mais alguns aspectos linguísticos da gesticulação a partir de trabalhos que exploraram diferentes funções da gesticulação manual acompanhando produções narrativas por crianças e adultos em sua língua materna e em língua estrangeira. Dadas a extensão e importância dos tópicos levantados a seguir para a discussão de nosso estudo experimental 1 e argumentação da tese, contamos com a generosa paciência e continuada atenção de nosso leitor.

CAPÍTULO 8

GESTOS NA PRODUÇÃO NARRATIVA

Ambos adultos e crianças lançam mão de variadas estratégias discursivas verbais e gestuais durante a produção narrativa. Variam muito de indivíduo para indivíduo na maneira como o fazem, etretanto. Explicações para a grande variabilidade na frequência e no uso dos gestos podem residir em parte no desenvolvimento de habilidades cognitivas narrativas em acompanhamento ao desenvolvimento linguístico etário, no grau com que falantes se apóiam em recursos visuo-espaciais, no nível de proficiência linguística e narrativa, bem como em diferenças linguístico-culturais.

Trazemos em detalhe a seguir trabalhos recentes que exploraram a emergência de diferentes naturezas de gestos durante a produção narrativa por crianças e adultos em diferentes línguas e em diferentes níveis de proficiência. Sendo a habilidade narrativa um forte preditor de sucesso no letramento escolar (Vilà-Giménez & Prieto, 2019), a ampliação do conhecimento sobre o funcionamento de gestos na recontação de histórias pode trazer contribuições pedagógicas importantes.

8.1 Gestos no desenvolvimento linguístico-narrativo etário

Gestos aparecem cedo nas histórias de crianças em acompanhamento a um florescimento da capacidade narrativa. Já aos três anos de idade, crianças passam a se servirem de gestos para identificação de personagens⁵ (O'Neill & Holmes, 2002).

Dos 4 aos 10 anos de idade, verificam-se avanços importantes em suas habilidades narrativas: apresentam clareza no rastreamento de referentes discursivos; incorporam foco narrativo através da representação de ações pela perspectiva de personagens⁶ (Demir, Levine & Goldin-Meadow, 2015); passam também a contar histórias com

⁵ A função de identificação de personagens através da gesticulação por crianças a partir de seus três anos de idade tende a uma diminuição em anos seguintes (O'Neill & Holmes, 2002).

⁶ Crianças que apresentam incorporação de foco narrativo aos 5 anos tendem a produzir narrativas futuras mais complexas e melhor estruturadas já aos 6 a 8 anos de idade (Demir, Levine & Goldin-Meadow, 2015).

começo, meio e fim, com sequência lógica e coesão interna através do uso consistente de marcadores temporais e de aspectos verbais (Berman & Slobin, 1994; Wigglesworth, 1997). Em acompanhamento à complexificação estrutural narrativa, verificam-se também alterações na maneira como gesticulam as crianças: entre 5 a 10 anos de idade, passam a apresentar variação funcional no uso de gestos, atendendo agora também a necessidades de marcações discursivas pragmático-estruturais (Colletta et al., 2015)⁷.

Alterações na funcionalidade de gestos pelo avanço etário foram estudadas por Zampini et al. (2016). Os autores analisaram a competência linguística verbal, a taxa e a função da gesticulação em tarefa narrativa por 45 crianças monolíngues falantes de italiano, variando entre 38 e 71 meses de idade. Testaram a habilidade linguística das crianças através de teste gramatical e de teste de competência narrativa a partir de um livro de 18 figuras elaborado para a produção de narrativas por crianças pré-escolares. Analizaram ainda aspectos lexicais (variedade lexical) e estruturais da produção narrativa, estes últimos a partir de índices de frequência de coordenação e subordinação oracionais, de períodos simples e compostos, e de frases sem verbo. Os gestos foram analisados quanto à sua natureza: gestos de apontamento (proximal e distal), gestos de mostração (ao aproximarem a imagem do livro, apresentado como estímulo, na linha de visão do ouvinte), gestos convencionais (ou emblemáticos), e gestos icônicos. Também classificaram os gestos quanto à sua função: gestos dêiticos, gestos complementares (que não adicionam informação) e gestos suplementares (que adicionam informação). Tanto para a natureza quanto para a função de gestos, utilizaram índices de ocorrência da categoria pelo número total de gestos. A hipótese principal dos autores era uma maior produção de gestos por crianças de menor competência linguística.

Resultados mostraram aumento da complexidade sintática e da variedade lexical das narrativas, bem como diminuição da frequência dos gestos com o aumento da idade das crianças. O gesto mais frequente para os 3 grupos etários (3, 4 e 5 anos) foi o gesto de apontamento proximal. Não se encontrou diferença significativa entre os grupos quanto a: ocorrência das diferentes categorias de gestos; número de orações e número total de palavras. Enquanto que, para crianças de 5 anos, verificou-se maior ocorrência de períodos complexos, em crianças de 3 anos verificou-se maior proporção de frases sem verbo. Não se verificou correlação entre a produção de gestos e a produção lexical;

⁷ Ver também Demir, Fisher, Goldin-Meadow & Levine, 2014, Demir, Levine & Goldin-Meadow, 2015, e Stites & Özçalışkan, 2017 para maior elaboração de gestos representacionais em acompanhamento ao aumento de complexidade narrativa.

crianças com baixa pontuação no teste gramatical, contudo, produziram mais gestos. A taxa de gesticulação esteve negativamente correlacionada com a produção de períodos complexos e positivamente correlacionada com a proporção de frases sem verbos. Um resultado que chama a atenção, conforme se observa no gráfico abaixo (**Gráfico 2**), é a tendência de aumento de gestos com função complementar (em redundância com a fala) e de diminuição de gestos com função suplementar (com informação adicional à fala) e dêiticos pelo aumento da idade das crianças.

Gráfico 2. Funções comunicativas de gestos produzidos em associação com produções verbais (Zampini et al., 2016).

Concluíram os autores dos resultados de suas análises que gestos não tiveram função compensatória lexical (não substituíram palavras); antes, foram usados como apoio estrutural, principalmente por crianças com baixa competência morfofossintática e narrativa. Sugerem ainda os autores que gestos com função suplementar (de adição de informação) podem ser compreendidos como índice de dificuldades na produção linguística verbal, maiores em crianças de menor idade. O trabalho de Zampini et al. (2016) corrobora, portanto, resultados de estudos anteriores (Iverson & Braddock, 2011; Mainela-Arnold, Alibali, Hostetter & Evans, 2014) que sustentam que a produção de gestos comunicativos, antes de uma estratégia para resolução de problemas lexicais, é uma estratégia compensatória para crianças de baixa competência linguística.

Outro dado que chama atenção no estudo por Zampini et al. (2016), relacionado à conclusão final dos pesquisadores de que gestos não exerceram função compensatória lexical, é a baixa ocorrência de gestos icônicos: menos de 14% para os 3 grupos etários – a gesticulação icônica é tipicamente compreendida por representar visuo-espacialmente aspectos concretos de referentes discursivos como personagens, objetos, suas características físicas, bem como ações e aspectos de movimento, dentre outros. Uma questão que se coloca sobre esse dado em particular é se a caracterização da população amostral do estudo (crianças de 38 a 71 meses de idade) não explicaria tal resultado por seu repertório lexical, em função de tão tenra idade, não ter sido ainda suficientemente acrescido de traços neurais repertoriados sensória e motoramente, como se esperaria em um falante adulto. A replicação do estudo com crianças de idades mais avançadas, adolescentes e adultos poderia trazer elucidações para a questão.

A partir de tais evidências, é plausível assumirmos que alterações funcionais no uso da gesticulação relacionadas ao avanço etário reflitam, portanto, avanços na complexificação da capacidade narrativa das crianças (Colletta, Pellenq & Guidetti, 2010); mais ainda, poderiam ser tomadas como índices preditivos de desenvolvimento linguístico-cognitivo (Fasolo & D’Odorico, 2012; Thal & Tobias, 1992; Butcher & Goldin-Meadow, 2000; Capirci, Iverson, Pizzuto, & Volterra, 1996; Iverson, Capirci, Volterra, & Goldin-Meadow, 2008). O trabalho de Butcher & Goldin-Meadow (2000) descrito logo a seguir traz evidência contundente.

Butcher e Goldin-Meadow (2000), investigando crianças em fase de aprendizagem escolar, descobriram que, no esforço de enquadramento e acomodação conceitual, gestos podem veicular informação contrastante com a fala. Tal discrepância seria a manifestação de instabilidade e vulnerabilidade cognitivas pela administração de duas ideias (concorrentes) a cerca de um mesmo problema. Em seu estudo, crianças que produziram incompatibilidades entre gesto e fala (diferentemente de crianças que produziram apenas gestos compatíveis com o conteúdo da fala) avançaram em seu estágio de conhecimento sobre o conceito trabalhado (conceito de conservação volumétrica) ao receberem instrução na tarefa. Portanto, momentos de instabilidade e vulnerabilidade cognitivas percebidas no contraste gesto-fala poderiam revelar risco iminente de aprendizagem. Por outro lado, também professores se valeriam de informações da gesticulação de seus alunos para adequação da instrução em momentos de disponibilidade cognitiva de aprendizes (Goldin-Meadow, Kim & Singer, 1999).

Essa íntima relação entre a funcionalidade da gesticulação e a complexidade linguística persiste, contudo, em produções narrativas por adultos. Jenkins, Coppola & Coelho (2017), procurando compreender os potenciais efeitos facilitadores de gestos no planejamento complexo discursivo, analisaram a produção narrativa de 10 sujeitos sob as condições de livre gesticulação e de restrição de gestos. Os resultados de suas análises mostraram narrativas com significativamente mais subordinação oracional, episódios completos e melhor organização na condição experimental de livre gesticulação. Os pesquisadores interpretaram seus dados compreendendo que a gesticulação teria promovido alívio de carga cognitiva liberando recursos atencionais para melhor planejamento macro-linguístico⁸.

8.2 Elementos visuo-espaciais e a criatividade na gesticulação narrativa

Para além de diferenças relacionadas à complexificação de habilidades narrativas próprias ao avanço etário, a variação no uso de gestos também pode se dar em acompanhamento a diferentes estratégias narrativas relacionadas ao grau com que indivíduos se apóiam em recursos cognitivos visuo-espaciais (Smithson & Nicoladis, 2014).

Laurent, Smithson & Nicoladis (2020) testaram experimentalmente se crianças pré-escolares que gesticulam com maior frequência tendem de fato a se apoiar mais do que crianças espontaneamente pouco ou não-gesticuladoras em recursos cognitivos visuo-espaciais. Análises evidenciaram para ambos os grupos correlação positiva entre a extensão das histórias e o uso de memória visuo-espacial, com crianças gesticuladoras produzindo, conforme hipotetizado pelas autoras, histórias significativamente mais longas do que crianças espontaneamente não-gesticuladoras. A recontação de histórias exigiria, pois, a manutenção de informação visuo-espacial de episódios na memória (Rubin, 1995); maior apoio em recursos cognitivos visuo-espaciais levaria à lembrança de mais informação (Parrill, Bullen & Hoburg, 2010), e, como consequência, à produção de histórias mais extensas (Nicoladis, Marentette & Navarro, 2014). Gestos estariam

⁸ Vale notar no estudo de Jenkins, Coppola & Coelho (2017), contudo, seu pequeno tamanho amostral ($n = 10$), bem como salientar a grande variância nos dados em: a quantidade de gestos produzidos pelos sujeitos (de 3 a 81 gestos), bem como a extensão narrativa. Diferentemente de outros trabalhos tratados mais adiante, não encontraram os pesquisadores entre os grupos efeito de gesticulação sobre a extensão ou o conteúdo narrativo (ver Laurent, Smithson & Nicoladis, 2020, para efeito de gestos sobre extensão narrativa).

direta e positivamente correlacionados com o processamento visuo-espacial de conteúdo imagético na memória (Matthews-Saugstad, Raymakers & Kelty-Stephen, 2016)⁹.

Ainda sobre o estudo de Laurent, Smithson & Nicoladis (2020), as autoras analisaram também a criatividade das produções narrativas através do número de palavras de baixa frequência e da organização de cenas. Diferentemente do que previram, não encontraram diferenças significativas entre os grupos de crianças gesticuladoras e pouco gesticuladoras; ambos grupos produziram em torno da mesma proporção de palavras de baixa frequência (aproximadamente 10%). Explicaram as pesquisadoras, entretanto, pela pouca idade das crianças participantes no estudo (variaram de 36 a 70 meses de idade), sua população amostral não apresentou variabilidade lexical suficiente. Apesar do resultado negativo de sua análise quantitativa, concluíram o estudo apontando que a análise qualitativa das produções narrativas das crianças sugere que crianças não-gesticuladoras tendem a usar um processo de memorização mais mecânico para a realização da tarefa, recontando o desenho animado como uma série simples de descrições de eventos; em contraste, crianças espontaneamente gesticuladoras tendem a uma estratégia de organização de eventos mais criativa e divergente (ver também Kirk & Lewis, 2017, para correlação positiva entre a produção de gestos e maior criatividade narrativa).

Já Nicoladis, Marentette & Navarro (2014), por sua vez, estudaram a taxa de gesticulação apenas de natureza icônica em relação a ambos aspectos etário e de estratégias visuo-espaciais por crianças pré-escolares em tarefa narrativa. Os autores testaram a predictibilidade de 3 fatores na produção de gestos icônicos em tarefa de contação de histórias por crianças de 3 a 10 anos de idade, a saber: a idade, a complexidade narrativa e a dependência de imageamento visuo-espacial. A complexidade narrativa foi indexada pelo número de conectores diferentes (por especificarem relações temporais ou causais); o imageamento visuo-espacial foi indexado pela extensão da história, medida em número de palavras diferentes. O estudo foi conduzido com 27 crianças francesas, 26 inglesas e 26 chilenas. Conforme previram, seus resultados mostraram correlação positiva entre os 3 fatores analisados: idade, complexidade narrativa e imageamento visuo-espacial. Não encontraram, entretanto, correlação entre a taxa de gesticulação e a idade para nenhum dos 3 grupos linguístico-culturais; também não se observou diferença significativa na taxa de gesticulação entre eles. A taxa de

⁹ Ver também Rauscher, Krauss & Chen (1996) e Rimé, Shiaratura, Hupet & Ghyssels (1984) para redução de conteúdo imagético na fala pela restrição de gestos.

gesticulação, porém, correlacionou-se positivamente com o número de palavras diferentes em todos os 3 grupos linguístico-culturais, e com o número de conectores apenas para falantes de língua francesa. Uma análise de regressão linear mostrou apenas o número de palavras diferentes como forte preditor da taxa de gestos icônicos produzidos pelas crianças, os conectores ficando praticamente à margem da significância estatística.

A partir das análises quantitativas e qualitativas do estudo, Nicoladis, Marentette & Navarro (2014) concluíram que crianças que gesticulam ao contarem uma história contam-na de maneira muito vívida e sensorialmente rica, com gestos incorporando ações de personagens. Por outro lado, crianças espontaneamente pouco gesticuladoras tendem a apenas comentar sobre episódios, comentando eventos de maneira usual e sem detalhamento. Ademais, concluíram que seu estudo corrobora resultados de estudos anteriores (Nicoladis, Pika, Yin & Marentette, 2007; Rubin, 1995), que apontam para correlação positiva entre taxa de gesticulação e extensão narrativa, revelando ainda com outros trabalhos (Wesp, Hesse, Keutmann & Wheaton, 2001) a importante função do uso de gestos para ambos crianças e adultos na manutenção de informação visuo-espacial na memória operacional em tarefas de recontação de histórias.

Por fim, outro trabalho interessante que explorou, não a gesticulação icônica na produção narrativa, mas gestos de pulso é o estudo por Vilà-Giménez & Prieto (2019). Vilà-Giménez & Prieto (ibid.) testaram o quanto encorajar crianças a produzirem gestos de pulso pode melhorar seu desempenho narrativo. Estudaram a produção narrativa por 47 crianças catalãs de 5 a 6 anos de idade. Realizaram pré- e pós-teste de estrutura e fluência narrativas, com uma sessão-treino intermediária aos testes, na qual as crianças observaram e recontaram histórias a partir de 6 vídeos de contação de histórias (de 45s cada, aproximadamente) por uma contadora adulta que utilizou gestos de pulso para marcações discursivas (como advérbios, conjunções e outros conectores) e focalização de palavras de conteúdo (carregadas semanticamente, tais substantivos, adjetivos e verbos). Durante a recontação das histórias pelas crianças na sessão-treino, metade delas foram encorajadas a moverem as mãos à maneira da contadora de histórias nos vídeos assistidos. Analisaram, então, de suas produções narrativas pré e pós-teste, 4 características estruturais discursivas: a presença de um protagonista animado, a estrutura temporal, a estrutura causal, e a presença de uma ação direcionada a um objeto. Além das características estruturais, analisaram também a fluência da fala, por apreciação holística com pontuação de 1 a 7 por falantes nativos. Resultados mostraram que crianças do grupo encorajado a gesticular na fase de treinamento desempenharam melhor em

ambas estrutura narrativa e fluência do que crianças que simplesmente recontaram a história sem instrução sobre os gestos. Concluíram as autoras que seus resultados fornecem evidências de funções discursivas pragmáticas de gestos de pulso, melhorando planejamento linguístico e processos de aprendizagem subjacentes a discurso narrativo oral. As autoras sugerem que outros estudos sejam realizados com outros tipos de gestos, bem como para verificação de efeitos a longo prazo.

8.3 O planejamento discursivo

Em língua estrangeira, as etapas de processamento da fala (conceitualização, formulação, articulação e monitoramento; Levelt, 1989) demandam maior recrutamento de recursos atencionais e consciência sobre a própria produção, resultando em significativo aumento da carga cognitiva (Kormos & Trebits, 2012). Com capacidade atencional e de processamento limitadas, o aumento da carga cognitiva pode trazer prejuízo ao atendimento de algumas das etapas, em especial na conceitualização e na formulação. Permitir a falantes de língua estrangeira oportunidade de planejamento discursivo antes de sua realização pode trazer benefícios como diminuição da carga cognitiva da etapa de conceitualização, por facilitar a recuperação e o tratamento de informação, resultando em melhora de desempenho discursivo narrativo (Ellis, 2016; Abrams & Byrd, 2016; Ahmadian, Tavakoli & Vahid Dastjerdi, 2015; Cho, 2018; Nitta & Nakatsuhara, 2014). A diminuição da carga cognitiva permitiria, ademais, maior grau de redundância semântica inter-modal gesto-fala, observada, por exemplo, em um maior recrutamento de gesticulação icônica (Krauss, Chen & Gottesman; 2000).

Lin (2020) investigou a influência do planejamento e da gesticulação na realização de tarefa narrativa em língua estrangeira. Participaram de seu estudo 40 sujeitos experimentais falantes de mandarim, todos em nível avançado de inglês. Os participantes assistiram a um episódio de desenho animado em uma tela de projeção e, em seguida, narraram a história em língua estrangeira (inglês) a um ouvinte-expectador. Divididos em dois grupos (pré-planejamento e planejamento em curso), foi dado aos sujeitos experimentais do grupo pré-planejamento tempo livre de planejamento antes da realização da tarefa narrativa. Sujeitos de ambos os grupos narraram a mesma história duas vezes: uma vez com gestos, e outra vez sem gestos, aleatorizando-se entre os participantes a ordem para eliminação de efeitos de aprendizagem. Na condição sem gestos, os sujeitos

experimentais foram solicitados a manterem suas mãos sobre suas pernas ou em outra posição na qual se sentissem confortáveis. Lin analisou então os efeitos de planejamento pré-tarefa narrativa e planejamento em curso sobre a qualidade discursiva da narrativa e a funcionalidade da produção de gestos.

A qualidade discursiva foi analisada através do número de cenas completas descritas, e o número de proposições novas, cada proposição nova sendo considerada uma ação clara, não-repetida, e relacionada à história. A funcionalidade dos gestos produzidos pelos participantes na condição de livre gesticulação foi classificada entre: ativação, manipulação, empacotamento e exploração da informação (ver Kita, Alibali & Chu, 2017, para o papel conceitualizador de gestos). A ativação foi identificada quando o movimento gesticulatório criou ou manteve ativa uma representação de alguma característica perceptual do conteúdo imagético da fala (exemplo, altura, largura, direção, forma e ações) evitando seu esvanecimento da memória; a manipulação foi considerada quando o falante usou gestos para manipular mentalmente uma representação espaço-motora, para melhor expressão de seu conteúdo (por exemplo, uma cena complexa com vários eventos de movimento ou dificuldades na evocação de uma próxima cena); o empacotamento, quando, para facilitação de tratamento, múltiplos pedaços de informação complexa foram exibidos em diferentes unidades verbo-gestuais, por provavelmente não ter sido possível codificar todos aspectos relevantes da informação de forma simultânea ou em uma única proposição; por fim, a exploração foi identificada quando o narrador usou gestos para procurar informação potencialmente relevante na tentativa de decidir que partes da informação expressar verbalmente, por exemplo tentando lembrar as cenas ou procurando as palavras mais apropriadas durante a tarefa narrativa.

Resultados mostraram que, com o aumento da carga cognitiva (condição de planejamento em curso), os sujeitos experimentais tenderam a produzir maior volume de gestos, com significativamente mais gestos dêiticos; enquanto que na condição de pré-planejamento, gestos icônicos foram os mais frequentes. Para a autora, enquanto que gestos dêiticos ajudaram falantes a identificarem e localizarem entidades da história beneficiando o planejamento discursivo, o uso icônico de gestos teve função autorregulatória, apresentando alta redundância com o conteúdo da fala. O tempo de planejamento dado aos participantes na condição de pré-planejamento permitiu-lhes maior familiaridade com os elementos narrativos e a criação de representações mais elaboradas da história e com maior grau de redundância semântica. Quanto à função dos gestos, em ambos grupos planejamento em curso e pré-planejamento, a ativação foi a

função gesticulatória mais frequente (a ativação ajuda a fala na manutenção de representação de informação pre-existente, ativa nova informação e gera simulações; Kita, Alibali & Chu, 2017). As funções de exploração e de manipulação foram significativamente mais frequentes na condição de planejamento em curso, resultado que a autora do estudo interpretou como sugestão de possíveis processos consecutivos de tentativa e erro de formulações linguísticas, processos esperadamente mais frequentes em situação de aumento da demanda cognitiva - trabalhos anteriores de Chu & Kita (2011) e Hostetter, Alibali & Bartholomew (2011) também demonstraram aumento da manipulação do conteúdo mental por gesticulação manual com aumento da dificuldade em tarefas de resolução de problemas. Lin (2017) e Lin (2019) também relataram uso de gestos de manipulação como forma compensatória para informação proposicional incompleta codificada verbalmente diante de dificuldades linguísticas, sobretudo por sujeitos aprendizes de língua estrangeira com baixo grau de proficiência.

Quanto ao conteúdo e à organização narrativa, a condição de livre gesticulação resultou em maior número de cenas completas, em comparação com a condição de restrição de gestos, corroborando resultados de trabalhos anteriores que sugerem melhor planejamento discursivo pelo uso de gestos (Jenkins, Coppola & Coelho, 2017). Entretanto, não se observou diferença no número de cenas completas entre as condições de planejamento. Quanto ao número de novas proposições, verificou-se efeito de ambas condições de gesticulação e de planejamento, com maior número de novas proposições para a condição de pré-planejamento e de livre gesticulação. Lin (ibid.) interpretou seus resultados sugerindo que ambos pré-planejamento e livre gesticulação facilitaram a etapa de conceitualização compensando limitações atencionais e de memória operacional durante a produção das narrativas orais em língua estrangeira. Para a autora, a gesticulação é uma variável significativa para ambos o planejamento discursivo e a recuperação de informação proposicional. A gesticulação promoveria a ativação e a manutenção de informação espaço-motora, propulsionando narradores a expressarem mais conteúdos proposicionais através da fala. Por outro lado, a autora pondera que, compreendendo fala e gesto como partes de uma mesma economia cognitiva, o desencorajamento à gesticulação pode ter inibido recursos mentais narrativos, dificultando o acesso lexical e a organização discursiva, aumentando a carga cognitiva na tarefa e diminuindo assim o desempenho narrativo dos participantes.

8.4 O uso compensatório de gestos

O resultado observado no estudo por Lin (2020) de maior produção de gestos com aumento da carga cognitiva (condição de planejamento em curso) corrobora resultados de trabalhos anteriores que exploraram a noção compensatória de gestos para dificuldades linguísticas por falantes de língua estrangeira (Lin, 2019, 2017; Nicoladis, Pika, Yin & Marentette, 2007). Porém, a noção compensatória de gestos em auxílio a dificuldades linguísticas advém também de trabalhos que investigaram a produção narrativa em diferentes línguas por sujeitos bilíngues. Os trabalhos de Pika, Nicoladis & Marentette (2006) e de Nicoladis, Pika & Marentette (2009) sugerem maior recrutamento de gestos por sujeitos bilíngues do que por sujeitos monolíngues, com diferenças qualitativas no uso de gestos para cada uma das línguas, por exemplo, com maior número de gestos icônicos ou representacionais em sua língua dominante (ver também Gullberg, 1999; Laurent, Nicoladis & Marentette, 2010; Nicoladis, 2002; Nicoladis, Mayberry & Genesee, 1999). Falantes bilíngues tenderiam, portanto, a gesticular em atendimento à norma cultural da língua falada (conforme apontado por Laurent & Nicoladis, 2015, referido anteriormente no item 7.1 e desenvolvido mais extensamente a seguir nesta seção), mas também em atendimento a eventuais dificuldades linguísticas, por exemplo, dificuldades de acesso lexical na língua menos dominante ou em casos de não-fluência (ver também Gollan & Acenas, 2004; Kroll, Michel, Tokowics & Ducour, 2002).

Um fenômeno que chama a atenção em trabalhos que investigaram o uso de gestos em tarefas de alta demanda cognitiva, como resolução de problemas ou produções discursivas em língua estrangeira, é a ocorrência de um tipo específico de gesticulação auto-orientada, produzida em silêncio, e com caráter de manipulação e exploração conceituais, os chamados *co-thought gestures* (Alibali, Spencer, Knox & Kita, 2011; Chu & Kita, 2011; Eielts et al., 2018). *Co-thought gestures*, a que propomos nos referirmos como gestos conceituais exploratórios, permitem a exploração espaço-motora da imagem mental para melhor focalização e seleção de domínios relevantes de informação a serem transformados em proposições (Kita et al., 2017). Tais gestos são relatados ocorrerem com frequência na produção discursiva em língua estrangeira, como forma de compensação de dificuldades linguísticas (Stam, 2017).

Pela função compensatória de gestos observada experimentalmente na literatura, é esperado que a restrição da gesticulação prejudique em maior grau a produção linguística na língua não-dominante em falantes bilíngues, bem como em produções em

língua estrangeira. Foi o que testou experimentalmente Laurent & Nicoladis (2015), estudo já referido anteriormente e que agora desenvolvemos em maior detalhe.

Laurent & Nicoladis (2015) procuraram abordar o empacotamento de informação e o acesso lexical comparando produções narrativas por falantes bilíngues com produções narrativas por falantes monolíngues. Pediram a 33 bilíngues franco-anglófonos adultos e 15 adultos monolíngues falantes de inglês, todos canadenses, que contassem uma história duas vezes, sendo uma vez com gesticulação livre e outra com restrição dos gestos. Os sujeitos eram residentes de uma região de língua inglesa no Canadá com uma comunidade ativa falante de francês, de forma que foram expostos à segunda língua desde pequenos em relações sociais.

A tarefa experimental consistiu na recontação de 2 segmentos de desenho animado (7 minutos) nas duas condições experimentais, com e sem gestos, a um falante nativo da respectiva língua. Na condição sem gestos, os sujeitos sentaram-se sobre as mãos. Os sujeitos bilíngues foram solicitados a realizarem a tarefa em inglês e em francês com um intervalo de 10 dias entre as duas sessões. Das produções narrativas, analisaram-se 3 medidas linguísticas: a extensão, através do número total de palavras (foram excluídos falsos-inícios e auto-repetições); a complexidade, através do número total de palavras diferentes; e o número de cenas, cada cena sendo concebida a partir dos episódios como uma possível oração (dentre 40 cenas possíveis; falantes incluíram em média menos de 20). Também se mediu a hesitação a partir do número de interjeições (uhm, uh, etc), na tentativa de se verificar se a restrição de gestos diminuiria a fluência da fala. Da gesticulação, contabilizaram gestos icônicos, dêiticos, convencionais (ou emblemáticos) e de pulso; colapsaram, contudo, as últimas três categorias em uma única denominada gestos não-icônicos.

A hipótese do estudo era de que, se gestos viessem ao auxílio da produção linguístico-discursiva e do acesso lexical na língua não-dominante de sujeitos bilíngues, sua restrição deveria impactar mais suas produções em sua língua não-dominante (com produções mais curtas, mais hesitantes e menos ricas lexicalmente), em comparação com sua produção na língua dominante, bem como em comparação com produções de sujeitos monolíngues. Resultados mostraram maior produção de gesticulação icônica por sujeitos bilíngues em geral em comparação com o grupo de sujeitos monolíngues. Contudo, bilíngues produziram significativamente mais gestos icônicos quando narrando em língua francesa do que em língua inglesa, a despeito de dominância linguística. Também em língua francesa, a restrição dos gestos impactou todas as variáveis linguísticas analisadas;

levou a produções mais curtas e menos elaboradas. A restrição dos gestos não teve efeito sobre as variáveis linguísticas nem nas produções em língua inglesa do grupo de falantes bilíngues, nem nas produções do grupo monolíngue.

As autoras interpretaram os resultados de suas análises compreendendo que a gesticulação pode ter operado em auxílio linguístico (lexical e discursivo) para os falantes bilíngues apenas em língua francesa. Possíveis explicações dadas pelas autoras para tais resultados poderiam estar em parte na especificidade de sua população amostral (o estudo deveria ser replicado com falantes bilíngues residentes em locais de domínio da língua francesa) e em parte na metodologia empregada para determinação de dominância linguística no caso dos falantes bilíngues, pois todos sujeitos se autodeclararam fluentes em ambas as línguas; a dominância linguística foi operacionalizada a partir das próprias amostras. Outra interpretação dada pelas autoras aos resultados obtidos seria, não a contribuição dos gestos para a produção linguística em língua francesa em específico, mas sim diferenças entre normas culturais, com produções de histórias mais imageticamente elaboradas, mais vívidas, bem como mais extensas e com maior uso de gestos na cultura de língua francesa (ver também Nicoladis, Nagpal & Marentette, 2011; Dart, 1992). Concluíram, por fim, Laurent & Nicoladis (2015) que o modo como bilíngues são afetados pela restrição de gestos pode depender da combinação de fatores relacionados a padrões linguístico-culturais de fala e de contação de histórias em cada comunidade linguística.

De fato, o estudo de Laurent & Nicoladis (2015) traz contribuições relevantes para a compreensão de diferenças linguístico-culturais no uso de gestos em produções narrativas. A hipótese das autoras de que seus resultados se explicariam por diferenças entre normas culturais, com produções de histórias mais imageticamente elaboradas, mais vívidas, extensas e com maior uso de gestos na cultura de língua francesa contradiz, porém, resultados reportados por Efron, 1972, que observou o desaparecimento de diferenças culturais no padrão gesticulatório em gerações assimiladas culturalmente. Ademais, apesar de um desenho experimental interessante proposto no estudo de Laurent & Nicoladis (ibid.), pela característica de sua população amostral (os sujeitos bilíngues eram todos fluentes em ambas línguas francesa e inglesa – alguns declararam ambas como línguas maternas), bem como pelo fato de terem colapsado em uma única categoria gestos dêiticos, de pulso e convencionais (ou emblemáticos), perdendo-se, portanto, informação qualitativa da ocorrência de gestos, não foi possível acessar sua função compensatória; considerando-se pouca ou nenhuma diferença de dificuldades linguísticas entre as línguas

para os falantes bilíngues, podemos assumir que simplesmente não houve controle da carga cognitiva da tarefa capaz de suscitar o efeito testado. Dessarte, permanece no respectivo estudo a questão sobre qual a qualidade ou natureza de gestos conversacionais exerceriam de fato a função compensatória frente ao aumento de carga cognitiva por dificuldades linguísticas em produções narrativas em língua estrangeira¹⁰. Também permanece aberto o perfil linguístico de falantes (monolíngue, bilíngue, falante de língua estrangeira de nível iniciante a avançado) e o contexto sócio-linguístico de para quem viria em auxílio compensatório a gesticulação manual. Outras tentativas interessantes de se flagar a função linguístico-compensatória de gestos foram empreendidas, como veremos nos estudos a seguir, explorando outras categorias de gestos em produções narrativas por falantes de língua estrangeira.

8.5 O uso compensatório de gestos e a competência em língua estrangeira

Gregersen, Olivares-Cuhat & Storm (2009) investigaram o fator de inexperiência na língua falada sobre a frequência e o uso compensatório de gestos para limitações linguísticas e lacunas de conhecimento em narrações por sujeitos aprendizes de língua estrangeira. Recrutaram 75 estudantes universitários estados unidenses aprendizes de espanhol em nível iniciante, intermediário e avançado. Os sujeitos foram filmados em conversa em duplas, primeiramente em espanhol, depois em inglês, simulando, um deles um policial investigador e o outro a testemunha de um roubo a banco, evento visualizado previamente apenas por esta última em uma tela de computador e sem áudio.

Para o levantamento da ocorrência dos tipos de gestos, utilizaram a tipologia de Ekman & Friesen (1969), incluindo ambos movimentos manuais dependentes da fala (ou ilustradores) e gestos independentes dela (emblemas, reguladores, adaptadores e mostras de estados emocionais – *affect displays*). No entanto, para a categoria dos gestos ilustradores, diferenciaram *gestos ilustradores* (aqueles em redundância com a fala, que colaboram com a construção do significado da elocução) de *gestos ilustradores compensatórios* (aqueles em auxílio por ocorrência de deficiências linguísticas). Como a produção dos sujeitos variou muito em extensão temporal, calculou-se uma taxa de

¹⁰ Ver Lin, 2020, Laurent & Nicoladis, 2015, para função compensatória por gestos dêiticos e de pulso, respectivamente, e Gregersen, Olivares-Cuhat & Storm, 2009, descrito na seção a seguir, por gestos icônicos e emblemáticos.

produção de gestos por minuto para cada estudante, e depois uma taxa média por grupo: iniciante, intermediário e avançado. Hipotetizaram que a diferença de proficiência entre as línguas aparecesse por diferenças no uso de gestos ilustradores, com maior ocorrência de gestos ilustradores em L1 (língua materna), e de ilustradores compensatórios em L2 (língua estrangeira). Hipotetizaram ainda que, sendo todos fluentes em língua inglesa (L1), não difeririam entre si os grupos (iniciantes, intermediário e avançado) quanto à frequência e uso dos gestos em L1 (língua materna).

Os participantes levaram para a realização da tarefa 3 minutos em média em conversa em L2 e 1 minuto em média em L1. Em L2, aprendizes em nível avançado usaram significativamente mais gestos relacionados à fala, de saliência de significado (gestos icônicos) do que iniciantes e intermediários; não se encontrou diferença entre os grupos quanto à frequência de gestos ilustradores compensatórios, contudo. O grupo iniciante produziu mais gestos emblemáticos do que avançados e intermediários em L2. Algo que destacaram os autores no grupo de nível de fluência intermediário foi o uso de mostras de expressão facial, altamente regulado por convenções culturais. Gestos de pulso foram utilizados por sujeitos em nível avançado em L2 para marcarem

pontuação, coincidindo com acento tônico primário. Ao se comparar a produção de gestos em L1 com a produção de gestos em L2, participantes usaram significativamente mais gestos como um todo em sua narrativa em inglês (L1) do que em espanhol (L2); produziram significativamente mais gestos ilustradores, ilustradores compensatórios e de mostras emocionais em L1 do que em L2 (o mesmo não se observou para gestos adaptadores, emblemáticos e reguladores). Curiosamente, sujeitos de nível avançado apresentaram competência sociolinguística em sua produção verbal e gesticulatória ao se dirigirem com deferência ao policial quando interpretando a testemunha; também os gestos utilizados pelos policiais foram em menor quantidade, mais rápidos e precisos. Conforme previsto, não foram encontradas diferenças entre os grupos quanto à frequência e tipo de gestos entre os grupos para L1 (inglês), o que permitiu aos pesquisadores interpretar seus resultados em atribuição a diferenças quanto ao nível de fluência dos sujeitos, e não quanto a aspectos idiossincráticos. Apresentando maior número de gestos ilustradores, aprendizes com maior nível de proficiência usaram suas mãos de forma a melhorar o significado, aumentando assim a probabilidade de comunicação efetiva entre os interlocutores; usaram as mãos para aumentar a redundância da informação transmitida verbalmente, reduzindo a ambiguidade no discurso. Usaram gestos para acentuar, complementar ou repetir o discurso oral. Por exemplo, observam os autores, enquanto falantes

avançados em L2 usaram a palavra “máscara” com um gesto ilustrador, falantes intermediários usaram gesto compensador apenas (sem fala) e iniciantes simplesmente evitaram o assunto.

Tais resultados sugerem que a competência gesticulatória apresente alterações em consonância com o desenvolvimento da competência verbal na aprendizagem de língua estrangeira. Por análise qualitativa dos vídeos das produções narrativas pelos sujeitos, notaram os pesquisadores que falantes de menor nível de proficiência tenderam a usar suas mãos de maneira improdutivo do ponto de vista comunicativo, com alta ocorrência de gestos adaptadores (de ajuste de roupa, toques no rosto e outras partes do corpo), geralmente em claro sinal de estresse e ansiedade. Com alguma ressalva, tal fenômeno foi atribuído à distração colocada pela resposta emocional ao ambiente e experiência de aprendizagem de língua estrangeira. Para os autores, a presença de uma câmera também teria colaborado para um aumento do nível de estresse dos participantes.

Como vimos pela contemplação dos trabalhos nas seções sobre produções narrativas até aqui, o aumento da frequência de ocorrência de gestos manuais pela variação da carga cognitiva na produção narrativa, em especial em língua estrangeira, pode ser levado em consideração mas não deve ser tomado como índice absoluto de uso compensatório para dificuldades linguísticas; há que se levar em conta as categorias gesticulatórias individualmente, pois podemos observar uso compensatório linguístico em gestos dêiticos, em gestos de pulso, ou ainda em parte da gesticulação icônica, bem como em outras naturezas de gesticulação não-conversacionais. A função compensatória dos gestos pode, ademais, se manifestar de formas diferentes em públicos diferentes, como em crianças e adultos, em função de seu nível de proficiência na língua, como em monolíngues e bilíngues, ou ainda em aprendizes de língua estrangeira.

8.6 Aumento da carga cognitiva e disfluências em gesto e fala

Para além de variações na frequência de ocorrência e nas categorias de gestos, o aumento da carga cognitiva em decorrência de dificuldades linguísticas pode produzir ainda outros fenômenos como a interrupção da realização dos gestos em passagens de hesitação, interrupção ou disfluências na fala (Gullberg, 1998; Mayberry, Jaques & DeDe, 1998; Ragsdale & Silvia, 1982; Seyfeddinipur & Kita, 2001; Graziano & Gullberg, 2013). Graziano & Gullberg (2013), por exemplo, investigaram características gesticulatórias de

gestos compensatórios durante fala fluente e não-fluente por falantes competentes e aprendizes adultos e crianças. Participaram do estudo: 33 crianças italianas, divididas em 3 grupos (de 4 a 5 anos, de 6 a 7, e de 8 a 10 anos de idade); 11 adultos italianos; e 11 adultos holandeses aprendizes de língua francesa em nível pré-intermediário. Para os grupos de crianças e adultos italianos, a tarefa consistiu na recontação de história a partir da observação de um desenho animado; para o grupo de holandeses aprendizes de língua francesa, a tarefa consistiu na recontação de histórias a partir de um livro de imagens. Das produções narrativas, analisaram-se: as pausas preenchidas e não-preenchidas, interrupções e alongamentos silábicos. Disfluências ocorrendo em limites oracionais ou depois de marcadores discursivos foram excluídas da análise, já que sinalizam processo de planejamento ao invés de dificuldade expressiva. Gestos foram classificados entre completos ou interrompidos, e entre referenciais (representacionais) ou pragmáticos (acompanhando funções linguístico-pragmáticas, por exemplo, de análise sintática da fala). As autoras hipotetizaram que a produção de gestos ocorreria predominantemente em passagens disfluente da fala.

Diferentemente de sua hipótese, no entanto, resultados mostraram que ambos falantes competentes e menos competentes, crianças e adultos, produziram gestos predominantemente durante extensões fluentes da fala; curiosamente, porém, no grupo de adultos falantes de língua estrangeira, o alinhamento de gestos com a fala fluente se revelou significativamente menos provável do que em todos os demais grupos. O grupo de falantes de língua estrangeira também produziu significativamente mais gestos em fala disfluente do que adultos competentes e crianças (**Gráfico 3**).

Gráfico 3. Taxa média de ocorrência de gestos com fala fluente e disfluente para os grupos: de crianças com idade de 4, 6 e 9 anos, adultos aprendizes de língua estrangeira (L2), e falantes nativos (L1; Graziano & Gullberg, 2013).

Durante disfluências da fala, gestos em curso foram em sua quase totalidade interrompidos, com muito pouco deles sendo levados a cabo, em ambos grupos de adultos e crianças. Com algumas variações, todos os quatro grupos produziram ambos gestos referenciais e pragmáticos (**Gráfico 4**). Não se observou diferença estatística entre os grupos devido ao reduzido tamanho da amostra e grande variância entre indivíduos.

Gráfico 4. Taxa média de ocorrência de gestos representacionais (R) e pragmáticos (P) para os grupos: de crianças com idade de 4, 6 e 9 anos, adultos aprendizes de língua estrangeira (L2), e falantes nativos (L1; Graziano & Gullberg, 2013).

As autoras concluíram do estudo que ambos crianças e adultos, falantes competentes ou aprendizes, produzem significativamente mais gestos em passagens fluentes da fala, em comparação com passagens de disfluência. Apesar de não diferirem estatisticamente, gestos em passagens disfluentes tendem a ser de modo geral representacionais, mais do que pragmáticos. Ademais, adultos aprendizes de língua estrangeira são mais suscetíveis de realizar gestos em disfluências do que qualquer outro grupo, presumivelmente porque usem gestos representacionais para solicitar ajuda lexical de seus interlocutores. As autoras sugerem o aprofundamento do estudo com crianças e adultos, nativos e aprendizes de língua estrangeira, através da categorização das disfluências, bem como possíveis correlações com os diferentes tipos de gestos.

8.7 Gestos e a hiperexplicitação de referentes discursivos

Um último aspecto relevante que abordaremos neste capítulo é a função da gesticulação na hiperexplicitação de referentes discursivos. A clareza na identificação de referentes é indispensável à inteligibilidade discursiva. Formas atenuadas de rastreamento de referentes, como o uso adequado de pronomes, no entanto, são aquisições linguísticas mais tardias (em relação à aquisição lexical, por exemplo), pois, ao demandarem planejamento linguístico complexo, revelam-se recursos linguísticos cognitivamente exigentes. Para se evitar uma sobrecarga cognitiva, sem, contudo, incorrer em ambiguidades que possam eventualmente comprometer a compreensão do discurso, é comum, como solução mais imediatamente disponível a aprendizes de uma determinada língua, o uso da hiperexplicitação para rastreamento de referentes discursivos, por exemplo, com o uso excessivo de frases nominais lexicais. Sabe-se que quanto menos preditivo, acessível e contínuo um tópico, mais material de codificação é utilizado para representá-lo em linguagem. No entanto, o uso excessivo de material linguístico codificador pela repetição de hiperexplicitação pode tornar confuso o status (novo ou conhecido) de referentes discursivos, atrasando a recuperação de informação pelo ouvinte com prejuízos para a compreensão.

Curiosamente, a hiperexplicitação também pode ser observada na gesticulação que acompanha a fala, e nela permanece mesmo quando ouvintes não têm acesso aos gestos, sugerindo que tenha função para a organização mental discursiva também para o falante; tende, no entanto, a desaparecer com a aquisição e uso adequado do sistema pronominal, sugerindo que a produção de gestos esteja também relacionada à aquisição de expressões referenciais (Gullberg, 1998, 2003, 2006). Em falantes nativos, gestos tenderiam a ocorrer mais em associação com informação imprevisível, por exemplo com primeiras menções de referentes discursivos em narrativas (em comparação com menções subsequentes; Levy & Fowler, 2000; Levy & McNeill, 1992; McNeill, 1992).

Yoshioka (2008) procurou compreender se a ocorrência da hiperexplicitação estaria de fato relacionada à aquisição de sistema pronominal complexo ou, antes, a um fenômeno mais geral de estruturação da informação no discurso em língua estrangeira. Estudou, para tanto, narrativas produzidas em língua nativa por holandeses e japoneses, bem como em língua japonesa por holandeses aprendizes de japonês - o holandês possui um sistema pronomial ativo, enquanto que o japonês carece de marcadores como artigos e pronomes; ademais, Há uma tendência de falantes nativos de língua japonesa de

focalizarem o protagonista de uma narrativa, de forma que o personagem ocupe uma posição tópica ao longo de uma extensão importante do discurso, o que permitiria o encadeamento de uma sequência anafórica sem marcadores (anáfora-zero); já aprendizes tendem a trocar o foco narrativo de um personagem para outro com maior frequência, marcando a reintrodução de referentes muito mais explicitamente com frases nominais lexicais do que o fazem falantes nativos.

Yoshioka (2008) analisou 12 produções narrativas em língua holandesa (L1), 15 produções narrativas em japonês (L1) e 12 produções narrativas em japonês (L2, falantes holandeses aprendizes de língua japonesa em nível pré-intermediário). A tarefa consistiu em recontar de memória a um falante nativo uma história exibida previamente em um livro de imagens. A história contém um personagem principal, 2 secundários, e vários periféricos. Como as narrativas em L1 holandês e em L2 japonês foram produzidas pelos mesmos sujeitos, dados foram coletados com um intervalo de 10 meses. Os dados foram analisados em termos de expressões linguísticas referenciais (locuções nominais lexicais, pronomes, e anáforas - zero-anaphora) e de taxa de gestos acompanhando os respectivos referentes da fala. O pesquisador diferenciou 2 tipos de rastreamento de referentes: manutenção e re-introdução. Manutenção refere-se à manutenção de um referente em posição de sujeito ou tópica introduzido por (ou mantido em relação a) oração precedente; reintrodução, quando o referente já fôra apresentado no discurso mas não corresponde ao sujeito/tópico da oração precedente. O pesquisador codificou então todos gestos em coocorrência com os referentes, incluindo palavras adjacentes. Calculou, a seguir, uma taxa de gestos por referente.

As análises revelaram uma diferença interlinguística quanto à taxa geral de gesticulação como um todo, com maior taxa de gestos para falantes de japonês L1 do que para falantes de holandês L1. Yoshioka interpretou tal resultado observando que, dada a ausência ou desuso de marcadores como artigos e pronomes em língua japonesa, a maior frequência geral de gestos observada em japonês (L1, em comparação com holandês L1) talvez ocorra em auxílio à fala na construção efetiva da informação discursiva, onde o status da informação dos referentes (novo ou dado) seja opaco na construção verbal.

Os dois grupos L1 (japonês e holandês) não apresentaram, no entanto, diferença entre si quanto à frequência de marcação gestual na reintrodução de referentes. Nas produções em holandês (L1), os referentes são mantidos principalmente por pronomes (75% das vezes); enquanto que em japonês (L1), os referentes são mantidos por anáfora-zero (também em 75% das vezes). Também não se observou variação entre as línguas

nativas L1 japonês e L1 holandês quanto à bi-modalidade (gesto-falo) nativa de reintrodução de referentes. Aprendizes, por sua vez, usaram frases nominais lexicais para reintrodução de referentes significativamente mais frequentemente do que falantes nativos. Tais frases foram frequentemente acompanhadas de gesto. Houve coocorrência de gestos junto com referências por frases nominais lexicais ou anáforas-zero. No caso de L1 japonês, a marcação gestual de correferência ocorreu onde a identidade do referente não veio clara do contexto linguístico precedente. Em contraste, em L2 japonês, falantes tenderam a produzir gestos acompanhando referentes de manutenção em passagens cuja identidade do referente já veio explicitamente marcada no contexto linguístico precedente. Yoshioka especula que tal gesticulação marcadora de reintrodução de referentes possa servir ao próprio falante.

Em holandês (L1), os referentes narrativos foram reintroduzidos principalmente por frases lexicais nominais (62%) e por pronomes (36%). De maneira semelhante, em língua japonesa (L1) os referentes narrativos também foram reintroduzidos em sua maioria por frases nominais lexicais (55%); contudo, diferiram do holandês (L1) pelo uso de anáforas-zero (44%, ao invés de pronomes). Já em japonês (L2), 86% dos referentes foram reintroduzidos por frases lexicais nominais, e apenas 13% por anáforas-zero.

Yoshioka (2008) conclui de seu estudo que aprendizes de língua estrangeira hiperexplicitaram a manutenção de referentes pelo uso de frases nominais lexicais, mais do que o fizeram falantes nativos, sem contudo, apresentarem aumento proporcional de gestos; aprendizes se assemelharam aos falantes nativos quanto à frequência da marcação gestual da manutenção de referentes, apesar de, diferentemente de L1, marcarem a manutenção de referentes onde a identidade do referente já estivesse explícita no contexto precedente. Yoshioka (Ibid) salienta que aprendizes foram mais abertamente precisos em marcar a reintrodução de referentes em ambos fala e gesto; hiperexplicitaram em ambas modalidades a reintrodução de referentes. Estudos com falantes e aprendizes de outras línguas serão necessários para confrontamento dos resultados apresentados por Yoshioka (Ibid).

CAPÍTULO 9

A GESTICULAÇÃO COMO ELEMENTO INERENTE À COMUNICAÇÃO E AO APRENDIZADO

Conforme apontamos já no primeiro capítulo, hoje mais do que nunca em nossa nova realidade de relações sociais virtualizadas, urge voltarmos nossa atenção para a relevância da interação social de nossos corpos. Em ambiente de aprendizagem virtual, alunos e professores não devem prescindir de recurso cognitivo social tão caro à interação humana como a gesticulação. Ainda que haja calorosos debates e resultados experimentais contraditórios sobre sua função, ignorá-la é negligenciar parte importante do processo de construção de um espaço cognitivo socialmente compartilhado.

Posicionamentos teóricos, no entanto, precisam ser melhor circunscritos na literatura, uma vez que se apoiam em resultados experimentais não conclusivos, alguns com contingências colocadas pelo próprio desenho experimental, tamanho amostral ou por limitações de alcance de seus resultados. Em se tratando de abordagens incorporadas da cognição, o posicionamento e a movimentação corporal são cruciais para a investigação, de modo que há de se pensar controles experimentais mais ecológicos, com menos desvio atencional e, sobretudo, sem adição de uma dupla tarefa. Também a inconclusão de resultados por razão de amostras populacionais pequenas (algumas poucas dezenas de sujeitos nas maiores amostras, com grande variância nos dados) exige a continuidade das investigações com replicação dos estudos em diferentes populações etárias e linguístico-culturais para avanço do conhecimento. Trabalhos que reportaram efeitos da gesticulação limitados a conteúdo espacial da fala não explicam ou exploram em seus respectivos estudos, por exemplo, dados como o aumento no tempo de produção da tarefa quando na condição de gesticulação, seja em tarefas de explicação conceitual, resolução de problemas matemáticos, ou de produção narrativa. Ademais, como se observa dos estudos referidos nesta introdução, encontramos uma relativa abundância de trabalhos na literatura que exploraram a produção de gestos em produções narrativas. Não obstante, os efeitos facilitadores de gestos no planejamento complexo discursivo macro-linguístico carecem de melhor descrição e explicação.

Um ponto divergente importante na literatura que merece esforço continuado de

investigação são, por um lado, evidências que sustentam a facilitação do acesso lexical como explicação principal da emergência dos gestos (Rauscher, Krauss & Chen, 1996; Krauss, 1998) e, por outro, a clara persistência de gestos icônicos em acompanhamento à fala mesmo após produções subsequentes de uma mesma narrativa (Beattie & Coughlan; 1998), quando o esperado, na perspectiva dos que defendem a função de acesso lexical dos gestos, seria sua diminuição uma vez realizadas as buscas lexicais.

Não podemos deixar de mencionar aqui também o campo talvez de maior disputa: a função comunicativa dos gestos. Enquanto alguns trabalhos trazem evidências que dão sustentação à perspectiva comunicativa dos gestos (Alibali, Kita & Young, 2000; Özyürek, 2002), outros, como o estudo por Krauss, Morrel-Samuels & Colossante (1991), trazem observações intrigantes, como, por exemplo, a de que a discriminação semântica dos gestos possa ser um efeito ilusório de construção *post-hoc* derivada da compreensão da própria fala, gestos veiculando informação largamente redundante e de forma pouco informativa. Como acomodarmos no campo teórico tais observações com resultados de trabalhos experimentais (Beattie & Shevelton, 1999; Pereira & Haddad, 2015) que mostram melhora significativa (18% e 10%, respectivamente) na compreensão de narrativas quando acompanhadas por gestos?

Como podemos depreender dos trabalhos descritos nos capítulos anteriores, apesar de evidências experimentais permitirem diferentes posicionamentos teóricos, não há ainda na literatura consenso sobre a razão da emergência de gestos em acompanhamento à fala, se emergiriam a serviço do próprio falante na elaboração de imagens mentais, por exemplo, ou a serviço da comunicação e em benefício da compreensão por seu interlocutor. Discrepâncias nos resultados dos estudos experimentais nos levam a pensar, de forma alternativa, que a gesticulação emerge como recurso cognitivo com diferentes funcionalidades para falantes distintos e em diferentes circunstâncias. Não obstante, e mais importante do que um posicionamento teórico último e definitivo, o debate ainda hoje muito fértil tem produzido através de novos estudos significativo avanço com produção de novos conhecimentos. Há, portanto, que se perseverar na investigação do custo-benefício cognitivo, linguístico e social de sua produção.

Esse esforço investigativo e de ampliação da compreensão do fenômeno da gesticulação pode resultar, ademais, em conhecimentos de grande valia para a pedagogia e a área da educação. Entretanto, apesar do valor informativo dos gestos sobre avanços conceituais por crianças em ambiente de ensino-aprendizagem (Butcher e Goldin-

Meadow, 2000)¹¹, bem como do valor informativo da gesticulação por professores nas diferentes etapas epistemológicas (Giordan, Silva-Neto & Aizawa, 2015)¹², poucos estudos se deram ao trabalho de descreverem e detalharem o funcionamento da gesticulação na interação professor-aluno no processo de ensino-aprendizagem; poucos estudos investigaram, por exemplo, a fixação ocular dirigida aos gestos em ambiente de aprendizagem preservando aspectos ecológicos das práticas de ensino, como em um ambiente de sala de aula. Ademais, alguns dados ficam inexplorados e carecem de investigação. De nosso conhecimento até esta publicação, nenhum trabalho prévio ao nosso investigou o tempo médio de fixação na região dos gestos; conhecer mais sobre a natureza da fixação ocular dedicada ao campo gestual pode nos dizer sobre nossa intencionalidade atencional, percepção e mesmo consciência de sua presença. Um conhecimento mais aprofundado desses aspectos e do fenômeno como um todo é potencialmente relevante para a passagem do planejamento à execução de atividades pedagógicas na complexa relação entre teoria, metodologia e prática de ensino (Giordan, Silva-Neto & Aizawa, 2015).

Expandirmos nosso conhecimento sobre as diferentes manifestações e funcionalidades dos gestos pode ainda nos ajudar a melhor compreendermos o fenômeno da cognição e da comunicação, bem como trazer contribuições pedagógicas importantes para o próprio desenvolvimento cognitivo de crianças e de adultos em ambiente de aprendizagem.

¹¹ Professores se valeriam da gesticulação das crianças para inferirem momentos de instabilidade cognitiva e potencial avanço conceitual (Goldin-Meadow et al., 1999).

¹² A gesticulação por professores tende a diminuir em grau de iconicidade em acompanhamento ao aumento no grau de abstração de operações epistêmicas em aula de ciências (Giordan, Silva-Neto & Aizawa, 2015).

ESTUDOS EXPERIMENTAIS

OBJETIVO GERAL

Avaliamos no presente estudo o papel da gesticulação na *produção* de narrativas em língua estrangeira. Poucos trabalhos na literatura investigaram o papel de gestos na produção oral livre de narrativas em língua estrangeira por adultos. Hipotetizamos que a gesticulação diminua a carga na memória operacional e melhore o desempenho gramatical-discursivo na tarefa de produção de narrativas por aprendizes de língua inglesa como língua estrangeira.

Buscando aprofundar nosso conhecimento e explicar nossos resultados experimentais (Pereira & Haddad, 2015), exploramos também neste trabalho a produção e o processamento visual de gestos ilustradores no contexto de ensino-aprendizagem de ciências. Objetivamos compreender: i) os diferentes tipos de gestos recrutados pelo professor durante o ensino de conceitos científicos; ii) quais gestos demandam alocação explícita de recursos atencionais (fixações oculares) por sujeitos aprendizes expectadores da aula; iii) o impacto dos diferentes tipos de gestos produzidos pelo professor e das diferentes estratégias de fixação ocular empregadas pelos aprendizes em sua aprendizagem dos conceitos científicos apresentados em aula. Hipotetizamos que: i) gestos de pulso sejam mais frequentemente recrutados em etapas de ensino com maior grau de abstração, como em introduções teóricas e definições conceituais; ii) que gestos icônicos, por outro lado, sejam mais frequentes em etapas de ensino de menor grau de abstração, como em descrições e explicações de exemplos. Hipotetizamos também que gestos icônicos demandem maior alocação explícita da atenção visual – foveamento – quando comparados com os demais tipos de gestos ilustradores, em especial com gestos de pulso. Hipotetizamos ainda que gestos acompanhados da fala – sobretudo gesticulação de pulso – sejam amplamente processados por regiões para-foveais da retina e, majoritariamente, pela via magnocelular e via visual dorsal.

EXPERIMENTO 1

Objetivo

Em continuidade ao estudo iniciado em nível de mestrado, em que investigáramos o impacto da gesticulação do narrador na *compreensão* de narrativas por aprendizes de língua estrangeira (Pereira & Haddad, 2015), conduzimos um novo experimento para avaliarmos o impacto da gesticulação dos aprendizes em sua *produção* oral de narrativas em língua estrangeira. Pretendemos avaliar através da qualidade de produção das narrativas benefícios e prejuízos cognitivos da gesticulação manual livre e espontânea durante a execução da tarefa experimental. Para tanto, controlamos a produção de gestos manuais.

Hipotetizamos que a livre gesticulação manual durante a realização da tarefa diminuísse a carga na memória operacional e melhorasse o desempenho geral dos participantes, por um menor tempo de realização da tarefa, menor número de pausas e hesitações, bem como por uma produção narrativa de melhor qualidade, com menor ocorrências de erros gramaticais e repetições. De forma complementar, hipotetizamos que a condição experimental *sem gestos* aumentasse a carga cognitiva do sujeito experimental durante a realização da tarefa por impedir a distribuição espacial no campo gestual dos referentes discursivos. Por tais razões, hipotetizamos que o tempo total de produção da tarefa na condição *sem gestos* superasse na média geral entre os participantes o tempo total de produção na condição *com gestos*. Também por inibir representações motoras facilitadoras do acesso lexical e dos episódios narrativos em memória, hipotetizamos que a inibição dos gestos por nosso controle experimental promovesse uma diminuição na qualidade gramatical-discursiva, por um maior número de ocorrências de repetições e de hesitações (interrupções sintáticas, alongamentos silábicos e interjeições), mas principalmente de erros gramaticais na condição *sem gestos*.

Participantes

Participaram voluntariamente de nosso estudo experimental 30 sujeitos (21 mulheres; média etária 27, $\pm$ 9), aprendizes de língua inglesa como língua estrangeira

matriculados no curso de extensão INCO-CEPEL, oferecido pela Faculdade de Educação da Universidade de São Paulo à comunidade universitária (funcionários, graduandos e pós-graduandos); todas as informações pessoais dadas pelos participantes para a realização do experimento foram mantidas em sigilo e confidencialidade. Os participantes não tinham informação prévia dos objetivos específicos e das medidas avaliadas no estudo experimental. Os procedimentos metodológicos experimentais foram devidamente submetidos e aprovados pelo Comitê de Ética em Pesquisa com Seres Humanos do Instituto de Psicologia da USP (CAAE 31781414.6.0000.5561, Parecer 2.023.665; **Anexo 1**).

Materiais e métodos

O experimento foi iniciado no Laboratório de Neurociências Cognitivas e Aprendizagem, sala 305 do Instituto de Biociências da USP. Pela dificuldade de recrutamento de voluntários de nosso público-alvo em número suficiente para nosso estudo experimental, mesmo após campanha de divulgação e convites pessoais em visita a salas de aula de cursos de língua estrangeira dentro e fora do campus universitário, optamos por encerrar nossa primeira coleta como estudo piloto.

A partir do estudo piloto, e agora em colaboração com pesquisadores e professores de outros laboratórios e institutos da USP, Laboratório de Ciências da Cognição (IBUSP), Laboratório de Pesquisa em Ensino de Química e Tecnologias Educativas (Lapeq - FEUSP) e o programa de extensão INCO-CEPEL (FEUSP), pudemos discutir mais amplamente e por diferentes perspectivas nossos objetivos e resultados preliminares, a adequação dos estímulos apresentados aos voluntários, nossa população amostral, o procedimento experimental, o local ideal de coleta e a ferramenta de análise de dados.

As sessões experimentais foram então realizadas na sala 306 do bloco A da Faculdade de Educação da USP, em parceria com a coordenação do curso de extensão acadêmica INCO-CEPEL de língua inglesa, oferecido na própria Faculdade de Educação a alunos e funcionários da universidade.

Como estímulo aos participantes, foram selecionadas três tirinhas de histórias em quadrinho. Adotamos como critérios de seleção: temáticas simples, elementos de cenário – objetos e ações – de baixo grau de dificuldade lexical, e mesmo número de quadros (**Figura 16a, b e c**).

A. Narrativa A.

B. Narrativa B.

C. Narrativa C.

Figura 16. Tirinhas de histórias em quadrinho apresentadas como estímulo aos participantes.

Os participantes foram recebidos individualmente e conforme agendamento na sala experimental; foram esclarecidos previamente por um dos experimentadores sobre o estudo, o experimento e demais obrigаторiedades determinadas pelo comitê de ética.

Procedimento experimental

A tarefa dos participantes consistiu na observação de uma das tirinhas exibida no monitor do computador (**Figura 17**). Após trinta segundos de exibição, o monitor era desligado e a cadeira do participante direcionada para a câmera, localizada à sua direita. O participante deveria então recontar o episódio contido na tirinha a um segundo experimentador, de modo que este, sem conhecê-la previamente, pudesse compreendê-la.

O experimentador 2, no entanto, encontrava-se visualmente isolado, sentado atrás de um painel na mesma direção da câmera. Com seu isolamento visual procuramos desestimular nos sujeitos experimentais a produção de gestos visuo-comunicativos; esperamos, dessa forma, a emergência espontânea de gestos oriundos apenas de seu esforço cognitivo de produção oral narrativa em língua estrangeira (**Figura 17**).

Figura 17. *Setup experimental.*

Com cada sujeito foram realizadas três sessões: uma sessão de treino, pela qual procuramos reduzir a ansiedade inicial dos participantes bem como diminuir o efeito de aprendizagem da tarefa; e duas sessões experimentais, uma *com gestos* e a outra *sem gestos*. Na sessão *treino*, bem como na condição experimental *com gestos*, os sujeitos narraram natural e livremente a história. Na condição experimental *sem gestos*, os sujeitos narraram a história segurando uma bola em suas mãos, em uma posição confortavelmente relaxada dos braços, repousados à altura do colo. Para inibirmos movimentos gesticulatórios de braços e mãos, observados nas sessões experimentais da fase piloto do experimento, envolvemos então a bola com uma sacola plástica e a atamos por entre as pernas do participante ao assento de sua cadeira, de forma que não fosse possível movê-la vertical ou lateralmente para além da posição de repouso dos braços sobre o colo do participante. Na tentativa de inibirmos também movimentos gesticulatórios dos dedos, envolvemos bola e mãos (à altura dos dedos) com uma faixa de velcro (**Figuras 18a e 18b**).

Figura 18a. Produção oral na sessão treino e na condição *com gestos*.

Figura 18b. Produção oral na condição *sem gestos*.

A ordem das condições experimentais (*com* e *sem gestos*), bem como das narrativas B e C foram aleatorizadas, cada participante atendendo a uma das quatro possíveis ordens de apresentação (**Tabela 1**).

Treino	1º estímulo	2º estímulo
Narrativa A <i>com gestos</i>	Narrativa B <i>sem gestos</i>	Narrativa C <i>com gestos</i>
Narrativa A <i>com gestos</i>	Narrativa B <i>com gestos</i>	Narrativa C <i>sem gestos</i>
Narrativa A <i>com gestos</i>	Narrativa C <i>sem gestos</i>	Narrativa B <i>com gestos</i>
Narrativa A <i>com gestos</i>	Narrativa C <i>com gestos</i>	Narrativa B <i>sem gestos</i>

Ordem de apresentação

Tabela 1. Possíveis combinações de apresentação dos estímulos utilizadas de forma aleatória entre os participantes.

A produção dos sujeitos experimentais foi registrada em áudio e vídeo para análise posterior à sessão experimental.

ANÁLISE

A produção oral dos participantes registrada em áudio e vídeo foi analisada utilizando-se o software ELAN versão 5.0.0-beta (**Figura 19**), desenvolvido por Max Plank Institute for Psycholinguistics (Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive; <https://archive.mpi.nl/tla/elan>). Foram analisados os seguintes parâmetros:

i) Quanto à realização da tarefa:

- tempo total da tarefa;
- tempo de fala;
- tempo de gesticulação.

ii) Quanto a parâmetros gramaticais:

- número de erros gramaticais;
- número de autocorreções;
- número de repetições;
- variabilidade de substantivos.

iii) Quanto ao discurso:

- número de hesitações;
- número de orações;
- número de interrupções sintáticas.

Figura 19. Análise realizada com software ELAN 5.0.0-beta.

Análise do discurso oral

Consideramos erros gramaticais: tanto i) escolhas lexicais e sintáticas que prejudicaram a comunicação e a compreensão da narrativa, sobretudo trocas de valor semântico em palavras tais substantivos, adjetivos, pronomes, verbos e advérbios, quanto ii) erros gramaticais de desacordo com a norma padrão da língua inglesa, ainda que não tenham comprometido a construção semântica narrativa, tais flexões de número, pessoa e tempo em verbos.

Consideramos repetições lexicais e sintáticas as estruturas puramente ecoadas, ou seja, aquelas que não foram objeto de reformulação. As repetições com reformulações foram por nós consideradas como esforço de adequação de forma-sentido, ou seja autocorrekções estruturais-semânticas, e classificadas, por isso, como autocorrekções gramaticais.

Em vista da grande variabilidade da qualidade fonológica na produção dos sujeitos, bem como da dificuldade de se estabelecerem valores seguros de referência para uma análise criteriosa com concessões a diferenças individuais de pronúncia e de acentuação frasal, optamos por desconsiderarmos aspectos fonético-fonológicos como critério de análise da qualidade linguística das produções.

Calculamos a variabilidade de substantivos como um índice do número de substantivos diferentes (sem as repetições) dividido pelo número total de substantivos (inclusas as repetições).

Como hesitações, incluímos alongamentos silábicos e interjeições, tais “ah...”, “hm...” e “é...” dentre outras.

Pela grande variância observada nos dados, para além da análise do *número de ocorrências* dos parâmetros linguísticos, optamos por analisarmos nossos dados também calculando-se para os respectivos parâmetros sua respectiva *taxa de frequência* de ocorrência por segundo de fala. Para a análise das *taxas de frequência* foram empregadas transformações de arcoseno (para transformações arcoseno, ver Howell, 2002). Apresentamos aqui, contudo, os dados em sua escala original para sua melhor interpretação.

Análise dos gestos

Quanto ao uso de gestos como parâmetro linguístico, para fins de computação do tempo de gesticulação, foram considerados apenas os gestos ilustradores, ou seja, gestos manuais espontâneos diretamente relacionados à fala, promovidos pelo esforço cognitivo de narração das histórias em língua estrangeira. Desconsideramos, portanto, gesticulação de outras partes do corpo, como da cabeça e de pernas, bem como gesticulação de qualquer outra natureza de gestos, como os gestos adaptadores ou de estados emocionais, gestos que pudessem, por exemplo, estar relacionados ao nível de ansiedade dos sujeitos durante a tarefa. A contabilização do tempo de gesticulação foi feita pela soma do tempo das ocorrências durante a produção oral do sujeito na tarefa, cada ocorrência sendo considerada do instante inicial da preparação do gesto até a conclusão de sua retração, ou seja, de repouso das mãos e braços pré- a repouso pós-golpe gesticulatório (para a anatomia dos gestos, ver McNeill, 2005).

Os sujeitos experimentais provieram de níveis de turmas de estudo de inglês distintos (I a VI) de um mesmo curso regular de língua inglesa oferecido pela Faculdade de Educação da Universidade de São Paulo. No entanto, o procedimento padrão do curso é de receber todos novos ingressantes como alunos de nível I, para somente após o primeiro semestre de estudo realizarem nivelamento dos alunos. Assim, recebemos para nosso estudo experimental participantes matriculados em nível I com nível de fluência de

elementar a avançado. Por tal razão, bem como pela ampla variância entre os participantes de experiências com a língua inglesa (diferentes modalidades e mídias, viagens ao exterior, uso profissional da língua, entre outros), também por nosso reduzido tamanho amostral, optamos por não tratá-los em subgrupos de níveis distintos de proficiência.

Efeito de aprendizagem e nível de estresse

Procuramos reduzir o efeito de aprendizagem da tarefa e a ansiedade inicial dos participantes inserindo na sequência experimental uma sessão inicial de treino. Não obstante, percebemos ainda a necessidade de explorarmos em nossos dados coletados possíveis índices de estresse promovido por nosso controle experimental. Em situação de estresse e ansiedade, sujeitos apresentam empobrecimento linguístico em produções narrativas orais, verificado pela relação entre o número de palavras diferentes e o número total de palavras utilizadas; o tratamento de elementos concretos no discurso pelo uso de substitutos estaria diretamente relacionado ao relaxamento ou rigidez corporal durante a realização da tarefa (Höweler, 1972). Elaboramos, assim, para cada produção oral um índice de variabilidade de substantivos, calculado a partir do número de substantivos diferentes (excluídas as repetições) pelo número total de substantivos (inclusas as repetições).

Análises estatísticas

Para as análises estatísticas, as premissas de normalidade de distribuição de dados e homogeneidade de variância foram verificadas pelos testes de Kolmogorov-Smirnov e Levine, respectivamente. Adotamos como critério de significância estatística 5% ($p \leq 0,05$). Para as análises de ocorrência e de frequência dos parâmetros linguísticos, utilizamos o teste de ranqueamento de amostras relacionadas de Wilcoxon (Related-Samples Wilcoxon Signed Rank Test), com correções de α pelo método de Bonferroni. Abrimos mão, porém, das correções de α ao abtermos como resultado falsos negativos, pela grande quantidade de itens analisados. Todas as análises foram conduzidas em IBM SPSS Statistics versão 20.0 (IBM Corp., 2011).

RESULTADOS

Geral

Excluímos de nossa análise um sujeito experimental por suas produções terem sido realizadas em língua portuguesa, não cumprindo, portanto, com a instrução da tarefa; e outros dois participantes por terem se sentado sobre as mãos durante a realização da tarefa e não terem, portanto, gesticulado na condição *com gestos*, invalidando assim nosso controle experimental.

Na média geral entre as produções (n=54), observamos os seguintes resultados:

Média geral (n=54)	Média	Mínimo	Máximo
<i>Tempo total</i> de produção (segundos)	57,8 ($\pm$ 5,6)	15,3	157,2
<i>Tempo de fala</i> (segundos)	26,31 ($\pm$ 1,9)	5,9	66,0
<i>Erros</i>	8,3 ($\pm$ 0,7)		
<i>Autocorreções</i>	1,9 ($\pm$ 0,3)		
<i>Repetições</i>	3,6 ($\pm$ 0,5)		
<i>Hesitações</i>	9,6 ($\pm$ 1,0)		
<i>Orações</i>	6,4 ($\pm$ 0,4)		
<i>Interrupções sintáticas</i>	23,8 ($\pm$ 2,0)		
<i>Substantivos TOTAIS (com repetições)</i>	9,4 ($\pm$ 0,6)		
<i>Substantivos DIFERENTES</i>	5,1 ($\pm$ 0,3)		

Tabela 2. Médias gerais (n=54) de desempenho na tarefa ($\pm$ erro padrão).

A

B

Gráficos 5a e b. Médias gerais (X) e medianas (traço) dos *tempos totais* e *de fala* (em segundos).

Gráfico 6. Médias gerais (X) e medianas (traço) do número de ocorrências linguísticas nas condições experimentais *com gestos* e *sem gestos*.

Notamos para alguns dos parâmetros analisados a presença de valores marginalmente *outliers*; nenhum, contudo, ultrapassou a marca de 3 desvios padrões. Optamos por mantê-los todos em nossas análises devido ao nosso reduzido tamanho amostral.

Uma análise de regressão linear múltipla foi utilizada para verificar a previsibilidade do *tempo total* de produção das narrativas pelas variáveis aqui consideradas como parâmetros linguísticos analisados. A análise resultou em um modelo estatisticamente significativo [$F(2,51) = 18,749$; $p = 0,001$; $R^2 = 0,424$], com o número de *hesitações* ($\beta = 0,470$; $t = 3,677$; $p = 0,001$) e o número de *repetições* ($\beta = 0,260$; $t = 2,038$; $p = 0,047$) como previsores do *tempo total* de produção das narrativas.

A análise de regressão linear múltipla também foi utilizada para verificar a previsibilidade do *tempo de fala* das narrativas pelos parâmetros linguísticos analisados. A análise resultou em um modelo estatisticamente significativo [$F(2,49) = 33,009$; $p < 0,001$; $R^2 = 0,729$], com o número de *autocorreções* ($\beta = 0,398$; $t = 4,525$; $p < 0,001$), o número de *orações* ($\beta = 0,369$; $t = 4,281$; $p < 0,001$), o número de *repetições* ($\beta = 0,231$; $t = 2,570$; $p = 0,013$), e o número de *substantivos diferentes* ($\beta = 0,201$; $t = 2,252$; $p = 0,029$) como previsores do *tempo de fala* das narrativas.

Uma análise não-paramétrica de amostras independentes (*Independent-Samples Median Test*) mostrou não haver efeito de narrativa (narrativa B X narrativa C) sobre o *tempo total* de produção da tarefa (Mann-Whitney $U = 354,0$; $Z = -0,182$; $p = 0,856$); tampouco sobre o *tempo de fala* (Mann-Whitney $U = 351,0$; $Z = -0,234$; $p = 0,815$). O teste de Mann-Whitney mostrou ainda não haver efeito de ordem de apresentação (*primeiro estímulo* X *segundo estímulo*) sobre o *tempo total* de produção da tarefa (Mann-Whitney $U = 331,0$; $Z = -0,580$; $p = 0,562$) ou sobre o *tempo de fala* (Mann-Whitney $U = 351,5$; $Z = -0,225$; $p = 0,822$). O teste de Mann-Whitney mostrou ainda não haver efeito de narrativa ou de ordem de apresentação dos estímulos sobre o número de ocorrência dos parâmetros linguísticos considerados ($p > 0,05$).

Entre as condições experimentais

Uma análise não-paramétrica de medidas relacionadas (*Related-Samples Wilcoxon Signed Rank Test*) mostrou diferença estatisticamente significativa entre a

produção dos participantes na condição experimental *com gestos* e a produção na condição experimental *sem gestos*, tanto para o *tempo total* de produção ($Z = -3,291$; $p = 0,001$), quanto para o *tempo de fala* ($Z = -2,972$; $p = 0,003$). Para nossa surpresa, em direção oposta à nossa hipótese inicial, o *tempo total* da tarefa na condição *com gestos* foi em média 38% maior do que na condição *sem gestos*; também o *tempo de fala* na condição *com gestos* foi em média 23% superior em comparação com a condição *sem gestos* (**Tabela 3, Gráfico 7**). No entanto, ao analisarmos a proporcionalidade do *tempo de fala* pelo *tempo total* da tarefa, ou seja, a *taxa de fala*, verificamos que esta não diferiu entre as condições experimentais (Wilcoxon; $Z = -0,024$; $p = 0,981$; **Tabela 3, Gráfico 8**).

Quanto à *taxa de gesticulação* (condição *com gestos*), os sujeitos gesticularam em média 88% ($\pm 0,1$) do *tempo de fala*, e 47% ($\pm 0,1$) do *tempo total* da tarefa. Contabilizamos e apresentamos aqui ambas *taxas de gesticulação* em relação ao *tempo de fala* e *tempo total* de produção da tarefa, pois os sujeitos participantes produziram gestos, via de regra, acompanhados de vocalização pela fala, porém, também em silêncio em algumas poucas ocasiões (**Tabela 3**).

A análise de parâmetros linguísticos gramaticais e discursivos mostrou uma tendência geral com maior número de ocorrências na condição experimental *com gestos* (**Tabela 3; Gráfico 10**), acompanhando, assim, o resultado das análises do *tempo total* e do *tempo de fala* acima descritas.

Investigando a qualidade linguístico-discursiva, fizemos o levantamento da *variabilidade de substantivos*, calculada pelo número de *substantivos diferentes* (não repetidos) dividido pelo número *total de substantivos* (repetidos ou não) presentes na narrativa. Interessantemente, conforme mostra a **tabela 3**, o número de *substantivos diferentes* utilizados na produção narrativa não diferiu entre as duas condições experimentais: 5,11 ocorrências para ambas condições *com gestos* e *sem gestos* ($\pm 0,4$ e $0,4$, respectivamente; Wilcoxon; $p > 0,05$). Já o número *total de substantivos* utilizados (incluindo substantivos repetidos) tendeu a um maior valor para a condição *com gestos* ($10,0$; $\pm 0,9$) em comparação com a condição *sem gestos* ($8,7$; $\pm 0,9$), sem, contudo, apresentar diferença estatística significativa (Wilcoxon; $Z = -1,351$; $p = 0,177$). No entanto, ao calcularmos a *variabilidade de substantivos* (número de *Substantivos diferentes* dividido pelo número *total de substantivos*), encontramos diferença significativa (Wilcoxon; $Z = -2,288$; $p = 0,022$) com média entre os participantes de $0,55$ ($\pm 0,0$) na condição *CG* e $0,67$ ($\pm 0,0$) na condição *SG*.

Sabendo-se distintas as produções das narrativas entre as condições experimentais pela análise do *tempo total* de produção da tarefa e do *tempo de fala*, procuramos responder quais dos parâmetros linguísticos analisados explicariam melhor a diferença observada entre as condições. Uma análise não-paramétrica de medidas relacionadas (Related-Samples Wilcoxon Signed Rank Test) mostrou diferença estatisticamente significativa quanto a: o número de *repetições* ($Z = -2,725$; $p = 0,006$), o número de *hesitações* ($Z = -2,670$; $p = 0,008$), e o número de *interrupções sintáticas* ($Z = -2,077$; $p = 0,038$). Diferentemente do que esperávamos, não encontramos diferença significativa quanto ao número de *erros* ($Z = -1,662$; $p = 0,096$); encontramos, porém, diferença marginalmente significativa entre as condições experimentais quanto ao número de *autocorreções* ($Z = -1,914$; $p = 0,056$). Não encontramos diferença, quanto a: o número de *orações* ($Z = -1,174$; $p = 0,241$), o número *total de substantivos* ($Z = -1,351$; $p = 0,177$), e o número de *substantivos diferentes* ($Z = 0,000$; $p = 1,000$; **Gráfico 3**).

Buscando ainda, dentre os parâmetros linguísticos analisados, fatores que explicassem a diferença entre as condições experimentais nos valores médios registrados para o *tempo de fala*, realizamos análises de regressão linear múltipla, resultando em um modelo estatisticamente significativo [$F(2,24) = 9,343$; $p = 0,001$; $R^2 = 0,438$] com o número de *repetições* ($\beta = 0,397$; $t = 2,319$; $p = 0,029$) e o número de *autocorreções* ($\beta = 0,380$; $t = 2,221$; $p = 0,036$) como fatores preditores do *tempo de fala* na condição *com gestos*; e em um modelo estatisticamente significativo [$F(2,22) = 140,380$; $p < 0,001$; $R^2 = 0,962$] com o número de *interrupções sintáticas* ($\beta = 0,410$; $t = 4,558$; $p < 0,001$), o número de *autocorreções* ($\beta = 0,312$; $t = 4,740$; $p < 0,001$), o número *total de substantivos* ($\beta = 0,252$; $t = 3,518$; $p = 0,002$) e o número de *orações* ($\beta = 0,180$; $t = 3,411$; $p = 0,003$) como fatores preditores do *tempo de fala* na condição experimental *sem gestos*.

Realizamos ainda outras duas análises de regressão linear múltipla: uma para o *tempo total* da tarefa e outra para o *tempo de fala*; porém, agora considerando tanto para a variável de previsão quanto para as variáveis preditoras a *diferença* entre as condições experimentais, ou seja, os valores considerados como os valores registrados para a *condição com gestos* menos os valores registrados para a condição *sem gestos*. A análise resultou em um modelo estatisticamente significativo [$F(2,23) = 27,733$; $p < 0,001$; $R^2 = 0,644$] com a diferença ($CG - SG$) no número de *hesitações* ($\beta = 0,595$; $t = 4,428$; $p < 0,001$) e de *repetições* ($\beta = 0,343$; $t = 2,550$; $p = 0,018$) como previsores da diferença ($CG - SG$) no *tempo total* entre as narrativas. Um modelo estatisticamente significativo [$F(2,23) = 25,427$; $p < 0,001$; $R^2 = 0,768$] com a diferença ($CG - SG$) no número de

orações ($\beta = 0,528$; $t = 4,504$; $p < 0,001$), de *repetições* ($\beta = 0,303$; $t = 2,756$; $p < 0,011$) e de *substantivos totais* ($\beta = 0,301$; $t = 2,410$; $p = 0,024$) como previsores da diferença ($CG - SG$) no *tempo de fala* entre as narrativas.

Diante dos resultados obtidos a partir das análises iniciais, compreendemos que seria esperado que uma produção mais longa (a condição *com gestos*) apresentasse proporcionalmente maior volume de ocorrências linguístico-discursivas. Buscamos então investigar se as produções diferiram entre as condições experimentais não apenas pela sua extensão e número de ocorrências que a acompanham, mas também se diferiram entre si qualitativamente. Para tanto, consideramos os parâmetros linguísticos analisados a partir de sua *taxa de frequência*, ou seja, seu número de ocorrências por minuto de *tempo de fala*. Uma análise não-paramétrica de medidas relacionadas (Related-Samples Wilcoxon Signed Rank Test) mostrou diferença estatisticamente significativa quanto a: *taxa de repetições* ($Z = -2,057$; $p = 0,04$) e *taxa de orações* ($Z = -3,123$; $p = 0,002$). Não encontramos diferença significativa quanto a: *taxa de erros* ($Z = -3,12$; $p = 0,755$), *taxa de autocorrekções* ($Z = -1,628$; $p = 0,104$), *taxa de hesitações* ($Z = -0,865$; $p = 0,387$), *taxa de interrupções sintáticas* ($Z = -0,889$; $p = 0,377$; **Tabela 3; Gráfico 11**).

Média geral		Com gestos	Sem gestos
<i>Tempo total de produção (segundos)</i>		**66,9 (± 9,2)	48,6 (± 5,9)
<i>Tempo de fala (segundos)</i>		**29,0 (± 3,0)	23,6 (± 2,4)
<i>Taxa de fala</i>		55,2% (± 0,04%)	55,8% (± 0,04%)
<i>Taxa de gesticulação</i>	<i>Por tempo de fala</i>	88% (± 0,1)	
	<i>Por tempo total</i>	47% (± 0,1)	
Ocorrências			
<i>Erros</i>		9,07 (± 1,1)	7,48 (± 0,9)
<i>Autocorreções</i>		2,33 (± 0,4)	1,48 (± 0,3)
<i>Repetições</i>		** 4,85 (± 0,9)	2,44 (± 0,5)
<i>Hesitações</i>		** 11,0 (± 1,7)	8,15 (± 1,1)
<i>Orações</i>		6,37 (± 0,6)	6,56 (± 0,6)
<i>Interrupções sintáticas</i>		* 26,07 (± 3,2)	21,48 (± 2,4)
<i>Substantivos TOTAIS</i>		10,0 (± 0,9)	8,7 (± 0,9)
<i>Substantivos DIFERENTES</i>		5,1 (± 0,4)	5,1 (± 0,4)
<i>Taxa de variabilidade de substantivos</i>		* 0,55 (± 0,0)	0,67 (± 0,0)
Taxas de ocorrências (por minuto de fala)			
<i>Taxa de erros</i>		19,9 (± 1,7)	21,7 (± 2,8)
<i>Taxa de autocorreções</i>		5,0 (± 0,9)	3,2 (± 0,6)
<i>Taxa de repetições</i>		* 9,7 (± 1,6)	6,0 (± 1,2)
<i>Taxa de hesitações</i>		20,7 (± 1,8)	20,1 (± 1,6)
<i>Taxa de orações</i>		** 14,4 (± 1,3)	18,4 (± 1,3)
<i>Taxa de interrupções sintáticas</i>		52,8 (± 2,5)	54,5 (± 2,3)

Tabela 3. Médias gerais de desempenho na tarefa (± erro padrão) para as condições experimentais *com gestos* e *sem gestos* (Wilcoxon; * $p < 0,05$, ** $p < 0,01$).

Gráfico 7. Tempo total e tempo de fala (em segundos) nas condições experimentais com gestos e sem gestos.

Gráfico 8. Taxa de fala (tempo de fala dividido pelo tempo total da tarefa) nas condições experimentais com gestos e sem gestos.

Gráfico 9. Taxa de gesticulação (calculadas pelo tempo de fala dividido pelo tempo total da tarefa) nas condições experimentais com gestos e sem gestos.

Gráfico 10. Número de ocorrências linguísticas nas condições experimentais *com gestos* e *sem gestos*.

Gráfico 11. Taxa de desempenho (número de ocorrências por minuto de fala) nas condições experimentais *com gestos* e *sem gestos*.

DISCUSSÃO PARCIAL

O resultado de nossas análises mostrou diferença significativa entre as condições experimentais com produções narrativas mais longas na condição experimental com gestos, acompanhadas, de forma geral, por maior número de ocorrências dos parâmetros linguísticos analisados. A *taxa de fala*, no entanto, foi praticamente a mesma para ambas condições experimentais. Também não diferiram significativamente entre as condições experimentais as taxas ou frequências de ocorrência por minuto de fala dos parâmetros linguísticos analisados, à exceção da *taxa de repetições* (superior na condição *com gestos*) e da taxa de orações (inferior na condição *com gestos*). De maneira geral, o número de *hesitações* e o número de *repetições* se mostraram fatores preditores do *tempo total* de realização da tarefa. Na condição *com gestos*, o número de *repetições* e o número de *autocorreções* se revelaram fatores preditores do *tempo de fala*; e, na condição *sem gestos*, os preditores do *tempo de fala* se revelaram o número de *interrupções sintáticas*, o número de *autocorreções*, o número *total de substantivos* e o número de *orações*.

Buscamos em nosso experimento 1 compreender o impacto da gesticulação manual na produção de narrativas por aprendizes de língua estrangeira. Nosso desenho experimental foi extensamente discutido por nosso grupo de pesquisa e outros colaboradores com base em outros trabalhos e em nosso estudo piloto. Uma análise qualitativa dos vídeos das produções narrativas dos participantes mostrou que nosso controle experimental foi bem-sucedido no controle da gesticulação manual dos voluntários. Contudo, mesmo após diversas adequações experimentais, não foi possível eliminar completamente movimentos corporais de natureza gesticulatória na condição experimental *sem gestos*; ainda que com as mãos atadas à bola, observamos – para nossa surpresa – movimentos de pequena amplitude dos dedos e braços dos voluntários na condição *sem gestos* durante seu esforço linguístico discursivo. Tais movimentos compreendemos, porém, não se realizarem como gestos por sua execução mínima (preparação, golpe e retração; consultar McNeill, 2005, para anatomia dos gestos) não ter sido levada a cabo, justamente por impedimento de nosso controle experimental; sua emergência, contudo, pareceu-nos ter mesma origem cognitiva e motora dos gestos.

Tal fenômeno corrobora a proposição de McNeill (2005) de que gestos operem em uma mesma economia cognitiva integrados à fala. Conforme observado na análise qualitativa dos vídeos das produções dos sujeitos, a inibição da execução dos gestos por nosso controle experimental não impediu em alguns casos a participação da componente motora na elaboração cognitiva da narrativa. Não sendo possível, portanto, o controle sobre o planejamento dos gestos, a simples inibição de sua realização foi suficiente em nosso experimento para observarmos interferências significativas na qualidade linguístico-discursiva das produções. Para controle sobre o planejamento motor, poderíamos ter instruído expressamente os participantes a não gesticularem ou a não produzirem movimentos corporais durante a realização da tarefa na condição *sem gestos*; a instrução, no entanto, adicionaria à condição experimental uma segunda tarefa, aumentando-lhes muito provavelmente a carga cognitiva, resultando em piora no desempenho.

Estudos anteriores também se depararam em seu desenho experimental com o mesmo desafio de controle dos gestos sem menção explícita à gesticulação e adição de uma segunda tarefa: Rauscher, Krauss & Chen (1995), por exemplo, utilizaram-se de falsos eletrodos, instalados aos tornozelos dos participantes na condição *com gestos*, e na cadeira, à altura dos apoios de braço, na condição *sem gestos*. Aos sujeitos participantes era dito que sua atividade eletrodérmica seria monitorada durante o experimento. Na condição *sem gestos*, porém, os sujeitos eram instruídos a manterem as palmas das mãos em contato com os eletrodos nos apoios de braço. A solução encontrada pelos pesquisadores de controle da gesticulação durante a execução da tarefa experimental seria a princípio bastante válida por controlar também um possível fator estressante causado pela presença dos eletrodos (mãos atadas à bola em nosso caso) ao prevê-la em ambas condições experimentais. No entanto, não foi possível aos pesquisadores contornarem o problema da adição de uma segunda tarefa na condição *sem gestos*: na condição *sem gestos*, os participantes eram instruídos a manterem as palmas das mãos em contato com os eletrodos nos apoios de braço. Em nosso desenho experimental, o controle por nós adotado das mãos atadas à bola resolveu a questão da dupla tarefa; por outro lado, o fator de potencial estresse pela presença da bola poderia ter sido ele também controlado simplesmente atando-se as pernas ou os tornozelos dos sujeitos à bola na condição *com gestos*.

Com a finalidade de verificarmos se nossos resultados entre as condições experimentais *com gestos* e *sem gestos* pudessem ter sofrido efeito de um eventual fator

estressante de nosso controle experimental, buscamos na literatura possíveis índices de estresse em produções orais. Como uma primeira opção, cogitamos examinar a variação da altura média da voz dos sujeitos em suas produções orais entre as condições experimentais; alguns trabalhos mostram correlação entre a altura média da voz e o nível de ansiedade e estresse. Essa primeira possibilidade de análise foi, todavia, logo descartada devido à baixa qualidade do registro de áudio das produções orais dos participantes em nosso experimento; durante as sessões experimentais o áudio fôra capturado pelo microfone interno à câmera, e esta posicionada à distância dos participantes, com prejuízo à qualidade do áudio pela presença de ruídos externos à sala experimental. Um futuro estudo com desenho experimental que utilize microfones auxiliares e isolamento acústico, que permita, portanto, o controle da qualidade de áudio, poderá contribuir na elucidação da questão.

Encontramos, porém, na literatura, solução alternativa para nossa questão através da variedade de substantivos. Uma maior variabilidade no uso de substantivos, calculada pelo número de substantivos diferentes dividido pelo número total de substantivos enunciados (repetidos ou não), indicaria uma produção de maior riqueza vocabular, obtida em uma condição ambiental de menor tensionamento e um estado mental de maior relaxamento e descontração; uma condição de menor estresse, portanto. Inversamente, uma menor variabilidade de substantivos (ou maior número proporcional de repetições) indicaria uma produção em condição ambiental de maior pressão e um estado mental submetido a maior tensionamento, uma condição estressante, portanto (Höweler, 1972).

Procedemos então ao levantamento do número *total de substantivos* (inclusas as repetições), do número de *substantivos diferentes* (excluídas as repetições), e extraímos o índice de *variabilidade de substantivos*, calculado pelo número de *substantivos diferentes* dividido pelo número *total de substantivos* presentes na produção. Para nossa surpresa, e um pouco divergente dos resultados reportados por Goldin-Meadow, Nusbaum, Kelly & Wagner (2001), encontramos exatamente mesmo valor médio de *substantivos diferentes* para as duas condições experimentais ($5,11, \pm 2,04$ para a condição *com gestos*, e $5,11, \pm 2,03$ para a condição *sem gestos*; Goldin-Meadow et al., 2001, reportaram maior evocação de memória por alívio de carga cognitiva dos gestos), um forte indício de uma mesma quantidade substantiva de informação (de elementos concretos do cenário) mobilizada para a produção das narrativas em ambas condições experimentais. Entendemos, contudo, que, para além das diferenças quanto ao desenho e tarefa experimentais em relação ao trabalho de Goldin-Meadow et al. (2001), em nosso estudo, as narrativas se

constituíram de narrativas simples com poucas personagens e elementos concretos, de forma que o aumento da demanda cognitiva na tarefa esteve antes relacionado com o grau de complexidade de construções linguísticas em língua estrangeira do que com a evocação de imagens e episódios da memória.

Já quanto ao número *total de substantivos* (inclusos substantivos repetidos), encontramos uma tendência de maior número de ocorrências na condição *com gestos* (10,0, $\pm 0,9$ para a condição *com gestos* e 8,7, $\pm 0,9$ para a condição *sem gestos*), corroborando o resultado das análises dos dados de *repetições* e apontando aqui também para a interpretação de que a diferença encontrada na produção das narrativas entre as duas condições experimentais se explicaria menos por diferenças qualitativas de acréscimo informacional e mais por diferenças quantitativas por repetições, com maior grau de redundância em produções na condição experimental *com gestos* (**Tabela 3**).

Em relação à *taxa de variabilidade de substantivos*, apesar de não nos permitir descartar a hipótese de nossos resultados se explicarem por um fator estressante de nosso controle experimental na condição *sem gestos*, ao mostrar menor *variabilidade de substantivo* na condição *com gestos* (0,55, $\pm 0,0$ para a condição *com gestos* e 0,67, $\pm 0,0$ para a condição *sem gestos*; $p < 0,05$), corrobora nossa interpretação principal dos resultados de uma produção oral mais repetitiva e expletiva, ruidosa e custosa linguisticamente na condição experimental *com gestos*.

Permanece, portanto, na literatura, não mensurada a carga cognitiva advinda dos diferentes controles experimentais empregados nos trabalhos até agora para controle do fator gesticulação: a instrução à não-gesticulação, o espalmar das mãos sobre a mesa, o uso de eletrodos, o cruzamento dos braços, o sentar-se sobre as mãos, o uso de objetos como bola, o impedimento de apenas uma das mãos, dentre outros. Traria grande contribuição para a área um estudo que comparesse os diferentes controles experimentais e realizasse essa mensuração. Medidas indiretas como a condutância galvânica da pele e comportamento ocular (o número de fixações como índice de simulações mentais no campo visual associadas à sobrecarga na memória operacional; Pow et al., 2016) podem ser de grande auxílio na investigação.

Ademais um potencial estresse pelo controle experimental, admitimos em nosso experimento aumento de carga cognitiva por a realização da tarefa de produção narrativa ser solicitada em língua estrangeira. Sobre este ponto, a replicação do experimento com a adição de sessões experimentais com produções narrativas em língua materna facilmente forneceria uma linha de base comparativa. Ademais os parâmetros linguístico-

discursivos, a frequência de gestos adaptadores e de estados emocionais poderia, nesse caso, ela também ser utilizada como índice comparativo de carga cognitiva.

Verificamos também através da análise qualitativa das produções, na direção apontada pela literatura (Goldin-Meadow, 2003), que outras partes do corpo (como cabeça, ombros e membros inferiores) assumem movimentos gesticulatórios quando mãos e braços estão tomados por outra função ou impedidos de se movimentarem, sugerindo em alguns dos casos analisados mesmo uma inexorabilidade de movimentos gestuais acompanhando a fala. Para fins de delimitação de nosso objeto de estudo, desconsideramos em nossa análise a gesticulação da cabeça e pernas. Seria, contudo, de grande interesse investigar se tais movimentações corporais de menção gestual substituem em natureza e frequência a gesticulação manual inibida por nosso controle experimental na condição *sem gestos*.

Perguntamo-nos por fim sobre nosso desenho experimental se, estivessem os sujeitos frente a frente e em contato visual com o experimentador-expectador, apresentariam os sujeitos durante a realização da tarefa aumento na carga cognitiva e se de fato – e de que forma – a natureza da gesticulação sofreria alterações. A replicação de nosso experimento controlando o contato visual dos sujeitos com seu interlocutor poderá nos fornecer algumas dessas respostas.

Participaram de nosso estudo experimental 30 sujeitos (21 mulheres; média etária $27, \pm 9$), tendo sido três dentre eles excluídos por não cumprimento da tarefa. Com tempo livre para realização da tarefa, observamos de maneira geral entre os participantes grande variabilidade de *tempo de fala* (média de $26,31; \pm 1,9$) e de *tempo total* (média de $57,8$ segundos; $\pm 5,6$). Também encontramos grande variância na *taxa de gesticulação* (na condição *com gestos*): os sujeitos gesticularam em média 88% do *tempo de fala* ($\pm 0,1$), e 47% do *tempo total* da tarefa ($\pm 0,1$; **Tabela 3**). Rauscher, Krauss & Chen, 1991, no entanto, reportaram taxas de gesticulação inferiores aos valores por nós observados: média de 51% do tempo de fala para conteúdo espacial e 17% para conteúdo não-espacial. A explicação para a diferença entre os resultados reportados por Rauscher et al. (ibid.) e os resultados por nós observados quanto à taxa de gesticulação poderia estar relacionada tanto ao fator língua estrangeira (no trabalho de Rauscher et al., os sujeitos narraram em sua língua nativa, o inglês), bem como em possíveis diferenças no critério de computação do tempo de gesticulação (os autores não explicitaram a categoria dos gestos levados em consideração no estudo experimental, tampouco seus critérios de computação do tempo de execução).

Nossos resultados também reiteram resultados de estudos anteriores que relataram grande variância entre sujeitos experimentais quanto à frequência da gesticulação e à extensão de produções narrativas em produções orais livres (ver Jenkins, Coppola & Coelho, 2017; e Laurent, Smithson & Nicoladis, 2020). É importante notar que Laurent, Smithson & Nicolades (2020) estudaram, porém, produções narrativas por crianças pré-escolares; de forma complementar, nosso estudo, assim como o trabalho de Jenkins, Coppola & Coelho (2017), analisou produções narrativas por adultos. A comparação entre trabalhos conduzidos com crianças e adultos é relevante para uma melhor compreensão do fator etário na gesticulação. Quanto à variância na frequência de gesticulação pelos sujeitos durante a produção narrativa reportada nos trabalhos, esta poderia ser explicada pelo grau com que lançam mão em suas construções narrativas de recursos cognitivos visuo-espaciais: um maior uso desses recursos implicaria em produções narrativas mais complexas e extensas (Smithson & Nicoladis, 2014; ver também Matthews-Saugstad et al., 2016).

Levamos em conta em nosso estudo ambas taxas de gesticulação por *tempo de fala* quanto por *tempo total* da tarefa, pois observamos gesticulação manual pelos participantes durante a produção da fala, em sua maior parte, mas também, por raras vezes, em silêncio durante a elaboração linguístico-discursiva. O fenômeno de gesticulação em silêncio, apesar de raro dentre as produções analisadas em nosso corpus, chamou-nos especial atenção por ilustrarem proposições de trabalhos como o de Kendon (1980) e de Butterworth & Beattie (1978) quando colocam que, apesar dos processos originários da fala e da gesticulação se originarem a um único e mesmo tempo, por a representação de ideias em movimentos gesticulatórios ser mais primitiva do que o tratamento hierárquico de unidades analíticas pela fala, frases gesticulatórias (sobretudo sua preparação) estariam disponíveis e emergiriam mais cedo do que o segmento da fala com seu afiliado lexical no instante da elocução. De fato, gestos foram produzidos em silêncio por participantes com baixo nível de proficiência na língua; com alta dificuldade de construção narrativa na língua estrangeira, tais participantes claramente recorreram aos gestos em ensaios mentais de elaboração discursiva (consultar Kita et al., 2017, para aprofundamento da discussão sobre os chamados *co-thought gestures*).

Uma análise de regressão linear múltipla que considerou a totalidade das produções narrativas de nossa população amostral, ao mostrar os parâmetros linguísticos de número de *repetições* ($\beta = 0,470$; $t = 3,677$; $p = 0,001$) e de número de *hesitações* ($\beta = 0,260$; $t = 2,038$; $p = 0,047$) como fatores preditores do *tempo total* de produção das

narrativas ao considerarmos o conjunto da totalidade das produções *com e sem gestos* [$F(2,51) = 18,749$; $p = 0,001$; $R^2 = 0,424$], sugere que a variabilidade entre os sujeitos quanto ao tempo gasto na realização da tarefa se explique em parte pela maior ou menor presença em suas produções orais de elementos discursivos ruidosos, que pouco acrescentam à qualidade narrativa. Por outro lado, ainda considerando nossa população amostral em conjunto, a análise de regressão linear múltipla que apontou como preditores do *tempo de fala* [$F(2,49) = 33,009$; $p < 0,001$; $R^2 = 0,729$] o número de *autocorreções* ($\beta = 0,398$; $t = 4,525$; $p < 0,001$), o número de *orações* ($\beta = 0,369$; $t = 4,281$; $p < 0,001$), o número de *repetições* ($\beta = 0,231$; $t = 2,570$; $p = 0,013$), e o número de *substantivos diferentes* ($\beta = 0,201$; $t = 2,252$; $p = 0,029$) sugere explicação de natureza mais qualitativo-informativa para a variabilidade do *tempo de fala* nas produções narrativas. Em vista disso, buscamos o quanto nos foi possível explorarmos nossos dados tanto quantitativa quanto qualitativamente.

Entre as condições experimentais

Entendendo a gesticulação manual como um recurso cognitivo recrutado pelo sistema linguístico durante a elaboração e produção da fala, esperávamos como resultado de nosso controle experimental, ou seja, uma vez inibida a gesticulação manual e os sujeitos participantes privados de seus recursos, que produzissem narrativas mais longas, com maior número de erros gramaticais, de hesitações, maior número e tempo de pausas em decorrência de maiores dificuldades linguístico-discursivas na condição experimental *sem gestos*. Surpreendentemente, a produção oral dos sujeitos na condição experimental *sem gestos* foi realizada em menor tempo: o *tempo total* de realização da tarefa na condição *com gestos* superou em 38% na média o tempo de produção na condição *sem gestos* (Related-Samples Wilcoxon Signed Rank Test; $Z = -3,291$; $p = 0,001$); e o *tempo de fala*, em 23% ($Z = -2,972$; $p = 0,003$). Tais resultados nos sugeriram de imediato duas possibilidades de interpretação e encaminhamento das análises: a de que a condição *sem gestos* teria, de forma inusitada, facilitado a produção narrativa; ou, em uma direção contrária, a de que nosso controle experimental pudesse ter induzido a um aumento da carga de estresse e, uma vez em condição de ansiedade, os participantes teriam produzido as narrativas de forma mais apressada, com menor engajamento e menor cuidado, de forma menos informativa e com mobilização de menos recursos cognitivos.

Interessantemente, Rauscher, Krauss & Chen (1996) também reportaram resultado inusitado com produções narrativas mais rápidas (maior número de palavras por minuto) em condição de restrição de gesticulação quando se tratando de conteúdo não-espacial. Seus resultados, porém, permaneceram inexplicados no trabalho e inexplorados na literatura como um todo. Discutiremos um possível fator de estresse na condição *sem gestos* mais adiante nesta seção.

Apontando na mesma direção de nosso primeiro resultado, todos os parâmetros que foram analisados (à exceção do número de orações) registraram maiores valores para a condição experimental *com gestos*, corroborando assim nosso primeiro achado de que as produções diferiram entre as duas condições experimentais (**Gráfico 5**). Compreendemos, porém, que, uma vez verificada a produção na condição *com gestos* ter sido tão mais longa, seria esperado que ela fosse acompanhada de maior volume de ocorrências linguísticas. Diante disso, afim de explorarmos mais amiudamente a natureza da diferença encontrada entre as duas condições experimentais, investigando se ela pudesse estar relacionada a mudanças no tempo de processamento linguístico ou a mudanças na qualidade informativa das narrativas, analisamos os parâmetros linguísticos considerando então sua *taxa* de frequência, ou seja, o número de ocorrências por minuto de *tempo de fala*.

Ao passarmos à análise das *taxas* de ocorrência dos parâmetros gramaticais por *tempo de fala*, ainda esperávamos encontrar maiores *taxas de erros* e de ruídos discursivos, como *repetições* e *autocorreções*, na condição experimental *sem gestos*. A uma primeira vista, os resultados pareceram-nos confundidores, pois apresentaram tendências não-homogêneas entre as condições experimentais, ou seja, alguns parâmetros tenderam a um maior valor médio para uma condição, outros, para a outra condição experimental; a *taxa de erros*, por exemplo, foi na média entre os participantes ligeiramente superior na condição *sem gestos* ($19,9, \pm 1,7$, para a condição *com gestos* e $21,7, \pm 2,8$, para a condição *sem gestos*; $p > 0,05$), enquanto que a *taxa de autocorreções*, maior na condição *com gestos* ($5,0, \pm 0,9$ para a condição *com gestos* e $3,2, \pm 0,6$ para a condição *sem gestos*; $p > 0,05$), sem contudo apresentarem diferença estatisticamente significativa. Também não encontramos diferença entre as condições experimentais quanto às *taxas de hesitações* ($20,7, \pm 1,8$, para a condição *com gestos* e $20,1, \pm 1,6$, para a condição *sem gestos*; $p > 0,05$) e de *interrupções sintáticas* ($52,8, \pm 2,5$, para a condição *com gestos* e $54,5, \pm 2,3$, para a condição *sem gestos*; $p > 0,05$). Não encontramos, portanto, em nossos resultados disfluência na fala por impedimento à gesticulação, como o reportado por Morsella & Krauss (2004), muito embora observemos pequena tendência a uma maior taxa de

ocorrência de *interrupções sintáticas* na condição *sem gestos*. Encontramos diferenças significativas, entretanto, entre as condições experimentais quanto a: *taxa de repetições*, com valor superior na condição *com gestos* ($9,7, \pm 1,6$ para a condição *com gestos* e $6,0, \pm 1,2$, para a condição *sem gestos*; $p < 0,05$), e *taxa de orações*, superior na condição *sem gestos* ($14,4, \pm 1,3$, para a condição *com gestos* e $18,4, \pm 1,3$, para a condição *sem gestos*; $p < 0,01$; **Tabela 3**).

A partir da análise das taxas de ocorrências, portanto e diferentemente do que esperávamos, deparamo-nos na condição *com gestos* com uma produção linguística quantitativamente menos informativa (ou menos econômica), por apresentar menor *taxa de orações*, bem como mais redundante e expletiva por uma maior *taxa de repetições* em comparação com a produção na condição *sem gestos*. Tais resultados parecem, assim, apontar para a mesma direção de nossas análises regressivas preliminares (da totalidade das produções narrativas; $n = 54$), sugerindo, não uma variação do grau de dificuldade entre as condições experimentais relacionada, por exemplo, a uma maior ou menor ocorrência de erros gramaticais ou autocorreções (essa era nossa hipótese inicial), mas antes de elementos discursivos aparentemente ruidosos, na medida em que apresentam redundância entre si, como no caso das repetições, não necessariamente contribuintes para a qualidade informativa das narrativas, influenciando no tempo de produção das tarefas. Referimo-nos às repetições como elementos *aparentemente* ruidosos pois, se, a uma primeira análise sua redundância discursiva implica em menor grau informativo e maior custo linguístico (uma menor economia linguístico-discursiva, portanto), quando conjugados com a gesticulação revelam-se efetivos no aumento da compreensão das narrativas por sujeitos interlocutores; dito em outras palavras, o custo linguístico-discursivo promovido pela produção de gestos manuais em acompanhamento à fala no ato de elocução é socialmente compensado no ato comunicativo por uma melhora efetiva na compreensão das narrativas por sujeitos ouvintes expectadores (ver Pereira & Haddad, 2005).

Curiosamente, observamos que na média, a *taxa de fala* (*tempo de fala* dividido pelo *tempo total* da tarefa) permaneceu praticamente constante entre as duas condições experimentais (55,2% na condição *com gestos* e 55,8% na condição *sem gestos*; $p > 0,5$; **Gráfico 8**); a mesmo resultado chegaríamos pela análise do dado complementar ao considerarmos a taxa de silêncio, ou seja, o tempo de silêncio dividido pelo tempo total de realização da tarefa (44,8% na condição *CG* e 44,2% na condição *SG*). A constância da *taxa de fala* (ou taxa de silêncio) nos sugere que a proporção de tempo gasto durante a tarefa para elaboração discursiva permaneceu constante entre as condições experimentais;

em outras palavras, nosso controle experimental não interferiu no tempo de planejamento discursivo. Corrobora a interpretação de que nosso controle experimental não tenha interferido no planejamento linguístico o resultado obtido pela análise de outras duas taxas: a *taxa de hesitações*, compreendidas aqui como enunciações, vocalizações ou prolongamentos silábicos sem qualquer contribuição informacional para o desenvolvimento narrativo, mas estreitamente relacionadas com o tempo de acesso lexical e elaboração discursiva; e a *taxa de interrupções sintáticas*, também esta diretamente relacionada ao acesso lexical e planejamento discursivo (como *taxa de hesitações* obtivemos média de 20,72 [$\pm$ 1,8] para a condição *com gestos* e 20,10 [$\pm$ 1,6] para a condição *sem gestos* [$p > 0,05$]; como *taxa de interrupções sintáticas* obtivemos média de 52,8 [$\pm$ 2,5] para a condição *com gestos* e 54,5 [$\pm$ 2,3] para a condição *sem gestos* [$p > 0,05$]; **Tabela 3 e Gráfico 11**).

Nossos resultados quantitativos nos remeteram, ademais, à análise dos vídeos das produções narrativas dos sujeitos experimentais. A análise qualitativa das produções orais sugerira por um lado maior assertividade informativa na condição experimental *sem gestos*, corroborada por resultados quantitativos como maior *taxa de orações* (**Gráfico 6**), medida indicativa do grau informativo da produção (quantidade de informação por tempo de fala), e maior *variabilidade de substantivos*, medida indicativa do grau de variabilidade substantiva de informação (**Gráfico 7**); por outro lado, a análise qualitativa sugerira maior prolixidade com certa redundância informativa na condição *com gestos*, corroborada pelos dados quantitativos de maior *tempo total* na produção da tarefa (39% maior do que na condição *sem gestos*; $p < 0,01$), e de maior *tempo de fala* (23% maior do que na condição *sem gestos*; $p < 0,01$; **Tabela 3**), mas também, e sobretudo, por maior *taxa de repetições* ($p < 0,05$). Um estudo experimental com mesmo desenho que o nosso, porém com qualidade de áudio controlada, poderia, em um momento seguinte, colocar à prova de julgamento qualitativo (por exemplo, de assertividade ou qualidade narrativa) os registros de áudio das produções narrativas por avaliadores cegos para as condições experimentais.

Em suma, o conjunto de nossos dados mostraram, portanto: i) que o fluxo de fala (calculado pela *taxa de fala*) se manteve entre as condições experimentais, sugerindo que a restrição da gesticulação por nosso controle experimental não interferiu no custo temporal do planejamento discursivo em curso durante a tarefa; ii) que a produção oral narrativa na condição *sem gestos* é menos extensa, bem como, por apresentar menor taxa de repetição e maior taxa oracional, menos ruidosa e proporcionalmente mais informativa; iii) que a produção narrativa na condição *com gestos*, por outro lado, pela maior

ocorrência de *hesitações* e de *interrupções sintáticas*, é uma produção menos econômica, mais expletiva, por apresentar maior grau de redundância observado pela *taxa de repetições*.

Como podemos observar, nossos resultados parecem nos conduzir na contra-mão de resultados reportados na literatura, que sugerem gestos manuais melhorarem a produção linguística oral por aumentarem-lhe a fluência (Graham & Heywood, 1975; Rauscher, Krauss & Chen, 1996), diminuírem-lhe as pausas preenchidas (Krauss, 1998; Rauscher et al., 1996), ou ainda por aliviarem carga cognitiva (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Ping & Goldin-Meadow, 2010; Wagner, Nusbaum, & Goldin-Meadow, 2004). É importante observarmos, no entanto, que os trabalhos logo acima mencionados, bem como outros que verificaram melhora na produção linguística oral pela presença de gestos manuais, utilizaram quase todos eles *conteúdo espacial* (quando não tarefa de resolução de problemas) como estímulo à produção linguística oral em seu desenho experimental, observando limites ou ponderações de seus resultados para passagens com conteúdo não-espacial (ver, por exemplo, Beilock & Goldin-Meadow, 2010; Morsella & Krauss, 2004; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Rauscher, Krauss & Chen, 1996). Ademais, alguns deles utilizaram para controle experimental dos gestos uma segunda tarefa: explicitamente, instruíram os sujeitos experimentais a não gesticularem, o que, como já discutimos, certamente lhes aumentara a carga cognitiva sobrecarregando a memória operacional no cumprimento da tarefa na condição de restrição dos gestos. Também diferem de nosso estudo experimental por terem analisado produções orais em língua materna dos sujeitos participantes; conforme mostramos no capítulo introdutório na seção sobre aumento da carga cognitiva e disfluência em gesto e fala, a gesticulação manual pode se comportar distintamente entre produções narrativas em língua materna e em língua estrangeira por um mesmo indivíduo.

Mas voltando a nossos resultados e avançando na direção por eles apontada, entendemos que, se por um lado, apontam para uma menor qualidade das produções narrativas quando na presença de gestos, por outro, todavia, tais produções tendem a ser acompanhadas por maior uso de elementos metalinguísticos, como as *autocorreções*, que, aliás, em si também poderiam ser compreendidas como ruídos disruptores do fluxo narrativo e, portanto, como custo linguístico; *autocorreções*, no entanto, são também fortes evidências de alto engajamento do sujeito no cumprimento da tarefa, através do emprego de recursos metacognitivos de viés correccional.

A análise de regressão linear múltipla que resultou em um modelo significativo [$F(2,24) = 9,343$; $p = 0,001$; $R^2 = 0,438$] com os parâmetros linguísticos de número de *repetições* ($\beta = 0,397$; $t = 2,319$; $p = 0,029$) e de *autocorreções* ($\beta = 0,380$; $t = 2,221$; $p = 0,036$) como preditores do *tempo de fala* na condição *com gestos* aponta para uma interpretação de que as produções narrativas acompanhadas de gesticulação manual variaram sobretudo no número de repetições lexicais e sintáticas e de reformulações linguísticas com caráter correccional. A variabilidade entre os participantes quanto ao número de *autocorreções* nos sugere diferentes graus de percepção e consciência de sua própria produção, diferentes graus de uma consciência metalinguística do processo de elaboração e enunciação narrativa; ao mesmo tempo que a variabilidade entre os sujeitos quanto ao número de *repetições* nos sugere diferentes graus de ruídos pouco contrubuintes para a melhora da qualidade da produção narrativa. De toda forma, a variabilidade observada tanto do número de *repetições* quanto do número de *autocorreções* na condição *com gestos* aponta para diferenças entre os sujeitos quanto ao seu esforço de maior ou menor engajamento no cumprimento da tarefa experimental.

Com objetivo de explicarmos a variância entre os sujeitos quanto à diferença encontrada entre suas duas produções narrativas *com* e *sem gestos*, realizamos também outras duas análises de regressão linear múltipla, uma para o *tempo total* da tarefa e outra para o *tempo de fala*, considerando tanto para a variável de previsão quanto para as variáveis preditoras a *diferença* entre as condições experimentais (*condição com gestos* menos a condição *sem gestos*). O resultado destas análises continuam a apontar em mesma direção de nossas análises precedentes, com modelo estatisticamente significativo [$F(2,23) = 27,733$; $p < 0,001$; $R^2 = 0,644$] com os fatores *hesitações* ($\beta = 0,595$; $t = 4,428$; $p < 0,001$) e *repetições* ($\beta = 0,343$; $t = 2,550$; $p = 0,018$) como previsores do *tempo total* entre as narrativas, sugerindo, portanto, maior ou menor presença de elementos ruidosos ou de redundância como explicação da variabilidade entre os indivíduos da diferença do *tempo total* das produções entre as duas condições experimentais, e um modelo estatisticamente significativo [$F(2,23) = 25,427$; $p < 0,001$; $R^2 = 0,768$] com número de *orações* ($\beta = 0,528$; $t = 4,504$; $p < 0,001$), de *repetições* ($\beta = 0,303$; $t = 2,756$; $p < 0,011$) e de *substantivos totais* ($\beta = 0,301$; $t = 2,410$; $p = 0,024$) como previsores da diferença (*CG – SG*) de *tempo de fala* entre as narrativas, sugerindo oscilações com maior tendência substancial (mas também de redundância) como explicação da variabilidade entre os indivíduos da diferença do *tempo de fala* das produções entre as duas condições experimentais.

O custo linguístico, observado em nossos resultados, do agenciamento de todo aparato cognitivo disponível, inclusos planejamento e realização motora de conteúdo imagético em gestos, no esforço de elaboração do espaço cognitivo a ser compartilhado com o interlocutor, portanto, parece ser o ônus de uma produção altamente engajada e mais consciente do próprio ato da produção. Tal custo seria, contudo, compensado pelo benefício à compreensão pelo ouvinte-expectador, quando a este seja dado acesso visual aos gestos do narrador (ver Pereira & Haddad, 2015, e Beattie & Shovelton, 1999). Por uma perspectiva evolutiva, com efeito, não faria sentido o aumento (ou manutenção) em nossa espécie do custo linguístico-cognitivo por acréscimo de componentes visuo-motores de movimentos corporais em acompanhamento à fala, não fossem os benefícios cognitivos sociais tais o engajamento atencional do interlocutor e melhora na comunicação, já reportados na literatura. Poderíamos argumentar, no entanto, que é possível se observar a gesticulação manual presente em produções linguísticas com interlocução à distância e sem contato visual, como ao telefone, por exemplo, ou mesmo em produções sem interlocução pessoal (“pensando em voz alta”, como dizemos) e ainda no escuro. Mas, ainda que desejássemos dispensar toda nossa gesticulação corporal durante o ato de elocução em uma produção oral individual isolada (de fato, conseguimos algum controle e mesmo a contenção de nossos gestos), a linguagem continuaria sendo, não obstante e inexoravelmente, um espaço e morada cognitiva social, pressupondo, portanto, a interlocução.

EXPERIMENTO 2

OBJETIVO

O objetivo do estudo experimental 2 foi investigar o impacto da gesticulação pela professora na alocação da atenção visual e na aprendizagem de alunos em uma aula de ciências. Para tanto, controlamos a visualização pelos sujeitos participantes dos gestos produzidos pela professora durante a aula. Registramos a fixação ocular dos participantes e lhes aplicamos um questionário de avaliação de aprendizagem.

Participantes

Participaram voluntariamente de nosso estudo experimental 62 sujeitos (29 mulheres; média idade 24,4; $\pm 5,2$), estudantes de diferentes cursos de graduação e pós-graduação da Universidade de São Paulo; todas as informações pessoais dadas pelos participantes para a realização dos experimentos foram mantidas em sigilo e confidencialidade. Os participantes não tinham informação prévia dos objetivos específicos e das medidas avaliadas no experimento. Os procedimentos metodológicos do experimento foram submetidos e aprovados pelo Comitê de Ética em Pesquisa com Seres Humanos do Instituto de Psicologia da USP (CAAE 19335518.8.0000.5561, Parecer 3.699.206; **Anexo 2**).

Materiais e métodos

Previamente ao experimento, preparamos um vídeo para ser apresentado como estímulo aos participantes. O vídeo consistiu em uma aula sobre etologia, de aproximadamente 15 minutos, com exibição de slides na porção central da tela e exibição da professora em um dos cantos inferiores (à esquerda ou à direita, aleatoriamente entre os participantes; para visualização da sequência dos slides, ver **Anexo 3**). A aula apresenta

definição, classificação e exemplos de comportamento animal (Lorenz, 1981; Alcock, 2013).

O vídeo foi editado de maneira que obtivéssemos 4 vídeos distintos: todos com a sequência de slides e exibição da professora palestrante, sendo 2 vídeos com enquadramento total da parte superior do corpo da professora (condição *com gestos*; **Figuras 20a e b**) e 2 vídeos com enquadramento apenas do ombro e cabeça da professora (condição *sem gestos*; **Figuras 23c e d**). Os sujeitos experimentais assistiram a um dos quatro vídeos, aleatorizados e balanceados entre os participantes.

Figura 20. Possíveis enquadramentos utilizados de forma aleatória entre os participantes.

Optamos pelo enquadramento da professora em separado do enquadramento dos slides como uma solução técnica simples com vistas ao nosso controle experimental: garantimos, assim, aos participantes a apresentação de mesmo conteúdo de aula com exatamente a mesma produção linguístico-discursiva da professora palestrante. Ademais, a filmagem da professora de forma isolada da projeção eliminou sua interação física com os slides, de modo que pudemos controlar interferências de outras modalidades

semióticas atuantes em uma aula regular mas que não são objeto de investigação do presente estudo, como, por exemplo, a proxemia (relação da posição da professora palestrante em relação ao objeto de ensino e ao público expectador; ver Giordan et al., 2015); também inibimos, assim, a produção de gestos dêiticos (de apontamento) com referência ao quadro de projeção, gestos estes que poderiam, em uma aula convencional, direcionar a atenção dos participantes. Procuramos favorecer, portanto, a produção de gestos próprios ao esforço explanatório do conteúdo apresentado pela professora.

Procedimento experimental

Os participantes foram recebidos individualmente e conforme agendamento em nosso laboratório e instruídos a assistirem atentamente aos 15 minutos de aula apresentada em projeção e a responderem, logo a seguir, a questões de compreensão sobre o conteúdo da aula (**Anexo 4**).

Durante a projeção da aula, registramos as fixações oculares dos participantes com equipamento e metodologia de rastreamento ocular. Pela análise das fixações oculares, esperamos identificar quais gestos (se de pulso ou icônicos, em especial) demandam maior número e tempo de fixação, ou se, ao contrário, dispensam direcionamento explícito de recursos atencionais visuais.

Após a exibição da aula, aplicamos um questionário de compreensão do conteúdo apresentado com questões objetivas e dissertativas. Os participantes tiveram tempo livre para resposta. Com o questionário, procuramos avaliar na aprendizagem do conteúdo da aula eventual prejuízo pela privação da visualização da gesticulação manual da professora palestrante pelos sujeitos participantes. Procuramos ainda identificar se diferentes estratégias de fixação ocular resultaram em diferenças na aprendizagem.

Para as análises estatísticas, as premissas de normalidade de distribuição de dados e homogeneidade de variância foram verificadas pelos testes de Kolmogorov-Smirnov e Levine, respectivamente. Adotamos como critério de significância estatística 5% ($p \leq 0,05$). Todas as análises foram conduzidas em IBM SPSS Statistics versão 20.0 (IBM Corp., 2011).

ANÁLISE

A fala e os gestos produzidos pela professora durante a apresentação da aula foram analisados por software ELAN versão 5.0.0-beta (**Figura 21**), desenvolvido por Max Plank Institute for Psycholinguistics (<http://tla.mpi.nl/tools/tla-tools/elan/>; ver Sloetjes & Wittenburg, 2008, para anotações).

Figura 21. Ilustração da análise da fala e da gesticulação da professora durante a aula de ciências.

Analisamos da fala da professora o seu tempo total de produção, a ocorrência de pausas sintáticas, o número de períodos discursivos (marcados por entonação conclusiva), bem como a frequência de acentuações tônicas silábicas marcando o ritmo prosódico da fala.

Com a finalidade de analisarmos o papel da gesticulação em passagens da aula com informações de naturezas distintas, identificamos e analisamos separadamente 5 etapas da sequência didática (**Tabela 4**): parte introdutória à aula (1:25 minutos); primeira sequência com definições e exemplos ilustrados de comportamentos inatos (5 minutos); segunda sequência com definições e exemplos ilustrados de comportamentos adquiridos (4 minutos); esquema ilustrado sobre o condicionamento clássico de Ivan Pavlov (1849-1936; 2 minutos); esquema gráfico sobre condicionamento operante (2:30 minutos); última etapa com as referências bibliográficas da aula (0:17 minutos).

Introdução	Sequência 1	Sequência 2	Esquema	Gráfico	Referências
1:25	5:00	4:00	2:00	2:30	0:17
4 slides	7 slides	5 slides	1 slide	1 slide	1 slide

Tabela 4. Etapas da sequência didática da aula sobre etologia.

Analizamos então, para cada slide de cada uma das etapas da sequência didática, o número de ocorrências e a frequência por tempo de fala dos gestos ilustradores (de pulso, dêiticos, metafóricos e icônicos) produzidos pela professora.

Também analisamos para cada um dos slides e etapa didática o tempo médio de fixação ocular dos participantes em três áreas de interesse durante a exibição da aula: a região dos gestos, a região da professora e a região do slide. A análise das fixações oculares foi realizada por software Tobii Studio 3.2.2 (Tobii Technology, Danderyd, Sweden).

Observamos, entretanto, como perda de informação de rastreamento ocular na média entre os participantes 38,3% do tempo total de aula. Em razão dessa perda e de nosso interesse em compreendermos as estratégias de alocação atencional empregadas pelos sujeitos aprendizes, optamos por considerar em nossas análises, não os valores registrados em unidade de segundos, mas a porcentagem de tempo dedicada por cada indivíduo a cada uma das três regiões analisadas: a região da professora, do slide e dos gestos. Excluímos de nossa análise participantes cujo tempo total de fixação ocular registrado tenha sido inferior a 15% do tempo total da aula ($n = 5$).

Para averiguarmos possíveis mudanças no comportamento ocular dos participantes dependentes da variável tempo de exibição do slide, analisamos tanto a porcentagem de tempo de fixação ocular para cada uma das três regiões estudadas quanto o tempo médio da fixação em cada uma delas levando-se em consideração a 1ª e a 2ª metades do tempo de exibição de cada slide.

Por fim, para possível benefício ou prejuízo na aprendizagem dos participantes por efeito de nosso controle experimental de exibição dos gestos da professora (condição *com gestos* e *sem gestos*), analisamos a pontuação média dos participantes no questionário de compreensão. Para fins de pontuações, atribuímos valor de 1 a 3 conforme o grau de correção (1 para respostas erradas ou não respondidas, 2 para respostas parcialmente corretas, e 3 para respostas totalmente corretas), tanto às questões objetivas quanto às

questões dissertativas. Como critério de correção das questões dissertativas, consideramos a coerência interna das respostas bem como coerência com o conteúdo apresentado em aula.

RESULTADOS

FALA

A aula foi composta por uma sequência de 19 slides (**Anexo 3**) com duração média de 48,3 segundos ($\pm 37,1$). A professora falou (vocalizou) 83% do tempo de aula, o restante do tempo sendo em sua maior parte silêncio durante transições de slides, e, em menor proporção, silêncios de pausas sintáticas (**Gráfico 12**).

Gráfico 12. *Duração dos slides e tempo de fala da professora (em segundos).*

Durante a fala da professora, computamos 1.532 acentuações tônicas frasais, marcando seu ritmo prosódico com frequência média de 2,0 acentos por segundo de fala. Frequências mais altas foram observadas para os slides iniciais e terminais da aula, e frequências mais baixas, concentradas nos slides intermediários (**Gráfico 13**).

Gráfico 13. Número de ocorrências de *acentuações tônicas frasais* por slide e por segundo de fala, para cada slide da sequência didática da aula.

Ainda analisando a composição temporal discursiva, computamos ao longo da aula 109 entonações conclusivas indicativas de finalizações de períodos. Verificamos uma média de 6,9 ($\pm 1,95$) segundos por período, destacando períodos mais longos para os slides 3 e 19 e períodos mais curtos para os slides 1, 5 e 12 (**Gráfico 14a**). Entre as sequências didáticas, observamos grandes variações no número total de períodos por etapa: 7 para a introdução, 40 para a sequência 1, 28 para a sequência 2, 18 para o esquema, 15 para o gráfico e 1 para as referências (**Gráfico 14b**).

Gráfico 14a. Número de ocorrências e duração média (segundos) de *períodos* por slide.

Gráfico 14b. Frequência de ocorrência de *períodos* e sua duração média (segundos) por sequência didática.

GESTICULAÇÃO

A professora gesticulou 78% do tempo total da aula e 94% do tempo de fala (Gráfico 15).

Gráfico 15. Tempo de gesticulação, de fala e duração dos slides.

Computamos um total de 597 golpes gesticulatórios (0,78 gesto por segundo de fala), todos caracterizados como gestos de pulso por acompanharem a acentuação

prosódica da fala. Dos 597, 151 deles compuseram marcações dêiticas discursivas (gestos dêiticos), 132 gestos metafóricos e 82 gestos icônicos, com *frequência de ocorrência* por segundo de fala de 0,79, 0,20, 0,18 e 0,11, respectivamente (**Gráfico 16**).

Gráfico 16. Frequência e ocorrência de gestos ilustradores por segundo de fala.

Pela análise da sequência de slides, observamos concentrações gesticulatórias acompanhando o tempo de duração dos slides (**Gráfico 17**). Para além da presença de base da gesticulação de pulso em toda a aula, as sequências didáticas se destacaram a grosso modo da seguinte forma: introdutória, por alta ocorrência de gestos dêiticos; sequência 1, por uma crescente de gestos metafóricos e decrescente de gestos dêiticos, e por alto volume de gestos icônicos; sequência 2, com baixa frequência de gestos icônicos, crescente de gestos metafóricos e dêiticos (estes últimos, de forma mais acentuada); sequência esquemática e gráfica, com alta frequência de todas as categorias de gestos, com pico de gesticulação icônica na sequência com o gráfico; por fim, as referências bibliográficas, com baixa ocorrência de todas as categorias.

Gráfico 17. Número de ocorrências de gestos ilustradores por slide.

Analisando-se a frequência de ocorrência dos gestos por segundo de fala ao longo da sequência de slides (**Gráfico 18**), observamos: uma frequência alta de gestos de pulso ao longo de toda a sequência de slides com picos nos slides 4, 9 e 14, e vales nos slides 2, 12 e 19; frequência regular de gesticulação dêitica, com picos nos slides 3 e 15 e decaimentos lentos subsequentes; oscilações importantes na gesticulação metafórica, com picos mais elevados nos slides 5 e 11 e vales nos slides 2, 6, 12 e 19; baixa frequência de gesticulação icônica com agrupamentos de frequências para os slides 3 e 4, 7 ao 10, e 17 e 18. De maneira geral, observamos grandes conjuntos de gesticulação delimitados entre fortes baixas (nula, em alguns casos) da frequência de ocorrência de gestos, a saber, os slides 2, 6, 12 e 19.

Gráfico 18. Frequência de ocorrências de gestos ilustradores por segundo de fala.

Também analisamos a frequência dos gestos para cada etapa da sequência didática (**Gráfico 19**). A análise mostrou que a frequência da gesticulação de pulso se manteve alta (0,79 por segundo de fala) e com pequena variação ao longo de toda sequência didática, marcando com consistência o ritmo da fala, diminuindo para 0,56 apenas no último slide com as referências bibliográficas. Gestos dêiticos iniciaram a parte introdutória da aula com frequência de 0,28, reduzindo para 0,14 na sequência 1 e mantendo-se a 0,24 para a sequência 2 e parte esquemática, decaindo lentamente ao final para 0,19 na etapa do gráfico, e terminando com 0,16 para a etapa com as referências bibliográficas. A gesticulação metafórica iniciou-se com frequência moderada, apresentando leve crescente ao início da sequência didática, com 0,13 na introdução e 0,18 na sequência 1, mantendo uma regularidade em torno desse último valor para as demais etapas didáticas, decaindo bruscamente na última etapa com as referências bibliográficas. A menor frequência de gestos foi observada para os gestos icônicos, que também apresentaram as maiores oscilações de frequência dentre os gestos, com maiores valores registrados para a sequência 1 e a sequência gráfica (0,22; **Gráfico 19**).

Gráfico 19. Frequência de ocorrências de gestos ilustradores por segundo de aula.

Uma análise correlacional de Spearman mostrou correlação positiva moderada entre a frequência de ocorrência de gestos icônicos e o tempo de duração dos slides ($\rho = 0,698$; $p = 0,001$; **Tabela 5**).

Correlation Spearman's ρ		Frq. Acentos Tonicos	Duração SLIDE	Dur. Med. Períodos	Frq. G. PULSO	Frq. G. DÊIT.	Frq. G. METAF.	Frq. G. ICÔN.
Fr. Acentos Tonicos	Correlation Coefficient	1,000	-,111	-,056	,301	,089	,187	-,068
	Sig. (2-tailed)	.	,652	,819	,210	,718	,443	,783
	N	19	19	19	19	19	19	19
Dur. SLIDE	Correlation Coefficient	-,111	1,000	,235	,366	,261	,407	,698**
	Sig. (2-tailed)	,652	.	,332	,123	,281	,083	,001
	N	19	19	19	19	19	19	19
Dur. Med Períodos	Correlation Coefficient	-,056	,235	1,000	-,014	,191	,132	,234
	Sig. (2-tailed)	,819	,332	.	,956	,434	,589	,336
	N	19	19	19	19	19	19	19
Frq. Gestos De PULSO	Correlation Coefficient	,301	,366	-,014	1,000	,032	,516*	,406
	Sig. (2-tailed)	,210	,123	,956	.	,896	,024	,085
	N	19	19	19	19	19	19	19
Frq. Gestos DÊITICOS	Correlation Coefficient	,089	,261	,191	,032	1,000	,072	,121
	Sig. (2-tailed)	,718	,281	,434	,896	.	,770	,623
	N	19	19	19	19	19	19	19
Frq. Gestos METAFÓRICOS	Correlation Coefficient	,187	,407	,132	,516*	,072	1,000	,164
	Sig. (2-tailed)	,443	,083	,589	,024	,770	.	,502
	N	19	19	19	19	19	19	19

Frq. Gestos	Correlation Coefficient	-,068	,698**	,234	,406	,121	,164	1,000
ICÔNICOS	Sig. (2-tailed)	,783	,001	,336	,085	,623	,502	.
	N	19	19	19	19	19	19	19

Tabela 5. Correlação (Spearman) entre a frequência de parâmetros linguístico-discursivos (de períodos e de acentuações tônicas frasais), de gestos ilustradores (de pulso, dêiticos, metafóricos e icônicos) e a duração dos slides (n = 19).

Por fim, apresentamos a seguir dois gráficos como tentativas de ilustrar a composição semiótica gestual-discursiva da aula: o primeiro (**Gráfico 20**) reunindo as frequências de ocorrências de gestos ilustradores (de pulso, dêiticos, metafóricos e icônicos) e de acentuações tônicas frasais por segundo de fala, bem como a duração média do período para cada slide; o segundo (**Gráfico 21**), com os mesmos parâmetros porém apresentados em número absoluto de ocorrências.

Gráfico 20. Frequência de ocorrência de gestos ilustradores e de acentuações tônicas frasais (por segundo de fala), e duração média do período de fala (em segundos).

Gráfico 21. Representação topográfica da composição semântica gesticulatório-discursiva de gestos ilustradores, de acentuações tônicas frasais, e de períodos de fala.

FIXAÇÕES OCULARES

Dos 917,9 segundos totais de tempo de aula, computamos média entre os 57 sujeitos participantes de 565,9 segundos ($\pm 23,5$) de registro total de fixação ocular; uma perda média de informação de 38,3% ($\pm 0,03$) do tempo total de aula, portanto. Dos 565,9 segundos ($\pm 23,5$) de tempo total médio registrado, 269,9 ($\pm 18,34$) segundos foram registrados na região da professora e 296,0 ($\pm 14,5$), na região do slide. Na média entre as porcentagens individuais dos participantes, os sujeitos dedicaram 46,5% ($\pm 2,0\%$) de seu tempo fixando a região da professora e 53,5% ($\pm 2,0\%$) fixando a região do slide (**Gráfico 22**).

Gráfico 22. Tempo de fixação ocular nas regiões da professora e do slide ($\pm$ erro padrão; $n = 57$).

Conduzimos análises de variância de uma via para verificarmos possíveis efeitos das variáveis não-controladas referentes à caracterização do perfil do público participante. Não encontramos efeito significativo na porcentagem de tempo de fixação na região da professora, do slide ou dos gestos pelas variáveis de sexo, idade, formação anterior e atual, ou conhecimento prévio do assunto. Uma ANOVA de uma via, contudo, evidenciou efeito do horário de realização do experimento ($F_{(2,50)} = 4,649$; $p = 0,014$), a porcentagem de tempo de fixação na região da professora para os sujeitos que realizaram a sessão experimental pela manhã ($57,5\%$; $\pm 0,04$; $n = 10$) distinguindo-se de ambas porcentagens dos sujeitos que realizaram a sessão experimental pela tarde ($42,4\%$; $\pm 0,02$; $n = 33$) e pela noite ($45,6\%$; $\pm 0,05$; $n = 10$), estas últimas não se distinguindo entre si (Gráfico 23).

Gráfico 23. Média do tempo de fixação ocular na região da professora ($\pm$ erro padrão; $n = 57$) das sessões experimentais realizadas nos períodos da manhã, tarde e noite.

As fixações oculares tiveram como duração média 0,49 segundos ($\pm 0,02$) para a região da professora, 0,23 segundos ($\pm 0,01$) para a região do slide, e 0,29 segundos ($\pm 0,01$) para a região dos gestos (este ultimo, apenas na condição *com gestos*). Uma análise de variância mostrou diferença significativa entre a duração média das fixações nas áreas da professora e do slide ($n = 57$; ANOVA; $F_{(1,112)} = 114,959$; $p < 0,001$; **Gráfico 24**).

Gráfico 24. Tempo médio da fixação ocular nas regiões da professora, do slide e dos gestos em segundos ($\pm$ erro padrão; $n = 57$).

Com interesse sobre a duração média da fixação ocular na região dos gestos, também verificamos por uma análise de variância diferença significativa entre a duração

média das fixações oculares nas áreas da professora, dos gestos e do slide para a condição *com gestos* ($n = 28$; ANOVA; $F(2,81) = 27,025$; $p < 0,001$). Testes post-hocs mostraram diferença significativa entre a duração média das fixações oculares nas áreas da professora e slide (Bonferroni; $p < 0,001$), professora e gestos (Bonferroni; $p < 0,001$), e marginalmente entre slide e gestos (Bonferroni; $p = 0,05$), todos para a condição *com gestos* ($n = 28$).

De maneira geral ($n = 57$), ao início da aula, os sujeitos privilegiaram de forma acentuada sua fixação ocular na região do slide (**Gráfico 25**). Após a etapa inicial introdutória, a distribuição do tempo de fixação tendeu gradualmente a um equilíbrio entre as duas regiões, privilegiando, contudo, praticamente durante toda a aula, a região do slide. A região do slide só será preterida pela região da professora na última seção da sequência didática, em que se apresenta o esquema gráfico. A sequência didática termina com uma forte baixa do tempo de fixação na região da professora e, de forma complementar, uma forte alta na região do slide para a última etapa contendo apenas informações bibliográficas (**Gráfico 25**).

Gráfico 25. Média geral da distribuição do tempo de fixação ocular nas regiões da professora e do slide ao longo da sequência didática ($\pm$ erro padrão; $n = 57$).

Se desconsiderarmos o último slide da aula, em que não mais é apresentado ou desenvolvido conteúdo da aula propriamente dito mas tão somente relacionadas as referências bibliográficas utilizadas, observamos ao longo da sequência didática tendências de aumento progressivo da porcentagem de tempo de fixação na região da

professora e de proporcional diminuição da porcentagem de tempo de fixação na região do slide (**Gráfico 26**).

Gráfico 26. Tendências gerais da distribuição do tempo de fixação ocular nas regiões da professora e do slide ao longo da sequência didática ($\pm$ erro padrão; $n = 57$).

Ainda analisando nossa população amostral como um todo ($n = 57$), ao considerarmos a sequência de slides (**Gráfico 27**) observamos oscilações importantes na proporção das fixações nas regiões analisadas. Ao considerarmos a região da professora (a região do slide apresentará comportamento inverso), a sequência se inicia com alta porcentagem de fixação (68%) durante a exibição do primeiro slide, contendo apenas o título da aula e a identificação da professora (ver **Anexo 3** para conteúdo e sequência de slides). Já a partir do segundo slide, em que se inicia a apresentação e desenvolvimento do conteúdo da aula propriamente dito, observamos um brusco decaimento (23%) na porcentagem de tempo de fixação na região da professora. Depois de algumas oscilações próximas ao intervalo entre 20% e 50% para os slides 2 a 5, notamos novo decaimento para 15% para o slide 6 e uma recuperação e manutenção da proporção do tempo de fixação próximo aos 50% para os slides 7 a 10, tendendo a um equilíbrio da distribuição das fixações entre as duas regiões da professora e do slide, com pequena vantagem para a região da professora. Contudo, o equilíbrio é interrompido por um novo decaimento progressivo acentuado de fixações na região da professora para os slides 11 e 12 (36% para o slide 11, e 6% para o slide 12), retornando logo após a um novo equilíbrio em torno dos 50%, com alternância de preferência entre as duas regiões (professora e slide) dentro

do intervalo de 40% a 60%. Por fim, durante a apresentação do último slide contendo as referências bibliográficas, a proporção de fixação na região da professora decaiu nova e bruscamente para a casa dos 18% (**Gráfico 27**).

Gráfico 27. Distribuição média do tempo de fixação ocular nas regiões da professora e do slide ao longo da sequência de slides ($\pm$ erro padrão; $n = 57$).

Também notamos a partir do **gráfico 27** acima a tendência a um baixo valor de porcentagem de tempo de fixação na região da professora para os slides 2, 4, 6, 12 e 19 (slides sem imagem – à exceção do slide 2 – e com esquema dos tópicos a serem desenvolvidos a seguir), acompanhando a tendência observada para os mesmos slides na análise da distribuição dos gestos ao longo da aula, com baixa ocorrência sobretudo de gesticulação dêitica e metafórica, mas também de gesticulação icônica e de pulso (**Gráfico 17**), bem como a baixa ocorrência de períodos para os referidos slides na análise da fala da professora (**Gráfico 14a**).

ENTRE AS CONDIÇÕES EXPERIMENTAIS

Considerando-se as condições experimentais, registramos: para a condição *com gestos* ($n = 28$), média de 303,1 segundos ($\pm 29,2$) de fixação ocular na região da

professora (sendo 16,3 [$\pm 3,5$] destes na região dos gestos) e 274,25 segundos ($\pm 19,5$) na região do slide; para a condição *sem gestos* (n = 29), média de 237,8 segundos ($\pm 21,8$) na região da professora e 317,1 ($\pm 21,4$) na região do slide (**Gráfico 28**).

Gráfico 28. Média de tempo de fixação ocular nas regiões da professora e do slide (segundos, $\pm$ erro padrão) para as condições experimentais *com gestos* (n = 28) e *sem gestos* (n = 29).

Levando-se em conta a média entre os participantes de suas porcentagens individuais de tempo de aula dedicado às regiões analisadas (**Gráfico 29**), observamos para a condição *com gestos* um equilíbrio da porcentagem de tempo de fixação entre a área da professora (51%; $\pm 2,7\%$) e a área do slide (49,0%; $\pm 2,6\%$). Já para a condição *sem gestos*, observamos um desequilíbrio de distribuição entre as áreas, com maior porcentagem do tempo de fixação na região do slide (57,8%; $\pm 2,8\%$) e menor porcentagem para a região da professora (42,2%; $\pm 2,8\%$). Conforme hipotetizamos, uma análise de variância (ANOVA) mostrou diferença significativa entre as condições experimentais *com gestos* e *sem gestos* quanto ao tempo fixação ocular na região da professora ($F_{(1,52)} = 4,992$; $p = 0,030$); mostrou ainda não haver diferença significativa do fator de aleatorização do lado da professora *à esquerda* e *à direita* do slide durante a apresentação da aula ($F_{(1,52)} = 0,316$; $p = 0,576$). A análise mostrou ainda não haver interação entre os fatores *condição experimental* e *lado da professora* ($F_{(1,52)} = 0,032$; $p = 0,860$).

Gráfico 29. Porcentagem média do tempo de fixação ocular nas regiões da professora e do slide nas condições experimentais com gestos e sem gestos ($\pm$ erro padrão; ANOVA $F_{(1,55)} = 4,992$; $p = 0,030$).

Assim como prevíamos, na condição *com gestos*, uma porcentagem diminuta do tempo de fixação ocular foi registrada na região do campo gestual da professora palestrante (**Gráfico 30**): em média, 2,9% do tempo total registrado ($\pm 2,9\%$). A análise dos dados mostrou um aumento progressivo da porcentagem de tempo de fixação na região dos gestos ao longo da aula, com (1,6%; $\pm 0,8\%$) para a etapa introdutória, 2,5% ($\pm 0,7\%$) para a sequência 1, 2,4% ($\pm 0,8\%$) para a sequência 2, 3,3% ($\pm 1,1\%$) para a etapa esquemática, 4,1% ($\pm 0,9\%$) para a etapa gráfica, e apenas 0,5% ($\pm 0,4\%$) para a última etapa com as referências bibliográficas.

Gráfico 30. Média geral da porcentagem de tempo de fixação ocular na região dos gestos da professora (condição *com gestos*; $\pm$ erro padrão; $n = 29$).

Ao longo da sequência didática (**Gráfico 31**), observamos entre as condições experimentais oscilações semelhantes de aumento e diminuição na porcentagem do tempo

de fixação ocular entre a região da professora e a região do slide. Contudo, ao longo de toda a sequência didática, apesar das oscilações observadas, a região do slide é privilegiada em todas as etapas para a condição *sem gestos*; enquanto que para a condição *com gestos*, pelas oscilações ocorrerem com valores mais aproximados entre as duas regiões, observa-se a alternância na dominância por uma ou outra região: a região do slide recebendo maior porcentagem de fixação nas etapas introdutória, esquemática e das referências da aula; e a região da professora, nas sequências 1 e 2 e na etapa gráfica. Por fim, observamos um forte efeito de decaimento do tempo de fixação na região da professora e aumento drástico do tempo de fixação na região do slide para ambas as condições experimentais *com gestos* e *sem gestos* durante a apresentação do último slide da aula, contendo as referências bibliográficas utilizadas.

Gráfico 31. Porcentagem média do tempo de fixação ocular para cada etapa didática ($\pm$ erro padrão) nas condições experimentais *com gestos* e *sem gestos*.

Em consonância com os dados obtidos pela análise das etapas da sequência didática acima apresentada, podemos observar mais precisamente através da análise da sequência de slides abaixo apresentada (**Tabela 6 e Gráfico 32**) valores de porcentagem de fixação ocular na região da professora superiores para a condição *com gestos* (em relação à condição *sem gestos*) ao longo de toda a sequência de slides. Destacamos os slides 13 em diante (à exceção do slide 19, com as referências bibliográficas), com maiores diferenças entre as condições experimentais *com gestos* e *sem gestos*.

No SLIDE	PROFESSORA				SLIDE			
	CG	EP	SG	EP	CG	SE	SG	SE
1	70%	3,4%	67%	4,3%	30%	3,4%	33%	4,3%
2	24%	2,6%	21%	3,2%	76%	2,6%	79%	3,2%
3	41%	4,4%	37%	4,0%	59%	4,4%	63%	4,0%
4	28%	4,4%	24%	4,8%	72%	4,4%	76%	4,8%
5	53%	3,8%	46%	4,9%	47%	3,8%	54%	4,9%
6	16%	2,9%	14%	2,3%	84%	2,9%	86%	2,3%
7	46%	3,5%	40%	3,5%	54%	3,5%	60%	3,5%
8	53%	3,7%	49%	4,0%	47%	3,7%	51%	4,0%
9	56%	3,0%	47%	4,2%	44%	3,0%	53%	4,2%
10	60%	3,7%	53%	3,9%	40%	3,7%	47%	3,9%
11	40%	4,5%	31%	4,7%	60%	4,5%	69%	4,7%
12	8%	2,2%	4%	1,3%	92%	2,2%	96%	1,3%
13	58%	3,7%	46%	4,5%	42%	3,7%	54%	4,5%
14	48%	4,4%	38%	4,0%	52%	4,4%	62%	4,0%
15	60%	4,0%	53%	4,3%	40%	4,0%	47%	4,3%
16	51%	4,7%	29%	4,3%	49%	4,7%	71%	4,3%
17	44%	3,8%	37%	4,6%	56%	3,8%	63%	4,6%
18	62%	3,2%	46%	4,8%	38%	3,2%	54%	4,8%
19	16%	2,4%	19%	3,3%	84%	2,4%	81%	3,3%

Tabela 6. Porcentagem média de tempo de fixação ocular nas regiões da professora e do slide ($\pm$ erro padrão) para a sequência de slides nas condições experimentais *com gestos* e *sem gestos*.

Gráfico 32. Porcentagem média de tempo de fixação ocular nas regiões da professora e do slide ($\pm$ erro padrão) para a sequência de slides nas condições experimentais *com gestos* e *sem gestos*.

Quanto à duração média (segundos) da fixação ocular nas áreas de interesse da professora e do slide, considerando-se agora as condições experimentais *com gestos* e *sem gestos*, observamos valores ligeiramente superiores para a condição *sem gestos* (**Gráfico 33**): para a região da professora, observamos 0,46 segundos ($\pm 0,03$) para a condição experimental *com gestos*, e 0,53 segundos ($\pm 0,04$) na condição experimental *sem gestos*; para a região do slide, observamos 0,21 segundos ($\pm 0,01$) para a condição *com gestos*, e 0,24 segundos ($\pm 0,01$) para a condição *sem gestos*. Uma análise de variância (ANOVA), no entanto, mostrou não haver diferença significativa entre as condições experimentais no que diz respeito à duração média da fixação nas regiões da professora ($F_{(1,55)} = 2,295$; $p = 0,136$); e uma análise Mann-Whitney (os valores para região do slide não apresentam normalidade em sua distribuição) mostrou não haver diferença significativa entre as condições experimentais para a região do slide ($Z = -1,699$; $p = 0,089$). A duração média da fixação na região dos gestos (condição *com gestos*) foi de 0,30 segundos ($\pm 0,03$).

Gráfico 33. Duração média da fixação nas regiões da professora, do slide e dos gestos, nas condições experimentais *com gestos* (- erro padrão) e *sem gestos* (+ erro padrão).

No **gráfico 34** abaixo, analisamos a duração média da fixação ocular nas áreas de interesse da professora e do slide para as condições experimentais *com gestos* e *sem gestos* levando-se em conta os valores observados para cada slide ($n = 19$). Uma análise não-paramétrica de medidas relacionadas mostrou diferença significativa entre as condições

com gestos e sem gestos para a região da professora (Wilcoxon Signed Ranks; $Z = -3,743$; $p < 0,001$) e para a região do slide ($Z = -3,219$; $p = 0,001$). O gráfico evidencia ainda oscilações e desvios padrões de maior amplitude para a condição *sem gestos*, tanto para a região da professora quanto para a região do slide. Ademais, observamos para os slides 12 e 19, e mais moderadamente para os slides 2, 4 e 6, a ocorrência de valores muito abaixo dos observados para os demais slides, indicando eventos de natureza distinta já apontados nas análises acima apresentadas (**Gráfico 34**).

Gráfico 34. Duração média da fixação ocular nas regiões da professora, do slide e dos gestos ($\pm$ erro padrão), para as condições experimentais *com* e *sem gestos* ao longo da sequência de slides.

Quanto à porcentagem de tempo da aula dedicado à fixação na região dos gestos, esperávamos encontrar maiores valores em momentos da aula com maior quantidade de gesticulação icônica produzida pela professora. Realizamos análises correlacionais entre a porcentagem de tempo de fixação ocular na região dos gestos e o número de ocorrências para cada tipo de gesto; bem como entre a porcentagem de tempo de fixação ocular na região dos gestos e a frequência de ocorrência por segundo de fala para cada tipo de gesto, ambas análises levando-se em consideração a unidade de slide ($n = 19$; **Gráfico 35**).

Uma análise correlacional de Spearman mostrou correlação significativa entre o

tempo de fixação ocular nos gestos e o número de ocorrências de gestos metafóricos (Spearman $\rho = 0,651$; $p = 0,003$); bem como correlação marginal entre a fixação ocular na região dos gestos e o número de ocorrências de gestos dêiticos ($\rho = 0,467$; $p = 0,044$) e de gestos de pulso ($n = 19$; Spearman $\rho = 0,447$; $p = 0,055$). No entanto, ao considerarmos a frequência de ocorrência (por segundo de fala) dos gestos, a análise evidenciou apenas correlação significativa entre o tempo de fixação ocular na região dos gestos e a frequência de gesticulação metafórica ($n = 19$; Spearman $\rho = 0,650$; $p = 0,003$).

Uma análise de regressão linear múltipla foi ainda utilizada para verificar se tanto o número de ocorrências quanto a frequência dos gestos poderiam prever o tempo de fixação ocular na região dos gestos. A análise resultou em um modelo estatisticamente significativo ($F_{(1,17)} = 6,993$; $p = 0,017$; $R^2 = 0,291$), apontando o número de ocorrências de gestos metafóricos ($\beta = 0,540$; $t = 2,645$; $p = 0,017$) como preditor do tempo de fixação ocular na região dos gestos; bem como de sua frequência de ocorrência por segundo de fala ($F_{(1,17)} = 9,095$; $p = 0,008$; $R^2 = 0,349$; $\beta = 0,590$; $t = 3,016$; $p = 0,008$).

Gráfico 35. Frequência de ocorrência de gestos de pulso, dêiticos, metafóricos e icônicos por segundo de fala e porcentagem de tempo de fixação na região dos gestos ($\pm$ erro padrão; $n = 19$).

Também por uma análise correlacional de Spearman, verificamos correlação significativa da porcentagem de tempo de fixação na região da professora na condição

com gestos e: o número de ocorrências de gestos metafóricos ($n = 19$; Spearman $\rho = 0,688$; $p = 0,001$), de pulso ($\rho = 0,602$; $p = 0,006$) e icônicos ($\rho = 0,495$; $p = 0,031$). Ao considerarmos a frequência de ocorrência (por segundo de fala) dos gestos, encontramos correlação positiva da porcentagem de tempo de fixação na região da professora na condição com gestos e a frequência de gestos metafóricos ($\rho = 0,530$; $p = 0,020$), e marginalmente de gestos icônicos ($\rho = 0,447$; $p = 0,055$; **Gráfico 36**). Uma correlação parcial com controle pela duração dos slides, no entanto, mostrou correlação entre o tempo de fixação na região da professora na condição *com gestos* e apenas a frequência de gestos metafóricos ($\rho = 0,656$; $p = 0,003$) e de pulso ($\rho = 0,517$; $p = 0,028$).

Uma regressão linear múltipla para verificar se a frequência de ocorrência dos diferentes tipos de gestos (de pulso, dêiticos, metafóricos e icônicos) seria capaz de prever o tempo de fixação na região da professora resultou em um modelo estatisticamente significativo ($F_{(1,17)} = 16,321$; $p = 0,001$; $R^2 = 0,490$), apontando a frequência da gesticulação metafórica ($\beta = 0,700$; $t = 4,040$; $p = 0,001$) como preditora da fixação ocular na região da professora na condição *com gestos*.

Gráfico 36. Porcentagem média de tempo de fixação ocular na região da professora ($\pm$ erro padrão) e frequência de ocorrência de gestos de pulso, dêiticos, metafóricos e icônicos por segundo de fala ($n = 19$).

A seguir, apresentamos um ensaio gráfico a título de ilustração da composição semiótica gestual-discursiva da sequência de slides com suas respectivas duração e

porcentagens de tempo de fixação ocular para as regiões dos gestos e da professora (**Gráfico 37**). Para melhor visualização da composição, os parâmetros se encontram fora de escala.

Gráfico 37. Gráfico ilustrativo da composição semiótica gestual-discursiva da sequência de slides com respectivas duração e porcentagens de tempo de fixação ocular para as regiões dos gestos e da professora (parâmetros fora de escala).

ENTRE AS METADES

Para verificarmos eventuais mudanças comportamentais no padrão de fixação ocular dos participantes dependentes da variável de tempo de exibição do slide, analisamos o tempo de fixação ocular nas regiões estudadas levando-se em consideração a 1ª e a 2ª metades do tempo de exibição de cada slide. Observamos média de 1,7% ($\pm 0,3$) de tempo de fixação na região dos gestos para a primeira metade, e 2,8% ($\pm 0,4$) para a segunda metade ($n = 19$; **Gráfico 38**). Uma análise não-paramétrica mostrou diferença marginalmente significativa entre a 1ª e a 2ª metade de apresentação dos slides quanto à porcentagem de tempo de fixação ocular dedicada à região dos gestos (Wilcoxon; $Z = -1,962$; $p = 0,050$; **Gráfico 38 e 39**).

Gráfico 38. Porcentagem de tempo de fixação ocular dedicada à região dos gestos para a 1ª e 2ª metades do tempo de apresentação dos slides.

Gráfico 39. Médias estimadas da porcentagem de tempo de fixação ocular na região dos gestos para a 1ª e 2ª metades do tempo de apresentação dos slides.

Considerando-se a região da professora, em ambas condições experimentais observamos maior tempo médio de fixação na região para a 2ª metade em comparação com a 1ª metade: na condição *com gestos*, observamos 33% ($\pm 3\%$) para 1ª metade e 55% ($\pm 5\%$) para a 2ª metade; na condição *sem gestos*, observamos 26% ($\pm 3\%$) para a primeira metade e 46,2% ($\pm 5\%$) para a 2ª metade (**Gráfico 41**). Uma análise GLM multivariada mostrou efeito da condição experimental (*com gestos* e *sem gestos*; $F_{(1,18)} = 33,774$; $p <$

0,001) e efeito das metades ($F_{(1,18)} = 72,371$; $p < 0,001$) sobre a porcentagem de tempo de fixação na região; mostrou também não haver interação entre os fatores condição experimental e metades ($F_{(1,18)} = 0,918$; $p = 0,351$). Como se observa no **Gráfico 40** abaixo, o fator metades dos slides (*1ª metade* e *2ª metade*) teve forte impacto no comportamento de fixação ocular dos participantes na região da professora.

Gráfico 40. Porcentagem do tempo de fixação na região da professora para as 1ª e 2ª metades do tempo de apresentação dos slides.

Gráfico 41. Médias estimadas da porcentagem de tempo de fixação ocular na região da professora para a 1ª e 2ª metades do tempo de apresentação dos slides.

Quanto à duração média da fixação ocular na região dos gestos, considerando-se a 1ª e 2ª metades de apresentação dos slides, também verificamos valores superiores para a 2ª metade: tempo médio de 0,24 segundos ($\pm 0,01$) para a 1ª metade, e 0,26 segundos ($\pm 0,02$) para a 2ª metade (ANOVA; $F_{(1,17)} = 4,768$; $p = 0,043$; **Gráfico 42**).

Gráfico 42. Duração média da fixação na região dos gestos para a 1ª e 2ª metades da duração dos slides.

Gráfico 43. Tempo médio de fixação ocular na região dos gestos para a 1ª e 2ª metades do tempo de apresentação dos slides.

Para a região da professora, não observamos grande discrepância na duração média da fixação entre as duas metades dos slides para a condição *com gestos*: média de 0,40 segundos ($\pm 0,02$) para a 1ª metade, e 0,42 segundos ($\pm 0,02$) para a 2ª metade. Já para a condição *sem gestos*, observamos maior discrepância entre os valores correspondentes às duas metades: 0,46 segundos ($\pm 0,03$) para a 1ª metade, e 0,56 segundos ($\pm 0,03$) em média para a 2ª metade (**Gráfico 44 e 45**). Uma ANOVA mostrou efeito da condição experimental *com gestos* e *sem gestos* ($F_{(1,18)} = 71,028$; $p < 0,001$) e das metades ($F_{(1,18)} = 9,784$; $p = 0,006$), bem como efeito de interação entre os fatores condição experimental e metades ($F_{(1,18)} = 6,613$; $p = 0,019$) no tempo médio de fixação na região da professora.

Gráfico 44. Duração média da fixação na região da professora para a 1ª e 2ª metades dos slides.

Gráfico 45. Tempo médio de fixação ocular na região da professora para a 1ª e 2ª metades do tempo de apresentação dos slides.

Quanto à duração média da fixação na região do slide, observamos: para a condição *com gestos*, pequena diferença com duração média da fixação de 0,20 segundos ($\pm 0,00$) para a 1ª metade e 0,21 ($\pm 0,00$) para a 2ª metade; para a condição *sem gestos*, diferença mais expressiva com 0,22 segundos ($\pm 0,01$) para a 1ª metade, e 0,26 segundos ($\pm 0,02$) para a 2ª metade (**Gráfico 46**). Uma ANOVA mostrou efeito da condição experimental ($F_{(1,18)} = 10,169$; $p = 0,005$) e das metades ($F_{(1,18)} = 10,685$; $p = 0,004$); mostrou ainda que não houve efeito de interação entre os fatores ($F_{(1,18)} = 2,032$; $p = 0,171$).

Gráfico 46. Duração média da fixação na região do slide para a 1ª e 2ª metades do tempo de exibição dos slides.

Tempo médio de Fixação no SLIDE

Gráfico 47. Tempo médio de fixação ocular na região da professora para a 1ª e 2ª metades do tempo de apresentação dos slides.

APRENDIZAGEM

QUESTIONÁRIO GERAL DE AVALIAÇÃO

No questionário geral de avaliação, sobre a facilidade de aprendizagem, 8 pessoas relataram *pouca* facilidade de aprendizagem, 33 *moderada* e 17 *intensa* (**Gráfico 48**).

Gráfico 48. Facilidade de aprendizagem.

Sobre se conheciam o conteúdo, 21 participantes responderam sim e 36 não (**Gráfico 49**).

Gráfico 49. Conhecimento prévio.

Do nível de conhecimento do assunto, 28 responderam não terem conhecimento prévio, 19 responderam pequeno conhecimento prévio, 6 médio, e 4 elevado (**Gráfico 50**).

Gráfico 50. Nível de conhecimento prévio.

Sobre dificuldade em alguma parte da aula, 24 relataram *sim* e 33 *não* (**Gráfico 51**). Dos que relataram alguma dificuldade, 7 relataram dificuldade com as *classificações dos comportamentos*, 7 relataram dificuldade na *compreensão de algum comportamento*

específico, 5 na compreensão de algum exemplo específico, 5 na conceitualização teórica geral de comportamento, 1 na parte final da aula por dispersão (Gráfico 51).

Gráfico 51. Dificuldade na aula.

Gráfico 52. Natureza da dificuldade da aula.

Sobre o nível de dificuldade da aula, 32 relataram *baixa* dificuldade, 25 relataram *média* dificuldade, e nenhum relatou *alta* dificuldade (**Gráfico 53**). Sobre se gostaram de ter participado da experiência de aprendizagem do conteúdo, todos relataram *sim*.

Gráfico 53. Nível de dificuldade da aula.

Para os itens da questão 8 do questionário de avaliação de aprendizagem, questões objetivas em que os participantes tiveram de classificar os exemplos de comportamento dados, observamos para os itens *a*, *b* e *d* uma regularidade com baixa ocorrência de respostas totalmente incorretas (média de 1,7 entre as questões), alta ocorrência de respostas meio-certas (média 39,3), moderada ocorrência de respostas totalmente certas (média 15), e baixa ocorrência de respostas não respondidas (média 1). O item *c* destoou dos demais itens com grande ocorrência de erros (32) e moderada ocorrência de respostas meio-certas (9), porém mantendo ocorrência moderada de respostas totalmente corretas (16) e baixa ocorrência de respostas não respondidas (0) semelhante aos itens *a*, *b* e *d* (**Gráfico 54**).

Gráfico 54. Índice de acertos na classificação dos comportamentos (questão 8).

A seguir, apresentamos um ensaio gráfico a título de exploração da relação da frequência de ocorrência dos quatro tipos de gestos analisados (de pulso, dêiticos, metafóricos e icônicos) com a porcentagem de tempo de fixação ocular dedicado às regiões dos gestos e da professora (para ambas condições *com* e *sem gestos*) para os slides de número 10, 13, 14 e 15, respectivamente relacionados aos itens *a*, *b*, *c* e *d* da questão 8 do questionário de aprendizagem. O gráfico sugere o item *c* da questão 8 diferir dos demais itens *b*, *c* e *d* pela ausência de ocorrência de gesticulação icônica e pequena baixa nos valores observados para o tempo de fixação na região da professora e dos gestos. Para melhor visualização, os parâmetros apresentados encontram-se fora de escala (**Gráfico 55**).

Gráfico 55. Frequência de ocorrência dos gestos e porcentagem de tempo de fixação ocular dedicada às regiões dos gestos e da professora para os slides 10, 13, 14 e 15, respectivamente relacionados às questões *a*, *b*, *c* e *d* da questão 8 do questionário de aprendizagem (parâmetros fora de escala).

Não encontramos correlação entre as pontuações dos sujeitos no questionário de aprendizagem e os parâmetros de fixação ocular analisados (tempo de fixação ou tempo médio de fixação para as regiões analisadas).

Os voluntários responderam ainda a duas questões dissertativas, uma sobre a importância de se estudar o comportamento animal e a outra sobre a vantagem do comportamento inato em relação ao comportamento aprendido. Observou-se efeito teto nas duas questões: para a primeira, 53 responderam de forma satisfatória e apenas 4 responderam

de forma parcialmente satisfatória. Para a segunda questão, 55 responderam de forma satisfatória e apenas 2 responderam de forma parcialmente satisfatória.

DISCUSSÃO PARCIAL

COMPOSIÇÃO DISCURSIVO-GESTICULATÓRIA

Neste segundo experimento investigamos o recrutamento e o impacto da gesticulação pela professora palestrante na alocação da atenção visual e na aprendizagem de alunos em uma aula de ciências. Controlamos a visualização dos gestos pelos sujeitos participantes.

Na preparação da aula apresentada como estímulo aos sujeitos experimentais, optamos pelo enquadramento da professora separado do enquadramento dos slides como uma solução simples para o controle experimental da exibição dos gestos. De outra forma, em uma filmagem de uma aula convencional na qual a professora se movimentasse diante à projeção e interagisse com os slides, seria tecnicamente mais complexo o controle experimental desejado sem a produção de áreas no vídeo com algum efeito distrator como montagens ou borramentos. Além disso, nossa opção de enquadramento reduziu a produção de gesticulação dêitica (de apontamento) pela professora durante a gravação da aula, permitindo-nos abordar, mais ao nosso interesse e durante toda a apresentação do conteúdo, o efeito da gesticulação ilustradora, sobretudo de gestos de pulso e icônicos, sobre a alocação da atenção dos estudantes na região da professora em concorrência com a região do slide; em uma aula convencional filmada, com interação da professora com os slides projetados, a professora forçosamente direcionaria com frequência a atenção dos participantes à região do slide através do direcionamento de seu próprio olhar, orientação corporal e gestos de apontamento (ver sobre proxemia em Giordan et al., 2015). Ademais, a produção dos estímulos de ambas condições experimentais a partir de uma única gravação da aula garantiu que os sujeitos experimentais recebessem não apenas o mesmo conteúdo de aula, mas também a mesma produção linguístico-discursiva, com mesmas construções, entonações, pausas e até expressões faciais.

Conforme observado pela análise da fala da professora, 82% do tempo de aula foram computados como fala e os 18% restantes como silêncio em pausas sintáticas e, em sua maior parte, silêncios durante transições de slides. O ritmo prosódico da fala da

professora manteve-se relativamente constante com frequência média de 1,68 acentos tônicos frasais por segundo de fala durante a aula (média de 2,04 entre os slides), com frequência mais alta nas extremidades inicial e final da aula (**Gráfico 13**). A maior frequência de acentuações tônicas frasais verificada na parte introdutória da aula sugere o início de um discurso teórico introdutório ensaiado, de fala rápida e sem necessidade de desenvolvimento ou explanação. Igualmente verificada na parte final da aula, a fala da professora apresenta uma frequência de acentuações tônicas frasais mais elevada, sugerindo também aqui uma fala acelerada, porém, qualitativamente diferente. Uma possível explicação seria a maior complexidade dos conteúdos dos slides finais, cujo desenvolvimento tenha exigido maior extensão temporal e um fluxo de fala contínuo e engajado - o slide esquemático de número 17 com duração de 121,3 segundos e o slide gráfico de número 18 com duração de 143,5 segundos, sendo a média geral entre os slides 48,3 ($\pm 8,5$) segundos de duração. Outro fator que possa ter contribuído para o aumento observado na frequência de acentuações frasais da fala durante os slides finais da aula seja um possível aumento do nível de ansiedade da professora pela proximidade de término da aula.

Através da análise das entonações conclusivas da fala da professora (**Gráfico 14b**), buscamos compreender a composição das unidades semântico-discursivas: o número de períodos indicando o volume de unidades semânticas, e sua duração média indicando seu grau de complexidade. Observamos claramente a concentração de maior volume de períodos ao interior das sequências didáticas 1, 2, esquemática e gráfica, sugerindo maior complexidade de conteúdo; de forma complementar, observou-se baixo volume de períodos ao longo de toda a sequência introdutória e para os slides 6, 11, 12 e 19, justamente os slides de menor duração e de pouco desenvolvimento.

Verificamos, no entanto, que a duração média dos períodos, diferentemente do número de períodos, não poderia ser compreendida como índice de complexidade do conteúdo do respectivo slide, pois os períodos nitidamente mais longos ocorreram para os slides 3 (da etapa introdutória) e 19 (com as referências bibliográficas), nenhum deles com complexidade que demandasse explanação ou desenvolvimento complexo de conteúdo (**Gráfico 20**). Especulamos a esse respeito se os períodos mais longos ao início da aula, sobretudo registrados para o slide 3, poderiam eles também ser explicados por um discurso retórico introdutório mais simples, de baixa complexidade, mais facilmente ensaiado e memorizado; bem como se a tendência de períodos mais longos para o último slide da aula pudesse estar associada a um aumento observado no fluxo da fala por um

possível aumento no nível de ansiedade da professora com a proximidade do término da aula.

Observamos ainda fenômeno semelhante para as sequências didáticas 1 e 2: slides iniciais com valores ligeiramente inferiores, próximos dos 5 segundos de duração média dos períodos, com leve tendência de aumento, aproximando-se dos 10 segundos, para os slides mais finais. No mais, e de maneira geral, os períodos apresentaram relativa constância, com média de duração variando entre 5 a 10 segundos ao longo da aula (**Gráfico 14a**).

Ainda no esforço de compreensão da composição das unidades semântico-discursivas da fala da professora, analisamos sua gesticulação manual ao longo de toda sequência didática. Foram contabilizados 597 golpes gesticulatórios. Esperávamos encontrar número de ocorrências e frequência aproximados entre golpes gesticulatórios e acentuações tônicas frasais. No entanto, o número de ocorrências de acentuações tônicas frasais (1532) superou em aproximadamente 2,6 vezes o número de ocorrências de golpes gesticulatórios (597), estes ocorrendo com frequência média de 0,65 por segundo de fala, e aqueles com frequência de 1,68 por segundo de fala (**Gráficos 13, 18 e 20**). Tal resultado sugere unidades rítmicas gesticulatórias de maior extensão temporal (de menor frequência) do que as unidades rítmicas prosódicas da fala, provavelmente devido a um custo cognitivo-motor superior em seu planejamento e execução, mas também e principalmente por, assim entendemos, sua ocorrência marcar unidades semântico-discursivas de maior abrangência do que aquelas marcadas pela acentuação tônica da fala.

A distribuição dos gestos, assim como esperado, deu-se conforme a organização dos conteúdos da sequência didática apresentada na aula e duração temporal dos slides, com maiores concentrações de ocorrências, especialmente em relação aos gestos metafóricos e icônicos mas também em relação aos gestos de pulso e dêiticos de modo mais geral, em slides explanatórios mais duradouros; e menores concentrações em passagens mais curtas, introdutória ou de transição entre as sequências didáticas, a saber, slides 1, 12 e 19, com menor volume informacional e conteúdo de baixa complexidade (**Gráfico 17**; correlação de Pearson, $n = 19$; $r = 991$; $p < 0,001$).

Esperávamos encontrar maior frequência de gesticulação de pulso em passagens teóricas, como na introdução e nas definições gerais dos comportamentos, e maior frequência de gesticulação icônica em passagens explicativas, descritivas e de exemplificações. A análise da frequência de ocorrência dos gestos por segundo de fala (**Gráfico 18**) mostrou, contudo, alta frequência da gesticulação de pulso ao longo de toda

a aula, com decaimentos mais agressivos para os slides 2, 12 e 19, justamente slides introdutórios (slides 2 e 12) e terminal (slide 19) de caráter mais teórico, portanto de baixa complexidade de compreensão e de menor duração, sugerindo, em direção contrária à esperada, que gestos de pulso tenham emergido antes associados ao engajamento e esforço cognitivo explanatório, de desenvolvimento do conteúdo pela professora palestrante durante a aula. Gestos de pulso apresentaram, ademais, a maior amplitude de oscilações dentre os quatro tipos de gestos.

Quanto à gesticulação dêitica, com valor organizacional dos referetes discursivos no campo gestual, fez-se presente ao longo de toda a aula (**Gráfico 18**), à exceção do primeiro slide, resultado compreensível, uma vez que os referentes discursivos não tivessem sido ainda apresentados. Dois momentos da aula concentraram as ocorrências de gesticulação dêitica, a saber, o slide 3, com a definição teórica de comportamento, sugerindo a apresentação e localização espacial de referentes discursivos a serem desenvolvidos durante a aula, e o slide 15, com um exemplo de comportamento de imitação, exemplo de movimentos complexos de manipulação de objetos (lavagem de batatas) por primatas.

Já a gesticulação metafórica tendeu a uma regularidade ao longo da aula (**Gráfico 18**), acompanhando em sua maior parte e com menor amplitude as tendências de oscilação da gesticulação de pulso e, de forma aproximadamente inversa em algumas passagens, as da gesticulação dêitica. A gesticulação metafórica apresentou, ademais, leve tendência a um aumento de sua frequência no interior das sequências didáticas (do meio para o fim das sequências 1 e 2), bem como fortes decaimentos para os slides 2, 6 e 12, slides teóricos introdutórios ou de transição entre sequências didáticas, e para o slide terminal 19, com as referências bibliográficas. Uma análise qualitativa da aula sugere que, assim como a gesticulação de pulso, a gesticulação metafórica tenha estado intimamente relacionada com o engajamento e o esforço cognitivo pela professora em passagens de desenvolvimento de conteúdo da aula.

Por fim, a gesticulação icônica apresentou baixa frequência de modo geral durante a aula, com as maiores ocorrências no interior da sequência didática 1 (slides 7 a 10) e slide gráfico (slide 18; **Gráfico 18**). Em suma, gestos icônicos também acompanharam – porém, bastante discretamente – a tendência geral da composição da frequência de gestos metafóricos com gestos de pulso.

Ao analisarmos a correlação entre a produção de gestos e a produção linguística levando-se em conta o *número* de ocorrências por slide, encontramos alta correlação

positiva entre todos os parâmetros analisados, resultado pouco informativo e já esperado dada a grande variância observada na duração dos slides (maior número de ocorrências de gestos e dos parâmetros linguísticos analisados, períodos e acentuações tônicas frasais, em slides com maior duração, e menor número de ocorrências em slides com menor duração).

A partir desse resultado, para melhor compreendermos a ocorrência de gestos na composição semântico-discursiva da aula, conduzimos, então, análise correlacional (Spearman) para a *frequência* (por segundo de fala) de ocorrência dos parâmetros em questão (**Tabela 5**). Interessantemente, observamos correlação positiva apenas entre a duração dos slides e a frequência de gestos icônicos ($\rho = 0,698$, $p = 0,001$; gestos metafóricos apresentando fraca tendência a uma correlação positiva com a duração dos slides, $\rho = 0,407$, $p = 0,083$). Tal resultado nos sugere maior produção de gesticulação icônica em passagens da aula de maior duração temporal, eventualmente com maior número de períodos e, sendo assim, de maior complexidade e desenvolvimento semântico – tal comportamento pode ser observado também ao analisarmos as etapas didáticas (**Gráfico 19**): percebe-se uma relativa constância na frequência dos gestos de pulso, dêiticos e metafóricos, e uma grande oscilação na frequência de gestos icônicos, com picos na sequência 1 e gráfica, e vales na etapa introdutória, sequência 2 e de referências bibliográficas. Cabe aqui perguntarmo-nos em que momento do desenvolvimento semântico-discursivo (se no momento de aumento de carga cognitiva em passagens de maior complexidade semântica, se em passagens expletivas de retomada de mesmo referente discursivo, ou outro) seriam disparadas as representações manuais de aspectos concretos de referentes discursivos, características da gesticulação icônica. Outras duas perguntas que se colocam neste momento de esforço de compreensão da composição semiótica gestual-discursiva são como medir a complexidade do conteúdo explanado pela professora em cada slide, e em que medida poderiam os gestos (e cada um dos quatro tipos de gestos ilustradores) ser tomados como índices dessa complexidade.

FIXAÇÕES OCULARES

Quanto às fixações oculares, dos 917,9 segundos totais de tempo de aula, registramos uma média entre os sujeitos participantes de 565,9 segundos ($\pm 23,5$) de tempo de visualização da aula, ou seja, uma perda informacional de 38,3% ($\pm 0,03$) do

tempo de aula. Tal perda poderia ser explicada por erro de calibração do equipamento causada por excesso de movimentação e mudança postural dos participantes ao longo da extensão da aula (extensão total de 15 minutos e 17 segundos). Alguns participantes relataram, no entanto, o hábito do desvio do olhar, normalmente para regiões superiores do campo visual fora da área de exibição do estímulo (portanto, fora da área de calibração do equipamento de rastreamento ocular), tanto como estratégia cognitiva de reflexão, digressão ou busca de informação em memória, como também – assim relataram alguns – uma estratégia de concentração de recursos atencionais no processamento de informações auditivas (da fala da professora), pela eliminação de estímulos visuais que pudessem lhe vir a ser distratores.

Tais relatos dos sujeitos participantes levam-nos a compreender que estímulos mais duradouros como a sequência didática de 15 minutos em nosso estudo em setups experimentais de natureza mais ecológica e exploratória, apesar de colocar maiores desafios ao controle experimental, podem nos fornecer pistas gerais importantes sobre estratégias cognitivas atencionais empregadas regularmente por estudantes em ambientes de aula; em nosso estudo, apesar da grande perda informacional, foi-nos possível, mediante também o relato dos participantes, termos acesso a tais estratégias, inicialmente não previstas por nosso desenho experimental mas com grande potencial de vir a ser objeto de investigação em trabalhos futuros. Um estudo experimental que controle o tempo de apresentação tanto dos slides quanto da totalidade da aula pode nos trazer dados interessantes diferentes daqueles por nós obtidos, bem como nos dizer de um possível índice de perda de calibração correlacionado à variável tempo, que possam nos ajudar a melhor interpretar nossos resultados e melhor compreender o fenômeno da aprendizagem. Também a replicação de nosso estudo com metodologia e equipamento de rastreamento ocular com tecnologia vestível, por exemplo, rastreadores oculares instalados em armações de óculos ao invés de equipamentos fixos à mesa à frente do sujeito experimental como em nosso experimento, possa nos trazer informações relevantes sobre tais estratégias atencionais de desvios do olhar.

Em vista dessa perda informacional, consideramos duas possibilidades de análise de nossos dados de fixação ocular. Uma primeira forma seria a contabilização e análise da somatória do dado bruto de tempo de fixação dos participantes em segundos - assim o fizemos inicialmente e de maneira geral a título de descrição dos dados, para o tempo total dedicado à região da professora e do slide, considerando-se a totalidade de nossa amostra ($n = 57$) e os grupos analisados, *com gestos* ($n = 28$) e *sem gestos* ($n = 29$). Dessa

forma, no entanto, nossa análise atribuiria aos participantes tanto maior peso informativo em nossos resultados quanto maior fosse sua porcentagem de tempo de visualização de aula registrado (ou menor perda de calibração). Uma alternativa encontrada, para que eventuais perdas de calibração por necessidades de mudança postural ou por estratégias cognitivas alternativas individuais de fixação ocular não diminuíssem o valor informativo em nossos resultados de cada um de nossos participantes (estivemos interessados em explorar o comportamento atencional geral dos participantes em função da presença/ausência da gesticulação da professora), foi calcularmos, a partir do tempo total de fixação registrado para cada indivíduo, sua porcentagem correspondente de tempo dedicado à região da professora e sua porcentagem correspondente de tempo dedicado à região do slide. Uma vez descritos os valores iniciais mais gerais de tempo de fixação nas regiões da professora e do slide em unidades de segundos (**Gráfico 22 e 29**), passamos ao cálculo das porcentagens individuais do tempo de fixação dedicado pelos sujeitos para cada uma das duas regiões como opção de tratamento dos dados para as análises subsequentes.

A título de descrição dos dados, portanto, em termos de valor temporal absoluto (em segundos), verificamos de maneira geral entre os 57 participantes uma prevalência pela área do slide em relação à área da professora: de uma média entre os participantes de 565,9 segundos totais registrados ($\pm 23,5$), 269,9 segundos ($\pm 18,34$) estiveram dedicados à região da professora e 296,0 segundos ($\pm 14,5$) à região do slide (**Gráfico 22**). Ao considerarmos as condições experimentais *com gestos* e *sem gestos*, no entanto, verificamos diferença na distribuição das fixações oculares: para a condição *com gestos* ($n = 28$), 303,1 segundos ($\pm 29,2$) na região da professora (sendo 16,3 segundos [$\pm 3,5$] destes na região dos gestos) e 274,25 segundos ($\pm 19,5$) na região do slide; para a condição *sem gestos* ($n = 29$), 237,8 segundos ($\pm 21,8$) na região da professora e 317,1 ($\pm 21,4$) na região do slide (**Gráfico 28**). Discutiremos daqui em diante, contudo, conforme as razões acima explicadas, valores calculados pela média entre os voluntários de suas porcentagens individuais registradas para as áreas de interesse analisadas.

Computamos como média geral entre os participantes de suas porcentagens individuais 46,5% ($\pm 2\%$) do tempo para a área da professora e 53,5% ($\pm 2\%$) para a área do slide (**Gráfico 22**). Os **gráficos 26 e 27** mostram uma tendência geral de aumento progressivo da porcentagem de tempo de fixação na região da professora ao longo da aula. De maneira geral, imediatamente após a fala de apresentação da professora durante a exibição do primeiro slide, a região do slide recebe alta preferência de fixação ocular,

concentrando em si uma média de 70% do tempo de fixação para os cinco slides seguintes. Contudo, no decurso da aula, observa-se aumento progressivo da porcentagem de tempo dedicado à região da professora, tendendo a um equilíbrio entre as duas regiões – eventualmente sugerindo uma possível inversão do interesse inicial da aula, a região da professora atraindo mais interesse que a região do slide a partir do slide gráfico, slide 18. A tendência ao equilíbrio entre as regiões é, no entanto, interrompida por fortes altas na região do slide para os slides 6 e 12 introdutórios às sequências didáticas 1 e 2, que serão discutidas mais adiante.

A mudança progressiva ao longo da aula de preferência entre as regiões nos coloca a especularmos sobre se haveria mudanças progressivas significativas na qualidade informacional dos slides apresentados ao longo da aula que justificassem uma progressiva queda de interesse pela região e consequente aumento de interesse pela região da professora; ou se, por outro lado, haveria fatores sociais interferindo em estratégias cognitivas atencionais de busca visual que entrariam em marcha a partir de um determinado limiar temporal. As análises das metades de tempo de exibição dos slides nos ajudarão no encaminhamento da questão. De qualquer maneira, uma análise exploratória mais detalhada da extensão temporal da exibição dos slides, por exemplo dividindo seu tempo de exibição em quartis ao invés de em duas metades apenas, que considere mais amiudamente as mudanças progressivas no comportamento das fixações oculares, possa nos elucidar melhor a questão.

Fenômeno semelhante ao observado para a extensão da aula também se verificou no interior do tempo de exibição dos slides ao considerarmos as duas metades (**Gráfico 40**): na média, os participantes dedicaram a maior parte do tempo inicial de apresentação do slide na varredura por informações visuais contidas na região do slide, dedicando mais intensamente atenção visual à região da professora apenas na segunda metade (discutiremos mais em detalhe as metades ao analisarmos mais adiante as condições experimentais). Tais resultados sugerem competição atencional visual entre as duas áreas de interesse e entre os canais visual e auditivo; a maior porcentagem do tempo de fixação na região do slide durante momentos iniciais de exibição do slide, bem como nas sequências iniciais da aula como um todo, sugere que motivações exógenas, como alterações dos estímulos visuais, e endógenas, tais busca por novas informações, levem vantagem atencional inicial sobre necessidades atencionais de atendimento social e de sincronização dos canais visual e auditivo para direcionamento da atenção à origem emissora do conteúdo da fala, o rosto da professora. Vale notar que a construção semiótica

pela fala é linear, dependente do decurso do tempo; enquanto que informações visuais contidas nos slides, ainda que também operem por unidades analíticas discretas, trazem-nas imediatamente disponíveis, distribuídas e organizadas espacialmente. Note-se também que a competição atencional entre as duas regiões do campo visual pode levar à competição entre os canais visual e auditivo se não houver sincronia entre o conteúdo da fala da professora e o conteúdo visualizado no slide pelos sujeitos, já que ambos estímulos mobilizam unidades linguísticas discretas em sua composição semiótica. Especulamos, contudo, a competição atencional entre as regiões e o atendimento social se darem de forma diferente fosse a aula interativa e presencial.

Perguntas que decorrem dessas observações são: i) qual o ganho (ou prejuízo) de compreensão advindo dessa competição atencional entre diferentes regiões do campo visual por movitações exógenas e endógenas; e ii) se não seria mais vantajoso do ponto de vista da aprendizagem e mais econômico em recursos cognitivos atencionais um sacrifício temporal inicial para reconhecimento e seleção de informações relevantes pelos aprendizes na apresentação de um novo slide antes do início da elocução pela professora. Por outro lado, qual seria o impacto da privação, não da professora ou de seus gestos como proposto em nosso experimento, mas da imagem do slide em momentos iniciais da aula (ou atraso de sua entrada ao início e nas transições entre diferentes slides) sobre a atenção visual e a aprendizagem dos sujeitos? Questões estas que imaginamos sejam possíveis de serem respondidas por estudos experimentais futuros.

Ainda sobre o comportamento geral da movimentação ocular dos participantes, observamos pulsos temporais com picos de tempo de fixação na região do slide; chamamos aqui de pulsos as fortes altas observadas na porcentagem de tempo de fixação na região do slide para os slides 2, 4, 6, 12 e 19 (**Gráfico 27**). Trata-se de slides esquemáticos introdutórios (ou terminal no caso do slide 19 com as referências bibliográficas), de menor extensão (duração média de 16,7 segundos, $\pm 2,6$), com baixa ocorrência de períodos (média de 1,8, $\pm 0,45$; sendo a média geral entre todos os slides de 5,7, $\pm 4,7$; **Gráfico 14b**) e baixa ocorrência de gestos (média de 11 gestos, $\pm 4,6$; para uma média geral entre todos slides de 50,6, $\pm 43,4$; **Gráfico 17**); slides, portanto, com baixo valor informacional, baixa complexidade de explanação e, por consequência, baixa dificuldade de compreensão (**Gráfico 21**). Como explicar, então, a região do slide receber próximo a 80% do tempo de fixação ocular durante sua exibição? Nossa interpretação é de que, apesar da baixa complexidade e valor informativo dos slides em questão, a mudança qualitativa informacional pela disruptura no padrão visual em transição de slide

mobiliza para a região o interesse e a atenção visual dos sujeitos; ademais, a curta duração dos slides em questão, conforme discutido anteriormente, não permite aos sujeitos expectadores o tempo necessário para a entrada em curso de mecanismos atencionais de atendimento social à professora palestrante; estes ocorrendo mais tardiamente no tempo de exibição do slide e no decurso da aula como um todo.

O fator temporal nos induz ainda a explicarmos também a alta porcentagem de fixação ocular observada na região da professora durante a exibição do slide gráfico de número 18, quando analisado em comparação com as etapas didáticas anteriores (**Gráficos 17 e 27**): o slide 18 é o mais duradouro de toda a aula, com 143,5 segundos de duração, enquanto que a média de duração dos slides é de 48,3 segundos. Entretanto, para a etapa imediatamente anterior, slide esquemático de número 17 com segunda maior duração (121,3 segundos), não observamos a mesma alta de fixação na região da professora; ao contrário, a região da professora registrou 41%, enquanto que a região do slide recebeu os outros 59% de tempo de fixação pelos voluntários (**Gráfico 27**). O decurso temporal explica, portanto, apenas em parte o aumento do tempo de fixação dedicado à região da professora, esta recebendo maior proporção de tempo de fixação ocular de forma progressiva ao longo da aula e na segunda metade do tempo de exibição de cada slide. Porém, como apontado acima, o tempo de exibição não explica tudo; observamos diferenças no comportamento atencional visual dos voluntários entre os dois slides mais longos (17 e 18; **Gráfico 27**), quando se esperaria, pelo raciocínio que vínhamos desenvolvendo, um comportamento atencional semelhante acompanhando a mesma tendência. Outra observação que limita o poder de explicação temporal para o aumento de tempo de fixação observado na região da professora é o fato de, assim como o slide 18, também os slides 10 e 15 apresentarem maior proporção de tempo de fixação na região da professora (ambos slides com 56% para a professora e 44% para a região do slide; **Gráfico 27**), tendo eles, no entanto, curta duração: apenas 76 e 56 segundos, respectivamente (**Gráfico 17**).

Tais observações sugerem haver fatores outros para além do fator temporal interferindo na alocação explícita da atenção dos sujeitos experimentais na região da professora. Seriam tais fatores aspectos intrínsecos aos slides, de maior ou menor natureza informacional, capazes de aumentar ou diminuir o interesse pela região? Estariam antes relacionados à necessidade de sincronização dos estímulos visual e sonoro pelo interesse dos sujeitos expectadores no conteúdo da fala da professora? Ou seriam aspectos sociais de atendimento social à professora palestrante? Por que não a combinação de todos eles

e, se sim, em que proporção? A análise das fixações oculares que considera as duas condições experimentais traz a seguir informações importantes sobre a questão.

ENTRE AS CONDIÇÕES EXPERIMENTAIS

Ao longo da aula (**Gráfico 31**), observamos entre as condições experimentais consistência nas oscilações de aumento e diminuição na porcentagem de tempo de fixação ocular nas regiões da professora e do slide, sugerindo uma tendência geral quanto ao interesse pelas áreas analisadas durante a aula. Dessa forma, em ambas condições experimentais, as maiores altas de porcentagem de fixação na região do slide foram observadas nas seções introdutória, esquemática e das referências bibliográficas; e na região da professora, nas sequências 1 e 2 e na etapa gráfica. Contudo, enquanto que na condição *sem gestos* a região do slide é privilegiada recebendo em todas as etapas didáticas maior porcentagem de tempo de fixação ocular em comparação com a região da professora, para a condição *com gestos*, pelas oscilações ocorrerem com valores mais aproximados entre as duas regiões, observou-se a alternância na prevalência por uma ou outra região. Destacamos as maiores diferenças entre as condições experimentais com maior aumento da porcentagem de tempo de fixação ocular na região da professora na condição *com gestos* para a sequência 2 e para a sequência gráfica.

Pela análise da sequência de slides (**Gráfico 32**), também observamos os fenômenos acima descritos, podendo notar as maiores diferenças nos padrões de fixação entre as condições experimentais a partir da segunda sequência didática, mais acentuadamente para os slides 13, 14 e 16 da sequência didática 2, e slide 18 da sequência gráfica. Nessa sequência de slides, podemos observar mais nitidamente o padrão de fixação tendendo a equilibrar a distribuição das fixações entre as áreas da professora e do slide na condição *com gestos*, e tendendo a um desequilíbrio privilegiando a área do slide na condição *sem gestos*. Essa diferença entre as condições experimentais se extinguirá pontualmente durante o último slide com as referências bibliográficas, ambas as condições experimentais registrando aumento drástico de porcentagem de fixação na região do slide.

Observamos, portanto, que, apesar de terem atraído para si apenas 2,9% ($\pm 0,5$) do tempo total de fixação ocular dos participantes (em mesma direção de resultados anteriores; Pereira & Haddad, 2015, Gullberg & Kita, 2009), os gestos contribuíram para

um aumento de 8,8% na porcentagem média de tempo de fixação ocular na região da professora (42,2%, $\pm 2,8$, para a condição *sem gestos* e 51%, $\pm 2,7$, para a condição *com gestos*; ANOVA $F_{(1,52)} = 4,992$; $p = 0,03$; **Gráfico 29**). A exibição dos gestos foi, assim, responsável por equilibrar o tempo de fixação ocular na região da professora com o tempo de fixação ocular na região do slide; uma vez ausentes, a região da professora perdeu em interesse, os sujeitos passando a privilegiar a região do slide.

Tamanho impacto no comportamento atencional visual dos sujeitos com reduzido tempo de fixação atraído para si reforça nossa hipótese inicial de que gestos sejam processados na maior parte do tempo por visão periférica. O fato de a projeção da gesticulação manual da professora recair como regra geral fora da região da fóvea na retina do sujeito expectador implica em que gestos sejam percebidos como movimento e processados com baixa resolução espacial preferencialmente pela via magnocelular. Poderíamos contra-argumentar nossos resultados apontando que Kelly et al. (2012), como já referido neste trabalho na seção sobre as vias do sistema visual, obtiveram como resultado em seu estudo experimental a melhora no processamento de gestos quando a imagem da gesticulação fora apresentada filtrada para alta frequência, resolução cujo processamento é otimizado pela atuação do sistema parvocelular. No entanto, é importante esclarecer que, assim como no desenho experimental de outros trabalhos (por exemplo, Tellier 2008), no estudo de Kelly et al. (ibid.) o estímulo apresentado ao participante consistiu apenas da gesticulação manual; ou seja, isolaram os gestos ao eliminarem da imagem apresentada como estímulo o rosto do falante. Ora, não sendo possível o atendimento social à região facial, é certo que a atenção visual do sujeito expectador se dirija ao único estímulo presente relevante: a gesticulação manual. Uma vez a fixação ocular forçosamente dirigida aos gestos, e sua projeção recaindo na região da fóvea, é esperado que a filtragem da imagem para alta frequência melhore o seu processamento que, nessas condições, se faz prioritariamente pela via parvocelular. Em nosso presente estudo, contudo, apresentamos aos sujeitos participantes a imagem integral da metade superior do corpo da professora; nessas condições, o atendimento atencional se dá prioritariamente à face, sobretudo aos olhos do falante e raramente aos gestos (ver também Pereira & Haddad, 2015; Kendon, 1990; Rimé et al., 1988; Argyle & Cook, 1976). Seria interessante para elucidação da questão a replicação de nosso estudo controlando-se o grau de deterioração da imagem do campo gestual durante a exibição da aula.

Ainda sobre o tempo de fixação ocular na região dos gestos, esperávamos como resultado a gesticulação icônica como principal categoria a atrair para a região as fixações oculares dos sujeitos participantes. No entanto, os resultados das análises correlacionais e de regressão linear que exploram os dados apresentados nos **gráficos 35, 36 e 37** apontaram os gestos metafóricos como preditores da fixação ocular na região dos gestos, seja ao considerarmos o número de ocorrências (Spearman $\rho = 0,651$; $p = 0,003$), seja ao considerarmos sua frequência por segundo de fala ($\rho = 0,650$; $p = 0,003$) – gestos dêiticos e de pulso apresentaram correlação positiva marginal e apenas para o número de ocorrências (respectivamente, $\rho = 0,467$, $p = 0,044$; e $\rho = 0,447$, $p = 0,055$). A direção apontada pelo resultado das análises correlacionais foi reafirmada pela análise de regressão linear múltipla subsequente, que resultou em um modelo estatisticamente significativo ($F_{(1,17)} = 6,993$; $p = 0,017$; $R^2 = 0,291$), apontando o número de ocorrências de gestos metafóricos ($\beta = 0,540$; $t = 2,645$; $p = 0,017$), bem como de sua frequência de ocorrência por segundo de fala ($F_{(1,17)} = 9,095$; $p = 0,008$; $R^2 = 0,349$; $\beta = 0,590$; $t = 3,016$; $p = 0,008$), como preditor do tempo de fixação ocular na região dos gestos.

Além de preditora do tempo de fixação na região dos gestos, a frequência da gesticulação metafórica também esteve correlacionada positivamente com as fixações na região da professora na condição *com gestos* ($\rho = 0,530$, $p = 0,020$), a frequência da gesticulação icônica ficando à margem da significância ($\rho = 0,447$; $p = 0,055$). Já ao considerarmos o número de ocorrências, verificamos correlação positiva entre as fixações na região da professora e: o número de ocorrências de gestos metafóricos ($\rho = 0,688$; $p = 0,001$), de pulso ($\rho = 0,602$; $p = 0,006$) e icônicos ($\rho = 0,495$; $p = 0,031$). Uma análise de regressão linear múltipla resultou em um modelo estatisticamente significativo ($F_{(1,17)} = 16,321$; $p = 0,001$; $R^2 = 0,490$), apontando a frequência da gesticulação metafórica ($\beta = 0,700$; $t = 4,040$; $p = 0,001$) como preditora da fixação ocular na região da professora na condição *com gestos*.

Em suma até aqui, verificamos a gesticulação metafórica como forte preditora das fixações em ambas regiões dos gestos e da professora.

A alocação da atenção visual no campo gestual da professora se comportou na média entre os participantes ($n = 28$) da seguinte forma: com menor valor de 1,6% para a sequência introdutória, 2,5% para a sequência 1, 2,6% para a sequência 2, 3,3% para a sequência esquemática, até atingir máxima de 4,1% na última sequência gráfica, antes de cair significativamente para 0,5% para as referências bibliográficas (**Gráfico 30**). Esse aumento progressivo da porcentagem de tempo dedicado à fixação dos gestos ao longo

da aula, observado também na análise por slide (**Gráfico 35**), é um primeiro indicativo de uma possível correlação positiva entre o tempo de fixação na região dos gestos e o tempo decorrido de aula; corrobora essa interpretação o resultado obtido da análise da fixação ocular dos participantes na região dos gestos para as metades do tempo de exibição dos slides: média de 1,7% ($\pm 0,3$) de tempo de fixação na região dos gestos para a 1ª metade, e 2,8% ($\pm 0,4$) para a 2ª metade (**Gráfico 38**).

O interesse pela região dos gestos acompanha, portanto, mesma tendência observada para a região da professora: também esta, ao analisarmos as duas metades do tempo de exibição dos slides, apresentou médias de porcentagem de tempo de fixação ocular maiores para a segunda metade do tempo de exibição dos slides (ANOVA, $F_{(1,18)} = 72,371$; $p < 0,001$; **Gráfico 40**) em ambas condições experimentais: na condição *sem gestos*, 26% ($\pm 3\%$) para a primeira metade e 46,2% ($\pm 5\%$) para a 2ª metade; enquanto que na condição *com gestos*, 33% ($\pm 3\%$) para a 1ª metade e 55% ($\pm 5\%$) para a 2ª metade (**Gráfico 41**). Note-se que esse retardamento do atendimento social à região da professora palestrante (inclusa a região dos gestos) observado tanto através da linha de tendência ao longo da aula (**Gráficos 27 e 36**) como através das análises das metades (**Gráfico 40 e 41**), é acompanhado por uma tendência geral de aumento progressivo da produção de gestos ao longo da aula (**Gráfico 17**). Novamente, perguntamo-nos sobre um possível aumento da frequência de gestos da 1ª para a 2ª metade dos slides, aumento com poder potencialmente explicativo do fenômeno em questão.

Tais resultados, corroboram mais uma vez a hipótese sugerida pela linha de tendência observada no **gráfico 26** já discutida, de que a área da professora no campo visual como área de interesse atencional é, como um todo (inclusa a área de gesticulação), dependente do fator tempo de apresentação do slide e do tempo decorrente da aula; sugerem também, por outro lado, uma correlação positiva entre o aumento progressivo na produção de gestos e o aumento do interesse visual pela área da professora verificado na alta progressiva da porcentagem de tempo de fixação ocular a ela dedicada ao longo da aula. Reforça-se, assim também, nessa linha de raciocínio, a hipótese de que o estímulo visual do slide seria mais atrativo atencionalmente nos momentos iniciais da apresentação de cada slide e da aula como um todo. Durante a exploração das informações visuais do slide, a atenção seria dirigida à região da professora com certa parcimônia. A região da professora passaria a receber maior interesse e maior tempo de fixação ocular somente após as mudanças qualitativas do estímulo visual do slide terem sido atendidas e esgotada nele a exploração por novas informações; ou, eventualmente, quando seja este superado

em interesse atencional pela construção semântico-discursiva fala-gesto da professora.

Perguntamo-nos a esse respeito se o deslocamento do interesse atencional e fixação na região dos gestos, para além do fator temporal, ocorreriam de forma semelhante ao que especulamos se dar para a região do slide em momentos de ruptura no padrão visual, por exemplo como o observado para os slides 2, 4, 6, 12 e 19 (**Gráfico 27**); ou seja, seria a atração da atenção visual dos sujeitos para a região dos gestos da professora motivada por rupturas na continuidade do estímulo visual por mudanças no padrão de movimentação das mãos, por exemplo por rupturas na qualidade, amplitude ou ritmo da gesticulação manual? Perguntamo-nos ainda se a fixação na região dos gestos poderia ocorrer diretamente a partir de uma fixação na região do slide, ou se dependeria necessariamente de uma fixação prévia (e um engajamento atencional social) na região da professora – ou ainda, eventualmente, de um direcionamento atencional aos próprios gestos realizado pela própria professora, como sugerem resultados de Gullberg & Kyta (2009). Considerando-se a possibilidade de sacadas e miradas na região dos gestos, não por razões de atendimento social de compartilhamento do foco atencional, mas por mecanismos neurocognitivos de níveis mais baixos disparados por mudanças significativas na qualidade do estímulo visual, como operariam, nestas circunstâncias, mecanismos de habituação sensorial às mudanças no estímulo visual provenientes do campo gestual da professora? Quais seriam os limiares de amplitude e velocidade angular da gesticulação capazes de atrair para si o foco atencional dos participantes? Sofreriam esses limiares interferência do conteúdo linguístico? Ou, seja, seriam eles, os limiares, absolutos, bastando-se para se verificar alocação explícita da atenção sobre os gestos ou seriam também modulados por conteúdo discursivo da fala? São questões relevantes capazes de serem respondidas em futuros estudos com controles experimentais adequados.

Vale retornarmos agora ao nosso esforço por explicação para os padrões distintos de fixação observados para os dois slides mais longos da sequência didática, os slides 17 e 18. Consideremos por ora o slide 18. Notemos em primeiro lugar que o slide 18 se assemelha aos slides 8, 9, 10, 13 e 15 tanto em relação à alta porcentagem de tempo de fixação na região da professora (**Gráfico 27 e 33**), quanto na rica composição de gestos metafóricos e icônicos (**Gráfico 17**). De fato, conforme já mencionado, uma análise correlacional de Spearman mostrou (**Gráfico 35**) correlação estatisticamente significativa entre a porcentagem de fixação na região dos gestos e a frequência de ocorrência de gesticulação metafórica ($n = 19$; Spearman $\rho = 0,650$; $p = 0,003$); uma análise de

regressão linear múltipla também revelou em um modelo estatisticamente significativo ($F_{(1,17)} = 6,993$; $p = 0,017$; $R^2 = 0,291$) o número de ocorrências de gestos metafóricos ($\beta = 0,540$; $t = 2,645$; $p = 0,017$) ser preditor do tempo de fixação ocular na região dos gestos. Na condição *sem gestos*, curiosamente, as fixações na região da professora apresentaram correlação significativa com a frequência de gesticulação icônica ($n = 19$; Spearman $\rho = 0,500$; $p = 0,029$) e correlação apenas marginal com a gesticulação metafórica ($n = 19$; Spearman $\rho = 0,450$; $p = 0,053$). Evidentemente, não seria possível atribuir aos gestos a fixação ocular na região da professora na condição *sem gestos*; contudo, o fato de o tempo de fixação ocular na região da professora se correlacionar mais acentuadamente com categorias distintas de gestos (gestos metafóricos na condição *com gestos* e gestos icônicos na condição *sem gestos*) sugere que o direcionamento atencional visual à região da professora atenda a conteúdos linguísticos distintos entre as duas condições experimentais, observados em diferenças de conteúdo gesticulatório; dito em outras palavras, sujeitos tenderiam a olhar para a região da professora em passagens distintas quanto à elaboração linguística quando o fazem na presença ou na ausência de gestos.

A partir de tais resultados especulamos que a atenção visual dos participantes seja dirigida à região da professora para sincronia dos canais sonoro e visual mas também, assim como à região dos gestos, por atendimento social – e não para extração de informação visual, como hipotetizamos inicialmente sobre as fixações na região dos gestos. A ausência dos gestos na condição *sem gestos*, por sua vez, além de diminuir o tempo de fixação na região da professora, parece influir também na estratégia de seleção da ocasião da fixação, mais frequentes agora em passagens discursivas acompanhadas de gesticulação icônica (ao invés de metafórica, como quando na presença dos gestos), alterando (talvez prejudicando) a sincronização dos canais sonoro e visual ou o atendimento social de compartilhamento da atenção. Para nossa surpresa, contudo, análises de regressão linear múltipla resultaram em modelos estatisticamente significativos apontando a frequência da gesticulação metafórica como preditora do tempo de fixação ocular na região da professora tanto para a condição *com gestos* ($F_{(1,17)} = 16,321$; $p = 0,001$; $R^2 = 0,490$), ($\beta = 0,700$; $t = 4,040$; $p = 0,001$), quanto para a condição *sem gestos* ($F_{(1,17)} = 9,861$; $p = 0,006$; $R^2 = 0,367$), ($\beta = 0,606$; $t = 3,140$; $p = 0,006$). Esse resultado, observado para ambas condições experimentais, a despeito dos resultados das análises correlacionais que diferem ligeiramente entre as condições experimentais, sinaliza que haja algo no disurso em passagens acompanhadas de gestos metafóricos (possivelmente conteúdo linguístico mais abstrato), influenciando fortemente – e para além

da visualização dos gestos – no padrão de fixação ocular registrado para a região da professora. Do ponto de vista da composição linguístico-discursiva, o que teriam os gestos metafóricos de diferente dos demais tipos de gestos? Gestos metafóricos estão a serviço do esforço de explanação de conteúdo de caráter mais abstrato; são carregados em acentuação prosódica (própria da gesticulação de pulso) e em expressividade do esforço de elaboração linguística da fala. Gestos metafóricos distinguem-se de gestos icônicos apenas quanto ao conteúdo linguístico que os acompanha; ou seja, não há diferenças intrínsecas e necessárias entre eles quanto ao seu aspecto morfológico.

Nossos resultados apontam em mesma direção de resultados da literatura (Gullberg & Kita, 2009) que mostram que fixações nos gestos tenham caráter de atendimento social de atenção compartilhada, sobretudo dirigido pela orientação atencional do próprio falante (professora gesticuladora, em nosso experimento), e não por necessidade de extração de informação visual. Poderíamos, portanto, atribuir por ora a essa composição gestual metafórico-icônica um fator atuante nas altas observadas nos valores registrados para a região da professora. Salientamos, contudo, que a alta na porcentagem do tempo de fixação na região da professora para os slides em questão (slides 8, 9, 10, 13, 15 e 18) é observada em ambas condições experimentais, o que nos sugere que aspectos gesticulatórios metafórico-icônicos, além de atrair a atenção visual dos participantes para a região dos gestos ou da professora como um todo na condição *com gestos*, sejam índices importantes, dentre outros índices linguísticos, de interesse atencional discursivo.

Porém, se por um lado a ocorrência de gestos metafóricos e icônicos nos ajudam a explicar a alta no tempo de fixação registrada para os slides 8, 9, 10, 13, 15 e 18, por outro, poderíamos contra-argumentar que semelhante ocorrência de gestos ricos em aspectos metafóricos e icônicos observada para o slide 17 (**Gráfico 17**) não produziu mesmo efeito (**Gráficos 27, 32 e 36**), muito embora difira daqueles por menor frequência de gesticulação icônica por segundo de fala (**Gráfico 18**). Parece-nos, portanto, que na passagem da sequência didática esquemática (slide 17) o interesse visual esteja de fato no slide, interesse que é vencido pela região da professora somente na segunda metade de tempo de exibição do slide na condição *com gestos*, fenômeno verificado de maneira mais acentuada também para o slide 16 (**Gráfico 40**).

Pelo que elaboramos até aqui, para bem além de apenas uma explicação, parece haver uma série de fatores interferindo na alocação da atenção visual durante a exibição da aula, tais fatores de tempo de exibição do slide e de decurso da aula, presença de

aspectos visuais relevantes do ponto de vista informacional, grau de abstração e conteúdo informacional também da fala, bem como a presença e frequência de gesticulação metafórica e icônica, capazes de atrair para a região da professora a atenção visual do sujeito expectador.

Há ainda um outro fator não controlado em nosso experimento mas que deveria ser levado em conta, ainda que de maneira especulativa: assim como especulamos sobre o engajamento da professora palestrante em seu esforço de apresentação e desenvolvimento do conteúdo a ser ensinado, que pelo próprio desenvolvimento linguístico-discursivo através da fala pode em si despertar interesse e atrair a atenção visual, é preciso considerarmos o engajamento do próprio sujeito expectador. O interesse e o engajamento na tarefa de compreensão e aprendizagem do conteúdo pelos sujeitos aprendizes podem também eles ter influído no direcionamento (ou não-direcionamento) de sua atenção visual para a região da professora. Haveríamos de verificar se fatores de empatia social relacionados à professora palestrante, como sexo, idade, vestuário, tom da voz, bem como o conhecimento prévio e presencial de sua pessoa pelos estudantes pudessem interferir no interesse visual pela região da professora durante a exibição da aula e no engajamento dos sujeitos experimentais na realização da tarefa, ou seja, em seu empenho de compreensão e aprendizagem do conteúdo ensinado. É possível que o contato físico e o conhecimento presencial prévios da professora palestrante pelo sujeito aprendiz venha a ter impacto significativo em seu engajamento atencional e aprendizagem à distância, como na modalidade de aula *online*, por exemplo. Nesse sentido, esperaríamos maiores diferenças de atendimento atencional visual entre as modalidades presencial e online do que entre diferenças informacionais entre slides.

Vimos ainda pela análise da duração média das fixações oculares (**Gráfico 24**) que ela diferiu qualitativamente entre as áreas de interesse: as fixações oculares na região da professora foram duas vezes mais duradouras do que as fixações na região do slide, com média de 0,49 segundos ($\pm 0,02$) para a região da professora, e 0,23 segundos ($\pm 0,01$) para a região do slide (ANOVA; $F_{(1,112)} = 114,959$; $p < 0,001$). Tal resultado revela duas naturezas distintas de comportamento visual atencional: uma com com fixações mais breves associada a uma estratégia de varredura do olhar por informação visual na região do slide, e outra com fixações mais prolongadas associada antes à necessidade de atendimento social na região da professora – estudos anteriores mostram preferência pela região dos olhos (Pereira & Haddad, 2015; ver também Gullberg & Kita, 2009). Explicitam-se, portanto, ao menos duas naturezas atencionais distintas de fixação ocular:

uma de caráter social e outra de caráter informacional. Ora, se as fixações oculares na região da professora têm caráter de atendimento social, qual seria então a relevância dessas fixações para a aprendizagem dos conteúdos apresentados, considerando sobretudo nosso contexto experimental de exibição de uma aula gravada, cuja própria natureza elimina a possibilidade de interação social (ao menos o *feedback* atencional) como seria esperada fosse uma aula presencial ou virtual mas em tempo real? Seria a imagem da professora dispensável? Seriam os gestos atratores atencionais para uma região não-necessária (a região da professora) à apreensão do conteúdo ensinado, contrariando, assim, resultados de trabalhos anteriores que corroboram a hipótese comunicativa dos gestos? Por outro lado, seria o atendimento social atencional dirigido à região da face e dos olhos tão inerente à comunicação humana ao ponto de sobreexistir às distâncias espacial, temporal e mesmo interacional impostas pelas condições de uma aula previamente gravada? Qual o impacto da presença da imagem da professora na acomodação do conhecimento nas memórias de curto e longo prazos? Essas são algumas das perguntas que suscitaram o conjunto de nossos resultados à luz de reflexões de estudos anteriores da área.

Ainda sobre a duração média das fixações oculares, não encontramos diferenças significativas entre as condições experimentais *com* e *sem gestos* (**Gráfico 33**) nem para a região da professora, nem para a região do slide (ANOVA, $F_{(1,55)} = 2,295$, $p = 0,136$, para a região da professora; e Mann-Whitney, $Z = -1,699$, $p = 0,089$, para a região do slide). Registramos, no entanto, tempo médio da fixação nas regiões da professora e do slide para a condição *sem gestos* aproximadamente 15% superiores em relação aos valores registrados para a condição *com gestos*: para a região da professora observamos 0,46 segundos ($\pm 0,03$) para a condição experimental *com gestos*, e 0,53 segundos ($\pm 0,04$) na condição experimental *sem gestos*; para a região do slide, observamos 0,21 segundos ($\pm 0,01$) para a condição *com gestos*, e 0,24 segundos ($\pm 0,01$) para a condição *sem gestos*. A análise da sequência de slides (**Gráfico 34**), no entanto, reforça claramente a natureza distinta da fixação ocular entre as regiões da professora e do slide em ambas condições experimentais, bem como mostra valores significativamente superiores com grandes oscilações e desvios padrões de maior amplitude para a condição *sem gestos*, tanto para a região da professora (Wilcoxon Signed Ranks; $Z = -3,743$; $p < 0,001$), quanto para a região do slide (Wilcoxon; $Z = -3,219$; $p = 0,001$). Especulamos a esse respeito se a estratégia cognitiva de busca visual na condição *sem gestos* sofreria algum tipo de instabilidade, revelando, por exemplo, algum grau de dificuldade por parte dos sujeitos

de definirem qual das áreas fixarem, em que momento e por quanto tempo, bem como de definirem a melhor estratégia de seleção e organização das informações visuais acessadas em ambas regiões do slide e da professora. Perguntamo-nos ainda, a partir de tais resultados, se a ausência de gesticulação na condição *sem gestos* possa ter causado algum prejuízo no atendimento social dos participantes à professora palestrante, dificultando ou diminuindo a qualidade da percepção da respectiva região por visão periférica, resultando, por exemplo, em fixações demasiadamente prolongadas – é o que sugerem diferenças observadas no tempo médio de fixação na região da professora, com valores bastante discrepantes entre as condições experimentais *com* e *sem gestos* (Wilcoxon Signed Ranks; $Z = -3,743$; $p < 0,001$; **Gráfico 34**). Uma análise do número de fixações oculares e visitas às respectivas áreas poderia trazer maiores elucidações à questão.

Ainda buscamos por análise qualitativa do conteúdo informacional contido em cada slide bem como na respectiva fala da professora possíveis razões para explicarmos por que as diferenças no tempo médio de fixação na região da professora variam mais para determinados slides, como os slides 5, 7, 15, 16 e 17, e menos para outros, como os slides 2, 4 e 6 (**Gráfico 34**). Parte da explicação acreditamos residir em seu maior ou menor valor informacional (valor introdutório ou de desenvolvimento, por exemplo), como já discutido anteriormente. Também desafiam nossa compreensão fortes picos no tempo médio de fixação observados, por exemplo, na condição *sem gestos* na região da professora para o slide 5, contendo como conteúdo um exemplo com uma imagem representando o padrão fixo de ação, bem como na região do slide para o slide de número 9, contendo um exemplo com imagem e texto de comportamento rigidamente programado, ocorrências para as quais ainda buscamos possível explicação.

Quanto ao tempo médio de fixação ocular na região dos gestos (0,29 segundos, $\pm 0,01$; **Gráfico 24 e 34**), este aproximou-se do tempo médio de fixação na região do slide, sugerindo por sua brevidade também uma natureza de fixação ocular de varredura por informação visual; contudo, seu desvio padrão difere qualitativamente do desvio padrão das fixações na região do slide: as fixações na região do slide mostram uma natureza bem definida com pequena variância (EP 0,01s), enquanto que o erro padrão do tempo médio de fixação na região dos gestos (0,03s) se aproxima daquele verificado para a região da professora. Tal resultado nos leva a pensar que a atenção poderia ser dirigida à região dos gestos tanto por razões exógenas de necessidades neurológicas de níveis mais baixos despertadas pelo contraste no padrão e amplitude de movimentos gestuais, quanto por necessidades neurocognitivas de níveis mais altos de atendimento atencional social

dirigido (ou não) pela própria professora palestrante para seu campo gestual à sua frente; uma vez varrida a área por informação visual relevante ou verificada a ausência desta, a fixação retornaria imediatamente ao centro de interesse atencional da região da professora, a região dos olhos, por continuidade ao atendimento atencional social.

Pela análise do tempo médio de fixação ocular considerado para cada uma das duas metades do tempo de apresentação dos slides, pudemos ainda flagrar indícios de mudanças comportamentais dependentes do tempo também para a duração média da fixação ocular. Uma ANOVA mostrou efeito das metades sobre a duração média da fixação na região dos gestos ($F_{(1,17)} = 4,768$; $p = 0,043$); observamos maior duração média da fixação ocular para a 2ª metade dos slides: tempo médio de 0,24 segundos ($\pm 0,01$) para a 1ª metade, e 0,26 segundos ($\pm 0,02$) para a 2ª metade (**Gráfico 42**).

Já para as fixações na região da professora (**Gráfico 12**), observamos alteração significativa do tempo médio de fixação entre as duas metades apenas para a condição *sem gestos*: média de 0,46 segundos ($\pm 0,03$) para a 1ª metade, e 0,56 segundos ($\pm 0,03$) para a 2ª metade (este último sendo significativamente elevado sobretudo por fortes altas nos valores observados para os slides 5, 9, 11 e 16, passagens com picos de frequência de gesticulação metafórica; **Gráfico 18**); enquanto que, para a condição *com gestos*, observamos maior regularidade com média de 0,40 segundos ($\pm 0,02$) para a 1ª metade, e 0,42 segundos ($\pm 0,02$) para a 2ª metade. Ademais, uma ANOVA mostrou efeito das condições experimentais *com gestos* e *sem gestos* ($F_{(1,18)} = 71,028$; $p < 0,001$), efeito das metades ($F_{(1,18)} = 9,784$; $p = 0,006$), bem como efeito de interação entre os fatores condição experimental e metades ($F_{(1,18)} = 6,613$; $p = 0,019$) sobre a duração média da fixação ocular na região da professora.

Para a região do slide (**Gráfico 46**), também observamos valores superiores de tempo médio de fixação ocular para a 2ª metade, mais moderadamente para a condição *com gestos* (com tempo médio de fixação de 0,20 segundos, $\pm 0,00$, para a 1ª metade, e 0,21, $\pm 0,00$, para a 2ª metade), e mais acentuadamente na condição *sem gestos* (com tempo médio de fixação de 0,22 segundos, $\pm 0,01$, para a 1ª metade, e 0,26, $\pm 0,02$, para a 2ª metade, com altos picos registrados para a 2ª metade para os slides 9 e 11). Uma ANOVA mostrou efeito da condição experimental ($F_{(1,18)} = 10,169$; $p = 0,005$) e das metades ($F_{(1,18)} = 10,685$; $p = 0,004$); mostrou, porém, não haver efeito de interação entre os fatores ($F_{(1,18)} = 2,032$; $p = 0,171$). Ainda buscamos explicações para tais fenômenos.

Por fim, esperávamos encontrar diferenças comportamentais (como tempo de fixação nas regiões estudadas), bem como na aprendizagem entre sujeitos de diferentes

formações e graus de conhecimento prévio sobre assunto. Análises de variância mostraram, contudo, não haver efeito estatisticamente significativo do perfil do público tais idade, sexo, formação anterior e atual, e conhecimento prévio sobre o assunto sobre o tempo de fixação ocular nas regiões da professora, do slide ou de sua gesticulação. Encontramos, contudo e curiosamente, efeito do horário de realização do experimento ($F_{(2,50)} = 4,649$; $p = 0,014$): o tempo de fixação na região da professora para os sujeitos que realizaram a sessão experimental pela manhã foi em média de 57,5% ($\pm 0,04$; $n = 10$), distinguindo-se de ambas porcentagens dos sujeitos que realizaram a sessão experimental pela tarde (42,4%; $\pm 0,02$; $n = 33$) e pela noite (45,6%; $\pm 0,05$; $n = 10$), estas últimas não se distinguindo entre si. Apesar de pequena a população amostral, tal resultado sugere que sujeitos estejam mais sensíveis ou dispostos ao engajamento atencional social no período matutino do que nos períodos vespertino ou noturno. Uma questão importante que imediatamente se coloca na direção de nosso estudo exploratório é o quanto a diferença observada no tempo médio de fixação ocular na região da professora para o período matutino (em comparação com os períodos vespertino e noturno) se daria por influência da gesticulação manual; ou seja, estariam os sujeitos aprendizes no período matutino mais sensíveis a pistas atencionais sociais do que nos demais períodos do dia? Futuros estudos que explorem a movimentação ocular de sujeitos aprendizes em aula e controlem o período de realização da tarefa podem trazer pistas do ecaminhamento da questão.

APRENDIZAGEM

Sobre o questionário de avaliação da aprendizagem dos sujeitos, a análise das respostas objetivas apresentada no **gráfico 54** apontou um dos itens da questão 8 (item *c*) diferir dos demais itens (*a*, *b* e *d*) por grande volume de respostas erradas e baixo volume de respostas meio certas. O **gráfico 55**, que explorou a relação entre a ocorrência de gestos e o tempo de fixação para a região da professora e dos gestos para os slides respectivos às perguntas da questão 8 (slides 10, 13, 14 e 15, respectivamente relacionados aos itens *a*, *b*, *c* e *d* da questão 8), sugere o item *c* diferir dos demais itens da questão 8 pela ausência de gesticulação icônica e por uma pequena tendência de menor tempo de fixação na região da professora e na região dos gestos. Nossos dados não permitem, contudo, atribuir a

diferença na pontuação nem à diferença observada na ocorrência de gesticulação, nem à tendência de menor tempo de fixação na região da professora e dos gestos.

De maneira geral, a avaliação de aprendizagem aplicada aos sujeitos foi pouco exploratória dos conteúdos e exemplos apresentados em aula; com um número pequeno de questões, sendo algumas delas de baixa dificuldade (observamos efeito teto em mais de uma questão), não nos permitiu investigar com propriedade possíveis efeitos de estratégia de fixação ocular sobre a aprendizagem dos sujeitos. Pensamos que um questionário que explore mais extensa e objetivamente os conteúdos abordados em cada um dos slides da aula, por diferentes naturezas de questões (como foi nossa tentativa com questões objetivas e dissertativas), possa melhorar a qualidade do estudo e trazer pistas de possíveis encaminhamentos da questão. Também um desenho experimental que controle a extensão da aula e a complexidade de seu conteúdo, possa trazer pistas importantes sobre efeitos de estratégias cognitivas de atenção visual sobre a aprendizagem dos sujeitos.

DISCUSSÃO GERAL

Avaliamos, em nosso estudo experimental 1, o custo-benefício linguístico da gesticulação manual na produção de narrativas por aprendizes de inglês como língua estrangeira. Nossas análises mostraram que a componente motora da gesticulação é inerente ao ato de elocução. Contudo, a grande variância em nossos dados de tempo de gesticulação e de extensão narrativa sugere que indivíduos difiram entre si no uso de recursos cognitivos e estratégias linguístico-discursivas, como a gesticulação, em produções narrativas. Em nosso estudo, o fator gesticulação não interferiu significativamente no tempo de planejamento discursivo, verificado através das taxas de fala (que permaneceu constante entre as condições experimentais), de hesitações e de interrupções sintáticas; tampouco interferiu substancialmente na qualidade narrativa, uma vez que não encontramos diferença significativa entre as condições experimentais quanto ao número de substantivos diferentes e taxa de erros. Entretanto, produções narrativas significativamente mais longas na condição de gesticulação (38% mais longas do que na condição *sem gestos*), com significativamente maior *taxa de repetições* por minuto de fala, levam-nos a compreender a produção narrativa em sua condição natural de fala-gesto como uma produção com redundância informacional. Esse aspecto elocucional pouco econômico é, contudo, como mostraram trabalhos anteriores (Pereira & Haddad, 2015, e Beattie & Shovelton, 1999), compensado socialmente pelo ganho em compreensão por um sujeito ouvinte-expectador quando a este lhe é dado acesso visual aos gestos do narrador em situação de interlocução. Ademais, a tendência a uma maior taxa de autocorrekções na presença de gestos sugere que, em condições naturais, estes favoreçam o engajamento e a consciência do sujeito sobre sua própria produção linguístico-discursiva.

Em nosso desenho experimental, buscamos um controle mais ecológico do fator gesticulação, com o cuidado de não introduzir uma segunda tarefa aos sujeitos experimentais ao evitarmos a instrução à não-gesticulação. Entretanto, não controlamos entre as condições experimentais um possível fator de estresse, possivelmente introduzido pelo recurso utilizado para controle dos gestos na condição *sem gestos*. Nossa análise da variabilidade de substantivos sugeriu, contudo, que nosso controle experimental não tenha interferido na carga de estresse durante a realização da tarefa de produção narrativa.

Em nosso experimento 2, exploramos o recrutamento e o impacto da gesticulação pela professora palestrante na alocação da atenção visual e na aprendizagem de alunos em uma aula de ciências. A análise da composição linguístico-discursiva da professora palestrante revelou frequência de acentuação prosódica da fala 2,6 vezes maior do que a frequência de golpes gesticulatórios, sugerindo gestos marcarem unidades discursivas de maior extensão e semanticamente mais abrangentes. Diferentemente do esperado, verificamos baixas na frequência de gestos de pulso em slides introdutórios e final, de caráter mais teórico, de baixa complexidade de compreensão e de menor duração; de maneira geral, observou-se maior ocorrência de gestos associados ao engajamento e esforço cognitivo explanatório em passagens com conteúdo de maior extensão e complexidade. Gestos icônicos apresentaram baixa frequência de ocorrência durante a aula e estiveram positivamente correlacionados com a duração dos slides, sugerindo maior ocorrência em passagens da aula com maior duração temporal, complexidade e desenvolvimento semântico.

Quanto à estratégia de fixação ocular, de maneira geral, os sujeitos deram maior preferência à região do slide (53,5%; $\pm$ 15,2%) em comparação com a região da professora (46%; $\pm$ 15,2%), levando-se em consideração o tempo de fixação ocular registrado. Observamos uma tendência inicial de maior tempo de fixação na região do slide e de aumento progressivo de interesse pela região da professora no decorrer da aula, os valores se alternando entre as duas regiões uma vez alcançado o equilíbrio entre elas. Semelhante tendência de aumento progressivo de interesse pela região da professora no decurso do tempo foi também verificada internamente aos slides através da análise das duas metades do tempo de sua exibição.

A região dos gestos (condição *com gestos*) recebeu apenas 2,9% do tempo de fixação ocular dos participantes, com aumento progressivo de interesse no decorrer do tempo de aula e entre as duas metades, semelhante ao verificado para a região da professora. Contudo, a presença da gesticulação manual durante o desenvolvimento do conteúdo da aula foi responsável por um aumento de 8,8% no tempo de fixação na região da professora, equilibrando, assim, a distribuição do tempo de fixação entre a região da professora e do slide na condição *com gestos*. Tais resultados reforçam nossa hipótese de que gestos sejam fortes atratores atencionais percebidos como movimento por visão periférica e processados com baixa resolução espacial e majoritariamente por via magnocelular.

Análises correlacionais e de regressão linear apontaram a frequência de gestos metafóricos como preditora da fixação ocular em ambas regiões dos gestos e da professora. A frequência de gestos icônicos ficou à margem da significância de correlação positiva com a fixação na região da professora.

Análises da duração média da fixação ocular mostraram que esta diferiu qualitativamente entre as áreas de interesse, com fixações duas vezes mais duradouras na região da professora do que na região do slide, com média de 0,49 segundos ($\pm 18,3$) para a região da professora, e 0,23 segundos ($\pm 4,2$) para a região do slide, sugerindo naturezas distintas de comportamento visual atencional: fixações mais breves no slide, provavelmente em estratégia de varredura do campo visual por informação relevante, e fixações mais prolongadas na região da professora, provavelmente em atendimento social. Quanto às fixações na região dos gestos, observamos valor intermediário entre o valor observado para a região da professora e o valor observado para a região do slide. Contudo, a grande variância nos dados de tempo médio de fixação ocular na região dos gestos sugere diferentes naturezas de fixações entre os sujeitos, algumas possivelmente em estratégia de varredura por informação visual, outras por atendimento e compartilhamento social da atenção. Verificamos ainda o tempo de apresentação dos slides interferindo no tempo médio da fixação ocular nas regiões de interesse, com fixações significativamente mais duradouras para a segunda metade de cada slide, tanto para a região da professora quanto para a região do slide, sendo tal diferença maior na condição *sem gestos*.

Não encontramos efeito do perfil do público, tal a idade dos participantes, sexo, formação e conhecimento prévio sobre o assunto, sobre o tempo de fixação ocular ou aprendizagem do conteúdo da aula. Aqueles que realizaram a sessão experimental pela manhã, entretanto, distinguiram-se dos demais com média de fixação ocular superior para a região da professora (57,5%; $\pm 13,8\%$), distinguindo-se de ambas porcentagens dos sujeitos que realizaram a sessão experimental pela tarde (42,4%; $\pm 13,3\%$) e pela noite (45,6%; $\pm 14,9\%$), estas últimos não se distinguindo entre si.

A avaliação de aprendizagem aplicada aos sujeitos participantes se mostrou de baixa dificuldade resultando em efeito teto. Não foi possível, portanto, avaliar o efeito de nosso controle experimental e da estratégia de fixação ocular dos participantes sobre sua aprendizagem.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aboitiz, F. (2012) Gestures, vocalizations and memory in language origins. *Front. Evol. Neurosci.* 4:2.
- Aboitiz, F., García, R. (2009) Merging of phonological and gestural circuits in early language evolution. *Rev. Neurosci.* 20, 71–84.
- Aboitiz, F., García, R., Bosman, C., Brunetti, E. (2006) Cortical memory mechanisms and language origins. *Brain Lang.*, 98, 40–56.
- Abrams, Z.L., Byrd, D.R. (2016) *Spoken corpus linguistics: From monomodal to multimodal*. London, Routledge.
- Adolphs, R. (1999) Social cognition and the human brain. *Trends Cognit. Sci.*, 3, 469–479.
- Ahmadian, M., Tavakoli, M., Vahid Dastjerdi, H. (2015) The combined effects of online planning and task structure on complexity, accuracy and fluency of L2 speech. *Language Learning Journal*, 43(1), 41–56.
- Alcock, J. (2013) *Animal Behavior: An Evolutionary Approach*. Oxford University Press.
- Alibali, M.W., Heath, D.C., Myers, H.J. (2001) Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. *Journal of Memory and Language*, 44, 169–188.
- Alibali, M.W., Kita, S., Young, A.J. (2000) Gesture and the process of speech production: we think, therefore we gesture. *Language and Cognitive Processes*, 15, 469–489.
- Alibali, M.W., Spencer, R.C., Knox, L., Kita, S. (2011) Spontaneous gestures influence strategy choices in problem solving. *Psychological Science*, 22(9), 1138–1144.
- Allison, T., Puce, A., McCarthy, G. (2000) Social perception from visual cues: role of the STS region. *Trends Cognit. Sci.*, 7, 267–278.
- Allison, T., Puce, A., Spencer, D.D., McCarthy, G. (1999) Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* 9, 415–430.
- Arbib, M.A. (2005) From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral & Brain Sc.*, 28, 105–167.
- Arbib, M.A. (2012) *How the Brain Got Language: The Mirror System Hypothesis*. Oxford University Press.
- Argyle, M., Cook, M. (1976) *Gaze and Mutual Gaze*. Cambridge, Cambridge University Press.
- Awh, E., Armstrong, K.M., Moore, T. (2006) Visual and oculomotor selection: links, causes and implications for spatial attention. *Trends Cognit. Sci.*, 10, 146–155.
- Awh, E., Jonides, J., Smith, E.E. (1996) Dissociation of storage and retrieval in verbal working memory: evidence from positron emission tomography. *Psychol. Sci.*, 7, 25–31.

- Baddeley, A.D. (1999) *Essentials of Human Memory*. Bristol, Psychology Press.
- Baddeley, A.D. (2000) Baddeley, A. D. The episodic buffer: a new component of working memory? *Trends Cognit. Sci.*, 4, 417–423.
- Baddeley, A.D. (2003) Working memory: looking back and looking forward. *Nat. Rev. Neurosci.*, 4, 829–839.
- Baddeley, A.D, Gathercole, S., Papagno, C. (1998) The phonological loop as a language learning device. *Psychological Review*, 105(1), 158–173.
- Baddeley, A.D, Hitch, G. (1974) Working memory. In Bower, G.A. (Ed) *The Psychology of Learning and Motivation*. New York: Academic Press, 8, 47-89.
- Baron-Cohen, S. (1995) *Mindblindness: An Essay on Autism and Theory of Mind*, MIT Press.
- Barros-Loscertales, A., Gonzalez, J., Pulvermüller, F., Ventura-Campos, N., Bustamante, J. C., Costumero, V., Parcet, M.A., Ávila (2012) Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. *Cereb. Cortex*, 22, 2554–2563.
- Bear, M.F., Connors, B.W., Paradiso, M.A. (2016) *Neuroscience*. 4a. edição. Wolters Kluwer.
- Beattie, G., Coughlan, J. (1998) Do iconic gestures have a functional role in lexical access? An experimental study of the effects of repeating a verbal message on gesture production. *Semiotica*, 119-3/4, 221-249.
- Beattie, G., Shovelton, H. (1999) Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*, 123, 1-30.
- Beilock, S.L., Goldin-Meadow, S. (2010) Gestures changes thought by grounding it in action. *Psychological Science*, 21, 1605-1610.
- Berman, R.A., Slobin, D. (1994) *Relating events in narratives: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
- Bonda, E., Petrides, M., Ostry, D., Evans, A. (1996) Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J. Neurosci.*, 16, 3737–3744.
- Bosworth, R.G., Bartlett, M.S., Dobkins, K.R. (2006) Image statistics of American Sign Language: Comparison with faces and natural scenes. *J Optical Society of America*, A 23(9): 2085–2096.
- Breitmeyer, B.G. (1975) Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Res.*, 15(12): 1411–1412.
- Brothers, L. (1990) The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Concepts Neurosci.*, 1, 27–51.
- Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.-J. (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.*, 13, 400–404.
- Bullier, J. (2001) Feedback connections and conscious vision. *Trends Cogn. Sci.*, 5(9): 369–370.

- Burswell, G.T. (1935) *How People Look at Pictures*. Chicago University Press.
- Butcher, C., Goldin-Meadow, S. (2000) Gesture and the transition from one- to two-word speech: When hand and mouth come together. In McNeill, D. (Ed) *Language and gesture*. Cambridge University Press, 235-257.
- Butterworth, B., Beattie, G. (1978) Gesture and silence as indicators of planning in speech. In R. N. Campbell & E T. Smith (Eds) *Recent advances in the psychology of language: Formal and experimental approaches*. New York: Plenum Press, 347-360.
- Callan, D., Jones, J.A., Munhall, K., Kroos, C., Callan, A.M. (2004) Multisensory integration sites identified by perception of spatial wavelet filtered visual speech gesture information. *J Cognit. Neurosci.*, 16: 805–816.
- Calvert, C.A. (2001) Cross-modal processing in the human brain. Insights from functional neuroimaging studies. *Cereb. Cortex*, 11, 1110–1123.
- Capirci, O., Iverson, J.M., Pizzuto, E., Volterra, V. (1996) Communicative gestures during the transition to two-word speech. *Journal of Child Language*, 23, 645–673.
- Carpenter, R.H.S. (1977) *Movements of the Eyes*. Pion.
- Cartmill, E.A., Beilock, S., Goldin-Meadow, S. (2012) A word in the hand: action, gesture and mental representation in humans and non-human primates. *Phil. Trans. Soc.*, B 367, 129-143.
- Cavanaugh, J.R., Bair, W., Movshon, J.A. (2002) Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J Neurophysiol*, 88(5), 2547-56.
- Cho, M. (2018) Task complexity, modality, and working memory in L2 Task performance. *System*, 72, 85-98.
- Chomsky, N. (1965) *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1975) *Reflections on Language*. New York, NY: Pantheon Books.
- Chu, M., Kita, S. (2011) The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, 140, 102-116.
- Citron, F.M.M. (2012) Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain Lang.*, 122, 211–226.
- Clark, A. (1998) *Being There: Putting Brain, Body, and World Together Again*. A Bradford Book.
- Clark, A. (2008) *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press.
- Colletta, J.-M., Guidetti, M., Capirci, O., Cristilli, C., Demir, Ö.E., Kunene-Nicolas, R.N., Levine, S. (2015) Effects of age and language on co-speech gesture production: An investigation of French, American, and Italian children's narratives. *Journal of Child Language*, 42 (1), 122–145.
- Colletta, J.-M., Pellenq, C., Guidetti, M. (2010) Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. *Speech Communication*, 52, 565–76.
- Condon, W.S., Ogston, W.D. (1966) Sound Film Analysis of Normal and Pathological Behaviour Patterns. *Journal of Nervous and Mental Disease*, 143, 4, 338-347.

- Cooperrider, K., Wakefield, E., Goldin-Meadow, S. (2015) More than meets the eye: gestures changes thought, even without visual feedback. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Crystal, D. (1969) *Prosodic Systems and Intonation in English*. Cambridge, Cambridge University Press.
- Cusick, C.G. (1997) The superior temporal polysensory region in monkeys. In Rockland, K. et al. (eds) *Cerebral Cortex: Extrastriate Cortex in Primates*, 12, 435–468. Plenum Press.
- Damasio, A.R. (1994) *Descarte's Error: Emotion, Reason, and the Human Brain*. Putnam, New York, 1994.
- Dart, S. N. (1992) Narrative style in the two languages of a bilingual child. *Journal of Child Language*, 19, 367–387.
- De Ruiter, J.-P. (2008) *Gesture and speech production*. Wageningen, Ponsen & Looijen bv.
- Dell, C. (1970) *A primer for movement description using effort-shape and supplementary concepts*. Princeton Book Company Publishers.
- Demir, Ö.E., Fisher, J.A, Goldin-Meadow, S., Levine, S.C. (2014) Narrative processing in typically developing children and children with early unilateral brain injury: Seeing gesture matters. *Developmental Psychology*, 50 (3), 815–828.
- Demir, Ö.E., Levine, S.C., Goldin-Meadow, S. (2015) A tale of two hands: Children's early gesture use in narrative production predicts later narrative structure in speech. *Journal of Child Language*, 42 (3), 662–681.
- Descartes, R. (1637) *The discourse on method and metaphysical meditations*. London: Walter Scott.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G. (1992) Understanding motor events: an neurophysiological study. *Exp. Brain Res.*, 91, 176–180.
- Dittmann, A.T., Llewelyn, L.G. (1969) Body movement and speech rhythm in social conversation. *Journal of Personality and Social Psychology*, 23, 283-292.
- Duchowski, A.T. (2007) *Eye Tracking Methodology: Theory and Practice*. Springer.
- Dudschig, C., DeLaVega, I., Kaup, B. (2014) Embodiment and secondlanguage: automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical Stroop paradigm. *Brain Lang.*, 132, 14–21.
- Efron, D. (1941) *Gesture and Environment*. Morningside Heights, NY, King's Crown Press.
- Efron, D. (1972) *Gesture, Race and Culture*. The Hague: Mouton & Co.
- Eielts, C., Pouw, W., Ouwehand, K., VanGog, T., Zwaan, R.A., Paas, F. (2018) Co-thought gesturing supports more complex problem solving in subjects with lower visual working-memory capacity. *Psychological Research*, 1-12.
- Ekman, P. (1964) Body position, facial expression and verbal behavior during interviews. *Journal of abnormal and social psychology*, 68, 3, 295-301.

- Ekman, P. (1965) Differential communication of affect by head and body cues. *Journal of Personality and Social Psychology*, 2, 5, 726-735.
- Ekman, P., Friesen, W. (1965) Personality, pathology, affect and nonverbal behavior. Trabalho apresentado em Western Psychological Association Convention. Havaí.
- Ekman, P., Friesen, W. (1967a) Head and body cues in the judgment of emotion: A reformulation. *Perceptual and Motor Skills*, 24, 711-724.
- Ekman, P., Friesen, W. (1967b) Nonverbal leakage and deception clues. Trabalho apresentado em encontro Western Psychological Association.
- Ekman, P., Friesen, W. (1968) Nonverbal behavior in psychotherapy research. In J. M. Shlien (Ed) *Research in psychotherapy*. American Psychological Association, 3, 179-216.
- Ekman, P., Friesen, W. (1969) The repertoire of non-verbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1.
- Ellis, R. (2016) Focus on form: A critical review. *Language Teaching Research*, 20(3), 405-428.
- Ellis, H.D., Gunter, H.L. (1999) Asperger syndrome: a simple matter of white matter? *Trends Cognit. Sci.*, 3, 192-200
- Engelkamp, J., Zimmer, H.D. (1985) Motor programs and their relation to semantic memory. *German Journal of Psychology*, 9, 239-254.
- Eschen, A., Freeman, J., Dietrich, T., Martin, M., Ellis, J., Martin, E. (2007) Motor brain regions are involved in the encoding of delayed intentions: A fMRI study. *International Journal of Psychophysiology*, 64, 259-268.
- Fasolo, M., D'Odorico, L. (2012) Gesture-plus-word combinations, transitional forms, and language development. *Gesture*, 12 (1), 1-15.
- Fischer, M.H., Zwaan, R.A. (2008) Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol.*, 61, 825-850.
- Fodor, J.A. (1976) *The language of thought*. Hassocks: Harvester Press.
- Fodor, J.A. (1987) *Psychosemantics: The problem of meaning in the Philosophy of mind*. Cambridge, MA: MIT Press.
- Gallagher, S. (2005) *How the Body Shapes the Mind*. Oxford University Press.
- Gallagher, H.L., Happé, F., Brunswick, N., Fletcher, P.C., Frith, U., Frith, C.D. (2000) Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, 38, 11-21.
- Gentilucci, M., Dalla Volta, R. (2008) Spoken language and arm gestures are controlled by the same motor control system. *Q. J. Exp. Psychol.*, 61, 944-957.
- Gerbella, M., Belmalih, A., Borra, E., Rozzi, S., Luppino, G. (2011) Cortical connections of the anterior (F5a) subdivision of the macaque ventral premotor area F5. *Brain Struct. Funct.*, 216, 43-65.

- Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*. Psychology Press.
- Giordan, M., Silva-Neto, A. B., Aizawa, A. (2015) Relações entre gestos e operações epistêmicas mediadas pela representação estrutural em aulas de química e suas implicações para a produção de significados. *Química Nova Escola*, 37(1), 82-94.
- Goffaux, V., Rossion, B. (2006) Faces are “spatial” – holistic face perception is supported by low spatial frequencies. *J Exp Psychol. Hum. Percept. Perform.*, 32(4):1023–1039.
- Goffaux, V. (2009) Spatial interactions in upright and inverted faces: reexploration of spatial scale influence. *Vision Res.*, 49(7), 774–781.
- Goldin-Meadow, S. (2003) *Hearing Gesture: How Our Hands Help Us Think*. Belknap Press.
- Goldin-Meadow, S., Alibali, M.W. (2013) Gesture’s role in speaking, learning, and creating language. *Annual Review of Psychology*, 64, 257-283.
- Goldin-Meadow, S., Cook, S.W., Beilock, S.L. (2010) Gesturing gives children new ideas about math. *Psychol Sci*, 2009, Mar, 20(3), 267-272.
- Goldin-Meadow, S., Kim, S., Singer, M. (1999) What the teacher’s hands tell the student’s mind about math. *Journal of Educational Psychology*, 91, 720-730.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S.D., Wagner, S. (2001) Explaining math: gesturing lightens the load. *Psychological Science*, 12, 516-522.
- Gollan, T.H., Acenas, L.-A.R. (2004) What is a TOT? Cognate and translation effect on tip-of-the-tongue states in Spanish–English and Tagalog–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 246–269.
- Goodwin, J. (2001) Teaching pronunciation. In M. Celce-Murcia (Ed) *Teaching English as a second or foreign language*, 103–116. Boston: Heinle & Heinle.
- González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., Ávila, C. (2006) Reading cinnamon activates olfactory brain regions. *Neuroimage*, 32, 906–912.
- Graham, J.A., Argyle, M. (1975) A cross-cultural study of the communication of extra-verbal meaning by gestures. *International journal of Psychology*, 10, 57-67.
- Graham, J.A., Heywood, S. (1975) The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Social Psychology*, 5(2), 189-195.
- Graziano, M., Gullberg, M. (2013) Gesture production and speech fluency in competent speakers and language learners. In *Tilburg Gesture Research Meeting (TiGeR)*, 2013. Tilburg, NL.
- Gregersen, T., Olivares-Cuhat, G., Storm, J. (2009) An Examination of L1 and L2 Gesture Use: What Role Does Proficiency Play? *The modern Language Journal*, 93, 195-208.
- Grèzes, J., Costes, N., Decety, J. (1998) Top-down effect of strategy on the perception of human biological motion: a PET investigation. *Cognit. Neuropsychol.*, 15, 553-582.
- Gullberg, M. (1998) *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*. Lund: Lund University Press.

- Gullberg, M. (1999) Communication, strategies, gestures, and grammar. *Acquisition et Interaction en Langue Etrangère*, 2, 61–71.
- Gullberg, M. (2003) Gestures, referents, and anaphoric linkage in learner varieties. In Christine Dimroth & Marianne Starren (Eds) *Information structure, linguistic structure and the dynamics of language acquisition*. Amsterdam: Benjamins.
- Gullberg, M. (2006) Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, 56 (1), 155–196.
- Gullberg, M., Holmqvist, K. (1999) Keeping an eye on gestures: visual perception of gestures in face-to-face communication. *Pragmatics & Cognition*, 7(1), 35-63.
- Gullberg, M., Kita, S. (2009) Attention to speech gestures: eye movements and information uptake. *Journal of Nonverbal Behavior*, 33, 251-277.
- Hadar, U. (1989) Two types of gesture and their role in speech production. *Journal of Language and Social Psychology*, 8, 221-228.
- Harwerth, R.S., Levi, D.M. (1978) Reaction time as a measure of suprathreshold grating detection. *Vision Res* 18(11), 1579–1586.
- Haviland, J. (1993) Anchoring, iconicity and orientation in Guugu Yimithirr pointing gestures. *Linguistic Anthropology*, 1, 3–45.
- Helene, A.F. (2006) *Aquisição e uso de memória implícita*. Tese de doutorado. Instituto de Psicologia da Universidade de São Paulo. São Paulo.
- Hewes, G. (1973) Primate communication and the gestural origins of language. *Current Anthropology*, 14, 5-24.
- Hickok, G., Poeppel, D. (2007) The cortical organization of speech processing. *Nature Rev. Neurosci.*, 8, 393–402.
- Hostetter, A. B., Alibali, M. W. (2019) Gesture as simulated action: Revisiting the framework. *Psychonomic Bulletin & Review*, 26(3), 721-752.
- Hostetter, A.B., Alibali, M.W., Bartholomew, A. (2011) Gesture as simulated action: Revising the framework. *Psychonomic Bulletin & Review*, 26(3), 721-752.
- Howard, R.J., Brammer, M., Wright, I., Woodruff, P.W., Bullmore, E.T., Zeki, S. (1996) A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain. *Curr. Biol.*, 6, 1015-1019.
- Höweler, M. (1972) Diversity of word usage as a stress indicator in an interview situation. *Journal of Psychological Research*, 1, 3, 243-248.
- Howell, D. (2002) *Statistical methods for psychology*. Wadsworth, University of Vermont.
- Hubbard, A.L., Wilson, S.M., Callan, D.E., Dapretto, M. (2008) Giving speech a hand: Gesture modulates activity in auditory cortex during speech perception. *Hum. Brain Mapp.*, 30, 1028–1037.
- Hurley, S. (1998) *Consciousness in Action*. Harvard University Press.

- Ingram, D. (1975) Motor asymmetries in young children. *Neuropsychologia*, 13, 1, 95-102.
- Iverson, J.M., Braddock, B.A. (2011) Gesture and motor skill in relation to language in children with language impairment. *Journal of Speech, Language, and Hearing Research*, 54, 72–86.
- Iverson, J.M., Capirci, O., Volterra, V., Goldin-Meadow, S. (2008) Learning to talk in a gesture-rich world: Early communication in Italian vs. American children. *First Language*, 28 (2), 164–181.
- Izumi, A., Kojima, S. (2004) Matching vocalizations to vocalizing faces in a chimpanzee (*Pantroglodytes*) *Anim. Cogn.*, 7, 179–184.
- Jacobs, N., Garnham, A. (2007) The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*. Elsevier.
- Jenkins, T., Coppola, M., Coelho, C. (2017) Effects of gesture restriction on quality of narrative production. *Gesture*, 16:3, 416-431.
- Johansson, G. (1973) Visual perception of biological motion and a model of its analysis. *Percept. Psychophys.*, 14, 202–211.
- Jonides, J., Reuter-Lorenz, P. A., Smith, E. E., Awh, E., Barnes, L. L., Drain, M., Glass, J., Lauber, E. J., Patalano, A. L., & Schumacher, E. H. (1996) Verbal and spatial working memory in humans. In Medin, D. (Ed) *The Psychology of Learning and Motivation*. Londres, Academic, 43–88.
- Jonides, J., Smith, E. E. (1997) The architecture of working memory. In Rugg, M.D. (Ed) *Cognitive Neuroscience*. Sussex, Hove, 243–276.
- Kelly, S., Hansen, B., Clark, D. (2012) “Slight” of hand: the processing of visually degraded gestures with speech. *Plos One*, 7(8), e42620.
- Kelly, S.D., Mcdevitt, T., Esch, M. (2009) Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Lang. Cogn. Process.*, 24, 313–334.
- Kendon, A. (1972) Some relationships between body motion and speech. In Siegman & B. Pope (Eds) *Studies in Dyadic Communication*. New York, Pergamon Press, 177-210.
- Kendon, A. (1980) Gesticulation and Speech: Two Aspects of the process of utterance. In Key, M. R. (Ed), *Relationship of verbal and nonbervel communication*, 207-228. Mouton.
- Kendon, A. (1990) *Conducting interaction*. Cambridge: Cambridge University Press.
- Kendon, A. (2004) *Gesture. Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, A. (2017) Pragmatic functions of gestures. *Gesture*, 16(2), 157-175.
- Kiefer, M., Pulvermüller, F. (2012) Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex*, 48, 805–825.
- Kimura, D. (1976) The neural basis of language qua gesture. In H. & H. A. Whitaker, *Studies in Neurolinguistics*, 2, 145-156.
- Kingdon, R. (1958) *The Groundwork of English Intonation*. London: Longmans.

- Kirk, E., Lewis, C. (2017) Gesture facilitates children's creative thinking. *Psychological Science*, 28(2), 225–232.
- Kita, S., Alibali, M. W., Chu, M. (2017) How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, 124(3), 245.
- Klein, C.L. (1986) Gaze and Eye Contact: A research Review. *Psychological Bulletin*, 100, 1, 78–100.
- Kormos, J., Trebits, A. (2012) The role of task complexity, modality, and aptitude in narrative task performance. *Language Learning*, 62(2), 439–472.
- Krauss, R.M. (1998) Why do we gesture when we speak? *Current Directions in Psychological Science*, 7, 54–60.
- Krauss, R.M., Chen, Y., Chawla, P. (1996) Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? *Adv. Exp. Social Psych.*, 28, 389–450.
- Krauss, R.M., Chen, Y., Gottesman, R. (2000) Lexical gestures and lexical access: A process model. In D. McNeill (Ed) *Language and gesture* (261–283) Cambridge, UK: Cambridge University Press.
- Krauss, R.M., Morrel-Samuels, P., Colasante, C. (1991) Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, 61, 5, 743–754.
- Kroll, J.F., Michael, E., Tokowicz, N., Dufour, R. (2002) The development of lexical fluency in a second language. *Second Language Research*, 18, 137–171.
- Langton, S.R.H., Watt, R.J., Bruce, V. (2000) Do the eyes have it? Cues to the direction of social attention. *Trends Cognit. Sci.*, 4, 50–58.
- Laurent, A., Nicoladis, E. (2015) Gesture restriction affects French-English bilinguals' speech only in French. *Language and Cognition*, 18(2), 340–349.
- Laurent, A., Nicoladis, E., Marentette, P. (2010) Discours, gestes et dominance linguistique: une étude développementale. *Enfance*, 3, 275–285.
- Laurent, A., Smithson, L., Nicoladis, E. (2020) Gestures tell a story creativity; non-gesture tell it like it happened. *Language Learning and Development*, 16(3), 292–308.
- Leavens, D.A. (2004) Manual deixis in apes and humans. *Interaction Studies*, JB, 5:3, 387–408.
- Levelt, W.J.M. (1989) *Speaking*. Cambridge, Massachusetts, MIT Press.
- Levy, E.T., Fowler, C.A. (2000) The role of gestures and other graded language forms in the grounding of reference in perception. In David McNeill (Ed) *Language and gesture*. Cambridge: Cambridge University Press, 215–234.
- Levy, E.T., McNeill, D. (1992) Speech, gesture and discourse. *Discourse Processes*, 15 (3), 277–301.
- Lickiss, K.P., Wellens, A.R. (1978) Effects of visual accessibility and had restraint on fluency of gesticulation and effectiveness of message. *Perceptual and Motor Skills*, 31, 533–552.

- Lin, Y.-L. (2017) Co-occurrence of speech and gestures: A multimodal corpus linguistic approach to intercultural interaction. *Journal of Pragmatics*, 117, 155-167.
- Lin, Y.-L. (2019) Speech-accompanying gestures in L1 and L2 conversational interaction by speakers of different proficiency levels. *International Review of Applied Linguistics in Language Teaching*, 1-15.
- Lin, Y.-L. (2020) A helping hand for thinking and speaking: Effects of gesturing and task planning on second language narrative discourse. *System*, 91.
- Lin, K.J., Loi, M.X., Shiao, C.C., Wu, C.L., Ho, J.H.C. (2013) Decrease in inner retinal thickness at para and perifoveal areas before vascular retinopathy in patients with metabolic risk factors. *Twain Journal of Ophthalmology*, 3, 3, 108-115.
- Lorenz, C. (1981) *Foundations of Ethology*. New York, Springer-Verlag.
- Macedonia, M. (2003) *Sensorimotor enhancing of verbal memory through “voice movement icons” during encoding of foreign language*. Ph.D. thesis, University of Salzburg, Salzburg.
- Macedonia, M. (2014) Bringing back the body into the mind: gestures enhance word learning in foreign language. *Frontiers in Psychology*, 5, 1467.
- Macedonia, M., Klimesch, W. (2014) Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind Brain Educ.*, 8, 74–88.
- Macedonia, M., Müller, K., Friederici, A. D. (2011) The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapp.*, 32, 982-998.
- Macedonia, M., Knösche, T.R., (2011) Body in Mind: How Gestures Empower Foreign Language Learning. *Mind, Brain, and Education*, 5, 196-211.
- Mahl, G.F. (1968) Gestures and body movements in interviews. Trabalho apresentado em *Third Research in Psychotherapy Conference* (Chicago, June 1966); V. III, American Psychological Association.
- Mainela-Arnold, E., Alibali, M., Hostetter, A.B., Evans, J.L. (2014) Gesture–speech integration in children with specific language impairment. *International Journal of Language & Communication Disorders*, 49 (6), 761–770.
- Masumoto, K., Yamaguchi, M., Sutani, K., Tsuneto, S., Fujita, A., Tonoike, M. (2006) Reactivation of physical motor information in the memory of action events. *Brain Research*, 1101, 102-109.
- Matthews-Saugstad, K.M., Raymakers, E.P., Kelty-Stephen, D.G. (2016) Gesturing more diminishes recall of abstract words when gesture is allowed and concrete words when it is taboo. *The Quarterly Journal of Experimental Psychology*, 70(7), 1099–1105.
- Mayberry, R., Jaques, J., DeDe, G. (1998) What stuttering reveals about the development of the gesturespeech relationship. *New Directions for Child Development*, 79, 77-87.
- McNeill, D. (2005) *Gesture and Thought*. University of Chicago Press.
- McNeill, D. (1992) *Hand and mind. What the hands reveal about thought*. Chicago: Chicago University Press.

- McNeill, D. (1998) Speech and gesture integration. In J. M. Iverson and S. Goldin-Meadow (Eds) *The nature and functions of gesture in children's communication*. San Francisco: Jossey-Bass, 11-27.
- McNeill, D., Duncan, S.D. (2000) Growth points in thinking-for-speaking. In D. McNeill (Ed) *Language and gesture*, 141-161. Cambridge University Press.
- Meunier, H., Vauclair, J., Fagard, J. (2012) Human Infants and Baboons Show the Same Pattern of Handedness for a Communicative Gesture. *PLoS One*, 7(3): e33959.
- Milner, A. D., Goodale, M.A. (1995) *The visual brain in action*. Oxford University Press.
- Montgomery, K.J., Isenberg, N., Haxby, J.V. (2007) Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Soc. Cognitive Affect Neurosci.*, 2: 114–122.
- Morrel-Samuels, P., Krauss, R.M. (1992) Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 615-622.
- Morsella, E., Krauss, R.M. (2004) The role of gestures in spatial working memory and speech. *American Journal of Psychology*, 117(3), 411-424.
- Moseley, R. L., and Pulvermüller, F. (2014) Nouns, verbs, objects, actions, and abstractions: local fMRI activity indexes semantics, not lexical categories. *Brain Lang.*, 132, 28–42.
- Munhall, K.G., Kroos, C., Jozan, G., Vatikiotis-Bateson, E. (2004) Spatial frequency requirements for audiovisual speech perception. *Percept Psychophys*, 66: 574–583.
- Nathan, M.J., Alibali, M.W. (2011) How gesture use enables intersubjectivity in the classroom. In Stam, G. & Ishino, M. (Eds.) *Integrating gestures: The interdisciplinary nature of gesture*. Amsterdam: John Benjamins.
- Nicoladis, E. (2002) Some gestures develop in conjunction with spoken language development and others don't: Evidence from bilingual preschoolers. *Journal of Nonverbal Behavior*, 26, 241–266.
- Nicoladis, E., Marentette, P., Navarro, S. (2014) Gesture frequency linked primarily to story length in 4–10-year-old children's stories. *J. Psycholinguistic Research*,
- Nicoladis, E., Mayberry, R.I., Genessee, F. (1999) Gesture and early bilingual development. *Developmental Psychology*, 35, 514-526.
- Nicoladis, E., Nagpal, J., Marentette, P. (2011) Language proficiency and culture affect bilinguals' gesture frequency. Presented at the American Association for Applied Linguistics, Chicago, IL.
- Nicoladis, E., Pika, S., Yin, H., Marentette, P. (2007) Gesture use in story recall by Chinese–English bilinguals. *Applied Psycholinguistics*, 28, 719–733.
- Nicoladis, E., Pika, S., Marentette, P. (2009) Do French–English bilingual children gesture more than monolingual children? *Journal of Psycholinguistic Research*, 38, 573– 585.
- Nilsson, L.G. (2000) Remembering actions and words. In Craik, F. I. M., Tulving, E. (Eds) *The Oxford Handbook of Memory*. Oxford University Press, 137-148.

- Nitta, R., Nakatsuhara, F. (2014) A multifaceted approach to investigating pre-task planning effects on paired oral test performance. *Language Testing*, 31(2), 147-175.
- Nöe, A. (2004) *Action in Perception*. The MIT Press.
- Nowak, L.G., Munk, M.H., Girard, P., Bullier, J. (1995) Visual latencies in areas V1 and V2 of the macaque monkey. *Visual Neurosci.*, 12(2), 371-384.
- O'Neill, D.K., Holmes, A.C. (2002) Young preschoolers' ability to reference story characters: The contribution of gestures and character speech. *First Language*, 22, 73-103.
- Oram, M.W., Richmond, B.J. (1999) I see a face – a happy face. *Nat. Neurosci.*, 2, 856-858.
- Özyürek, A. (2002) Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, 46, 688-704.
- Parker D.M., Dutch, S. (1987) Perceptual latency and spatial frequency. *Vision Res.*, 27(8), 1279-1283.
- Parrill, F., Bullen, J., Hoburg, H. (2010) Effects of input modality on speech-gesture integration. *Journal of Pragmatics*, 42(11), 3130-3137.
- Paulesu, E., Frith, C.D., Frackowiak, R.S. (1993) The neural correlates of the verbal component of working memory. *Nature*, 362, 342-345.
- Pereira, C.D., Haddad, H. (2015) *O papel dos gestos manuais na compreensão de narrativas em língua estrangeira: um estudo com rastreamento de movimentos oculares*. Dissertação de mestrado. São Paulo, Universidade de São Paulo.
- Perrett, D.I., Harries, M.H., Bevan, R., Thomas, S., Benson, P.J., Mistlin, A.J., Chitty, A.J., Hietanen, J.K., Ortega, J.E. (1989) Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.*, 146, 87-113.
- Perrett, D.I., Mistlin, A.J., Harries, M.H., Chitty, A.J. (1990). Understanding the visual appearance and consequence of hand actions. In Goodale M.A. (Ed) *Vision and Action: The Control of Grasping*. Norwood, NJ: Ablex, 163-342.
- Perrett, D.I., Hietanen, J.K., Oram, M.W., Benson, P.J., Rolls, E.T. (1992) Organization and functions of cells responsive to faces in the temporal cortex. *Philos. Trans. R. Soc. London, Ser. B* 335, 23-30.
- Piaget, J. (1976) Piaget's theory. In B. Inhelder, H. Chipman, C. Zwingmann (Eds) *Piaget and His School*. Berlin: Springer, 11-23.
- Pika, S., Nicoladis, E., Marentette, P. (2006) A cross-cultural study on the use of gestures: Evidence for cross-linguistic transfer? *Bilingualism: Language and Cognition*, 9, 319-327.
- Pilz, K.S., Bulthoff, H., Vuong, Q.C. (2009) Learning influences the encoding of static and dynamic faces and their recognition across different spatial frequencies. *Vis. Cogn.* 17: 716-735.
- Posner, M. and Dehaene, S. (1994) Attentional networks. *Trends Neurosci.*, 17, 75-79

- Ping, R., Goldin-Meadow, S. (2010) Gesturing saves cognitive resources when talking about non-present objects. *Cogn. Sci.*, 34(4), 602-619.
- Posner, M.I., Snyder, C.R.R., Davidson, B.J. (1980) Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160-174.
- Pouw, W.T., Mavilidi, M.F., Van Gog, T., Paas, F. (2016) Gesturing during mental problem solving reduces eye movements, especially for individuals with lower visual working memory capacity. *Cogn Process*, 17(3), 269-277.
- Premack, D. and Woodruff, G. (1978) Does the chimpanzee have a 'theory of mind'? *Behav. Brain Sci.*, 4, 515-526.
- Pulvermüller, F. (1999) Words in the brain's language. *Behav. Brain Sci.*, 22, 253- 279.
- Purves, D., Augustine, G., Fitzpatrick, D., Hall, W., Lamantia, A.-S., McNamara, J., Williams, S. (2004) *Neuroscience*. Sunderland, Massachusetts.
- Ragsdale, J. D., Silvia, C.F. (1982) Distribution of kinesic hesitation phenomena in spontaneous speech. *Language and Speech*, 25(2), 185-190.
- Rauscher, F.H., Krauss, R.M., Chen, Y. (1996) Gesture, speech, and lexical access: the role of lexical movements in speech production. *Psychological Science*, 7(4), Columbia University.
- Riedl, T.R., Sperling, G. (1988) Spatial-frequency bands in complex visual stimuli: American Sign Language. *J Optical Society of America*, A 5(4), 606-616.
- Rimé, B., Boulanger, B., d'Ydewalle, G. (1988) Visual attention to the communicator's nonverbal behavior as a function of the intelligibility of the message. Apresentado em *Symposium on TV Behavior*, 24th International Congress of Psychology, Sydney.
- Rimé, B., Schiaratura, L. (1991) Gesture and speech. In: R. S. Feldman, B. Rimé (Eds) *Fundamentals of nonverbal behavior*, 239-281. New York, Cambridge University Press.
- Rimé, B., Shiaratura, L., Hupet, M., Ghysseleincx, A. (1984) Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion*, 8(4), 311-325.
- Rizzolatti, G., Arbib, M.A. (1998) Language within our grasp. *Trends Neurosci.*, 21, 188-194.
- Rizzolatti, G., Riggio, L., Dascola, I., Umiltà, C. (1987) Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25, 31-40.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., Fazio, F. (1996) Localization of grasp representations in humans by PET: 1. Observation versys execution. *Exp. Brain Res.*, 111, 246-252.
- Rizzolatti, G., Luppino, G. (2001) The cortical motor system. *Neuron.*, 31, 889- 901.
- Rizzolatti, G.W., Craighero, L. (2004) The mirror-neuron system. *Annu. Rev. Neurosci.*, 27, 169-192.
- Rockland, K. (1997) Elements of cortical architecture: hierarchy revisited. In Rockland, K. et al. (Eds) *Cerebral Cortex, Extrastriate Cortex in Primates*, 12. Plenum Press, 243-293.

- Romanski, L.M., Averbeck, B.B., Diltz, M. (2005) Neural representation of vocalizations in the primate ventro lateral prefrontal cortex. *J. Neurophysiol.*, 93, 734–747.
- Rosenfeld, H.M. (1966) Approval-seeking and approval-unducing functions of verbal and nonverbal responses in the dyad. *Journal of Personality and Social Psychology*, 4(6), 597–605.
- Rowlands, M. (2010) *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. A Bradford Book.
- Rubin, D.C. (1995) *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. New York: Oxford University Press.
- Saussure, F.de (1916) *Cours de linguistique générale*. Lausanne: Balley & Séchehayé.
- Seyfeddinipur, M., Kita, S. (2001) Gestures and disfluencies in speech. In S. Santi, I. Guaïtella, C. Cavé et G. Konopczynski (Eds) *Oralité et gestualité. Communication multimodale, interaction*, 266-270. Paris: L'Harmattan.
- Schmolesky, M.T., Wang, Y., Hanes, D.P., Thompson, K.G., Leutgeb, S., Leventhal, A.G. (1998) Signal timing across the macaque visual system. *J Neurophysiol*, 79(6), 3272–3278.
- Sekiyama, K., Kanno, I., Miura, S., Sugita, Y. (2003) Auditory-visual speech perception examined by fMRI and PET. *Neurosci. Res.*, 47, 277–287.
- Shannon, C.E., Weaver, W. (1949) *The mathematical theory of communication*. Urbana, Illinois, University of Illinois Press.
- Shapiro, L. (2011) *Embodied Cognition*. Routledge.
- Slobin, D. (1987) Thinking for speaking. In J. Aske, N. Beery, L. Michaelis, H. Filip (Eds) *Paper from the 13th Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: BLS.
- Smith, E., Jonides, J., Koeppe, R.A. (1996) Dissociating verbal and spatial working memory using PET. *Cereb. Cortex*, 6, 11-20.
- Smithson, L., Nicoladis, E. (2014) Lending a hand to imagery? The impact of visuospatial working memory interference upon iconic gesture production in a narrative task. *Journal of Nonverbal Behavior*, 38(2), 247-258.
- Stam, G. (2017) Verb framed, satellite framed or in between? A L2 learner's thinking for speaking in her L1 and L2 over 14 years. In I. Ibarretxe-Antunano (Ed) *Motion and space across languages: Theory and applications*. Philadelphia/Amsterdam: John Benjamins, 329-365.
- Stites, L. J., Özçalışkan, Ş. (2017) Who did what to whom? Children track story referents first in gesture. *Journal of Psycholinguistic Research*, 46(4), 1019–1032.
- Stout, D., Chaminade, T. (2012) Stone tools, language and the brain in human evolution. *Phil. Trans. R. Soc.*, B 367, 75-87.
- Straube, T., Sauer, A., Miltner, W.H.R. (2011) Brain activation during direct and indirect processing of positive and negative words. *Behav. Brain Res.*, 222, 66–72.
- Sugase, Y., Yamane, S., Ueno, S., Kawano, K. (1999) Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, 400, 869–873.

- Tagliabata, J.P., Russell, J.L., Schaeffer, J.A., Hopkins, W.D. (2008) Communicative signaling activates 'Broca's' homolog in chimpanzees. *Curr. Biol.*, 18, 343–348.
- Tellier, M. (2008) The effect of gestures on second language memorization by young children. *Gesture*, 8, 219-235.
- Thal, D.J., Tobias, S. (1992) Communicative gestures in children with delayed onset of oral expressive vocabulary. *Journal of Speech, Language, and Hearing Research*, 35 (6), 1281–1289.
- Tokarev, K., Tiunova, A., Scharff, C., Anokhin, K. (2011) Food for song: Expression of C-Fos and Zenk in the zebra finch song nuclei. *PLoSOne*, 6, e21157.
- Tomasello, M. (2005) *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomkins, S.S., McCarter, R. (1964) What and where are the primary affects? Some evidence for a theory. *Perception and motor skills*, 18, 119-159.
- Trevarthen, C. (1977) Descriptive analyses of infant communicative behavior. In H. R. Schaffer (Ed) *Studies in mother-infant interaction*. London: Academic Press, 227-270.
- Ungerleider, L.G., Haxby, J.V. (1994) 'What' and 'where' in the human brain. *Curr. Opin. Neurobiol.* 4, 157–165
- Valenko, S., Cviljušac, V, Zlatić, S., Modrić, D. (2019) The Impact of Physical Parameters on the Perception of the Moving Elements in Peripheral Part of the Screen. *Technical Gazette*, 26, 5, 1444-1450.
- Van Gompel, R.P.G., Fischer, M.H., Murray, W.S., Hill, R.L. (2007) *Eye Movements: A Window on Mind and Brain*. Elsevier.
- Vilà-Giménez, I., Prieto, P. (2020) Encouraging kids to beat: Children's beat gesture production boosts their narrative performance. *Developmental Science*, Mar 28, e12967.
- Vygotsky, L.S. (1978) *Mind in Society*. M. Cole, V. John-Steiner, S. Scribner, & E. Soubelman (Eds), Cambridge: Harvard University Press.
- Vygotsky, L.S. (1986) *Thought and Language*. Cambridge: MIT Press.
- Wagner, S.M., Nusbaum, H., Goldin-Meadow, S. (2004) Probing the Mental Representation of Gesture: Is Handwaving Spatial? *Journal of Memory & Language*, 50(4), 395–407.
- Watson, C.E., Cardillo, E.R., Ianni, G.R., Chatterjee, A. (2013) Action concepts in the brain: an activation likelihood estimation meta-analysis. *J. Cogn. Neurosci.*, 25, 1191–1205
- Wesp, R., Hesse, J., Keutmann, D., Wheaton, K. (2001) Gestures maintain spatial imagery. *American Journal of Psychology*, 114, 591–600.
- Wigglesworth, G. (1997) Children's individual approaches to the organization of narrative. *Journal of Child Language*, 24, 279–309.
- Willems, R.M., Özyürek, A., Hagoort, P. (2007) When language meets action: the neural integration of gesture and speech. *Cereb. Cortex*, 17, 2322–2333.

Wright, R.D., Ward, L.M. (1994) Shifts of visual attention: An historical and methodological overview. *Canadian Journal of Experimental Psychology*, 48, 151-166.

Yarbus, A.L. (1964) *Eye Movements and Vision*. Plenum Press.

Yoshioka, K. (2008) Gesture and information structure in first and second language. *Gesture*, 8(2), 236-255.

Zadavko, R. (2013) *The Hand, an Organ of the Mind: What the Manual Tells the Mental*. The MIT Press.

Zampini, L., Zanchi, P., Suttora, C., Spinelli, M.F., Salerni, N. (2016) Gesture production in the narratives of preschool children. *Gesture*, 15(3), 306-320.

Zimmer, H.D., Helstrup, T., Engelkamp, J. (2000) Pop-out into memory: A retrieval mechanism that is enhanced with the recall of subject-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 658-670.

Software

ELAN (Version 5.0.0 beta) [Computer software]. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from <https://archive.mpi.nl/tla/elan>.

Tobii Studio (Version 3.2.2) [Computer software]. Danderyd, Sweden.

ANEXOS

ANEXO 1

USP- INSTITUTO DE PSICOLOGIA DA UNIVERSIDADE DE SÃO		
PARECER CONSUBSTANCIADO DO CEP		
DADOS DA EMENDA		
Título da Pesquisa: A gesticulação na compreensão e produção de narrativas em língua estrangeira		
Pesquisador: Charles Desena Pereira		
Área Temática:		
Versão: 2		
CAAE: 31781414.6.0000.5561		
Instituição Proponente: UNIVERSIDADE DE SAO PAULO		
Patrocinador Principal: MINISTERIO DA EDUCACAO		
DADOS DO PARECER		
Número do Parecer: 2.023.665		
Apresentação do Projeto:		
Trata-se de pesquisa em nível de mestrado, envolvendo como voluntários 64 estudantes universitários. A pesquisa envolve a condução de dois experimentos sobre o papel da gesticulação na compreensão e na produção de narrativas em língua inglesa e foi conduzida em um laboratório da universidade, com sala preparada para a condução do estudo. Os participantes assistiram a 2 vídeos em inglês e responderam a perguntas em português. Foram rastreados os movimentos oculares dos participantes durante a veiculação dos vídeos.		
Objetivo da Pesquisa:		
Avaliar o papel da gesticulação na compreensão e produção de narrativas por aprendizes de língua estrangeira.		
Avaliação dos Riscos e Benefícios:		
O projeto apresenta o mesmo nível de risco de uma atividade escolar de ensino de línguas, ou seja, risco mínimo. O pesquisador se compromete a manter o sigilo dos dados. Quanto aos benefícios, conforme informado no TCLE, seus resultados poderão auxiliar a "compreender melhor o papel dos gestos no processo de aprendizagem de língua estrangeira e pode contribuir para estratégias de ensino e também de aprendizagem."		
Comentários e Considerações sobre a Pesquisa:		
Trata-se de pesquisa de interesse científico e social, e que não apresentou intercorrências até o		
Endereço: Av. Prof. Mello Moraes, 1721 - Bl. "G" sala 27 Bairro: Cidade Universitária UF: SP Telefone: (11)3091-4182 Município: SAO PAULO CEP: 05.508-030 E-mail: ceph.ip@usp.br		

ANEXO 2

USP- INSTITUTO DE
PSICOLOGIA DA
UNIVERSIDADE DE SÃO

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: O processamento visual de gestos manuais no ensino de ciências

Pesquisador: Charles Desena Pereira

Área Temática:

Versão: 2

CAAE: 19335518.8.0000.5561

Instituição Proponente: UNIVERSIDADE DE SAO PAULO

Patrocinador Principal: UNIVERSIDADE DE SAO PAULO

DADOS DO PARECER

Número do Parecer: 3.699.206

Apresentação do Projeto:

Nas duas últimas décadas, produziu-se um enorme corpus de evidências da contribuição da gesticulação manual para o ensino-aprendizagem de diversas áreas do conhecimento. Algumas abordagens teóricas têm buscado explicar como padrões sensório-motores de movimentos corporais espontâneos integram a cognição e a comunicação humanas. Em uma abordagem incorporada da cognição humana, que enfatiza a capacidade do corpo em moldar nossa percepção e consciência, gestos manuais têm atraído especial interesse, pois seriam movimentos corporais espontâneos produzidos no ato da fala capazes de enriquecerem representações neurais lexicais, de operarem dialogicamente com a fala, atuando diretamente em mecanismos cognitivos de elaboração de sentido. Com a tecnologia hoje disponível, é possível acompanhar em tempo real processamentos cognitivos ocorrendo no sistema nervoso central. No entanto, pouco se sabe ainda sobre os processos neurocognitivos de integração linguística dos gestos. A partir de dados obtidos experimentalmente em trabalhos anteriores por nosso grupo de pesquisa, hipotetizamos que gestos manuais acompanhados da fala sejam processados majoritariamente por regiões periféricas da retina e pela via dorsal do sistema visual. Buscamos abordar inicialmente nossa hipótese avaliando experimentalmente durante a apresentação de uma aula de ciências o impacto da gesticulação manual da professora em ambos o comportamento ocular do aprendiz e o seu desempenho em tarefa de aprendizagem.

Hipótese:

Endereço: Av. Prof. Mello Moraes, 1721 - Bloco G - Sala 27

Bairro: Cidade Universitária

CEP: 05.508-030

UF: SP

Município: SAO PAULO

Telefone: (11)3091-4182

E-mail: ceph.ip@usp.br

ANEXO 3

COMPORTAMENTO ANIMAL

Leila Miguel Strávale
Instituto de Psicologia - USP

1

INTRODUÇÃO

Etologia
↓
Ciência que estuda o comportamento animal

Karl von Frisch
Konrad Lorenz

2

INTRODUÇÃO

Comportamento: é uma reação em forma de resposta das condições internas e externas do animal, em relação aos estímulos bióticos e abióticos do ambiente

3

TÓPICOS DA AULA

- Tipos de Comportamento
- ✓ Padrão Fixo de Ação
- ✓ Comportamento adquirido ou aprendido

Definições e exemplos

4

Padrão Fixo de Ação

- Comportamento inato: memória evolutiva da espécie.

5

Mecanismos desencadeadores inatos

Resistente às alterações

Padrão Fixo de Ação

Comportamento rigidamente programado

Comportamento no vazio

6

TIPOS DE COMPORTAMENTO Padrão Fixo de Ação

- Mecanismos desencadeadores inatos

Oulterio Larus argentatus

7

TIPOS DE COMPORTAMENTO Padrão Fixo de Ação

- Resistente às alterações

Padrão de comportamento onde o indivíduo repete sua ação mesmo que haja alguma alteração no local.

Pogonius curvipes

8

TIPOS DE COMPORTAMENTO Padrão Fixo de Ação

- Tipo: rigidamente programado

Padrão de comportamento onde o indivíduo executa-o apenas uma vez na vida.

Arachnida Opisthothelphusa

9

TIPOS DE COMPORTAMENTO Padrão Fixo de Ação

- Comportamento no vazio

Padrão de comportamento onde o indivíduo realiza uma ação mesmo que em um local inapropriado (vazio).

Equus caballus

10

Comportamento adquirido ou aprendido

- Resultado das experiências ao longo da vida

11

Habituação

Imitação

Condicionamento clássico

Imprinting

Ensinamento

Condicionamento operante

Comportamento adquirido ou aprendido

12

TIPOS DE COMPORTAMENTO Comportamento adquirido ou aprendido

- Habituação

Comportamento onde o indivíduo se habitua com alguma ação que acontece repetidas vezes.

Helix aspersa

13

TIPOS DE COMPORTAMENTO Comportamento adquirido ou aprendido

- Imprinting

Comportamento onde o indivíduo aprende algo após o nascimento e isto fica "impresso" em seu cérebro por toda a vida.

Pardal & Cinza-branca

14

TIPOS DE COMPORTAMENTO Comportamento adquirido ou aprendido

- Imitação

Comportamento onde o indivíduo imita a ação de outro, mesmo que isso não tenha uma função para ele.

Macaca fasciata

15

TIPOS DE COMPORTAMENTO Comportamento adquirido ou aprendido

- Ensinamento

Comportamento repetido entre indivíduos que sugira algum benefício.

Panthera leo

16

TIPOS DE COMPORTAMENTO Comportamento adquirido ou aprendido

- Condicionamento clássico

17

TIPOS DE COMPORTAMENTO Comportamento adquirido ou aprendido

- Condicionamento operante

18

Bibliografia

- ALCOCK, J. 2005. Animal behaviour. An evolutionary approach. 8ª ed. Sinauer Associates, Inc. 564 p.
- LORENZ, K. 1993. Os fundamentos da etologia. Editora UNESP, 466 p.

19

ANEXO 4

Nome: _____

Questionário Final

1. Na maioria das vezes, qual é a sua facilidade para aprender?

( ) nenhuma   ( ) pouca   ( ) moderada   ( ) intensa

2. Você conhecia o conteúdo da aula assistida?

( ) sim   ( ) não

Se sim, você considera o seu domínio prévio sobre o assunto:

( ) parcial- pequeno   ( ) parcial-médio   ( ) parcial-elevado   ( ) completo

3. Alguma parte da aula foi mais difícil de entender?

( ) sim   ( ) não

Se sim, qual?

_____

4. Qual é o nível de dificuldade da aula para você?

( ) baixa   ( ) média   ( ) alta

5. Você gostou da experiência de participar desse estudo?

( ) sim   ( ) não

6. Qual é a importância de estudar o comportamento animal?

_____
_____
_____
_____
_____
_____

7. Dê uma vantagem do comportamento inato em relação ao aprendido.

_____
_____
_____
_____
_____
_____

8. Leia os textos abaixo e classifique-os em comportamentos inatos ou aprendidos. Se possível, coloque o tipo de comportamento a que se refere:

- a. Ao ver um ovo fora do ninho, o ganso fêmea inicia um movimento repetido de arrastar o ovo com o bico e o pescoço. No entanto, se o ovo escapar ou se o experimentador retirá-lo, o ganso continua a efetuar os movimentos estereotipados mesmo na ausência do ovo até chegar ao ninho, quando então tem de começar tudo de novo. (Retirado de Etologia – Introducción Al Estudio Comparado de Comportamiento)

Resposta: _____

- b. Um sapo jovem vai se alimentar e captura uma libélula, depois uma mosca e por último uma abelha. Depois de um período, esse mesmo sapo passa a se esquivar de moscas e abelhas, alimentando-se apenas de libélulas.

Resposta: _____

- c. Logo que nascem, os filhotes de algumas aves estão aptos a reconhecer os pais. Fez-se um experimento em que os filhotes viram uma ave de papelão ao abrir os olhos pela primeira vez. Essa ave de mentira foi tomada como os parentais daqueles filhotes por toda a vida dos mesmos.

Resposta: _____

- d. Esquilos adultos abrem as nozes diferentemente dos jovens. Os adultos roem a casca em um ponto específico e ela logo se abre. Os jovens demoram um tempo maior e roem a casca de forma mais desordenada.

Resposta: _____

9. Identifique no seu repertório comportamental do dia-a-dia um exemplo de comportamento especificado na aula.

---

10. Crie um exemplo fictício que envolva de um a dois tipos de comportamento mostrados durante a aula.