

A bi-estabilidade nas representações de movimento humano a partir de estímulos visuais e auditivos

CATARINA MENDONÇA (*)

JORGE A. SANTOS (*)

INTRODUÇÃO

Os primeiros trabalhos científicos em percepção de movimento biológico (Johansson, 1973) permitiram a constatação de que um estímulo simplificado de movimento humano, denominado *Point-Light Walker* (PLW), produzia uma impressão vívida e imediata de uma pessoa em acção. Ou seja, representações com pontos de luz das articulações de uma pessoa em andamento são suficientes para proporcionar uma imagem clara de um humano em movimento. Contudo, esta imagem dá lugar à observação de um conjunto sem sentido de pontos quando o PLW está parado. Desde então, o PLW tem permanecido como o estímulo por excelência no estudo da percepção do movimento biológico, não só por permitir uma boa síntese e controlo de variáveis, como também pelas suas propriedades perceptivas e robustez.

A acção humana não é processada da mesma forma que o movimento de um objecto rígido. A sensibilidade ao movimento biológico aumenta

com o número de pontos luminosos associados às articulações, muito mais rapidamente que na detecção de outros padrões de movimento (Neri, Morrone, & Burr, 1998). Por outro lado, esta informação é integrada durante intervalos de tempo 8 vezes superiores aos observados no movimento de objectos rígidos. Uma boa detecção é possível até em condições altamente degradadas, como na diminuição do número de frames, na alteração das trajectórias de pontos ou na introdução de um fundo constituído por pontos similares em movimento aleatório (Blake & Shiffrar, 2006). Mesmo com os movimentos articulares misturados ou incongruentes e a velocidades lentas, há uma percepção vívida de um sujeito em movimento (Beintema, Olesiak, & Wezel, 2006).

O estudo da percepção do movimento biológico a partir do uso das metodologias descritas tornou possível conhecer a informação que pode ser deduzida a partir da observação do movimento de uma pessoa. Sabe-se hoje que a partir da visualização de um PLW é possível perceber a actividade desenvolvida, o género do actor (Pollick, Kay, Heim, & Stringer, 2005; Troje, 2002), a sua identidade (Jacobs & Shiffrar, 2005; Loula, Prasad, Harber, & Shiffrar, 2005; Richardson & Johnston, 2005; Troje,

(*) Escola de Psicologia, Universidade do Minho, Campus de Gualtar, 4700 Braga, Portugal; Telefone: 253604240; Fax: 253604224; E-mail: catarina.mendonca@hotmail.com / jorge.a.santos@psi.uminho.pt

Westhoff, & Lavrov, 2005) e até estados emocionais (Dittrich, Troscianko, Lea, & Morgan, 1996).

A robustez deste tipo de estímulo e as suas propriedades são melhor compreendidas à luz de dados neurofisiológicos. Os primeiros dados que indicaram a existência de uma área cerebral específica para a codificação de estímulos de movimento humano resultaram de estudos com primatas. Nestes estudos, observou-se que neurónios do sulco temporal superior (STS) respondiam selectivamente a caras (Perrett, Rolls, & Caan, 1982), assim como a formas e movimentos humanos (Perret, Smith, Mistlin, Chitty, & Head, 1985). A área STS é um ponto de convergência das vias visuais ventral e dorsal, com funções de processamento de forma e de movimento, respectivamente, e tem ligações com a amígdala e com o córtex orbitofrontal, regiões implicadas no processamento de estímulos de importância social e emocional (Puce & Perrett, 2003).

Num estudo com tomografia de emissão de positrões (PET), identificou-se que área posterior do STS (STSp) responde apenas quando as pessoas vêem PLW coerentes (Bonda, Petrides, Oery, & Evans, 1996). Na mesma área, observou-se uma activação mais forte para os PLW de pé que para os invertidos (Grossman, Blake, & Kim, 2004). Tanto figuras humanas estáticas, como movimentos rígidos complexos produzem pouca ou nenhuma activação da área STSp (Beauchamp, Lee, Haxby, & Martin, 2002; Peuskens, Smith, Mistlin, Chitty, & Head, 2005). As células desta área que respondem ao movimento de corpo inteiro são selectivas para a direcção do movimento e activam-se maioritariamente perante movimentos de corpo apresentados em orientações frontais, ainda que algumas células respondam a movimentos de costas (Puce & Perrett, 2003). Esta selectividade de activação celular poderá estar na base do viés perceptivo encontrado para posições frontais, e que será analisado em maior detalhe mais abaixo.

A área STSp parece responder não só a imagens, como a sons relacionados com movimento biológico. Num estudo com ressonância magnética (fMRI), observou-se a activação desta área durante a percepção auditiva de vários tipos de passos humanos, mas não para sons não

relacionados (Bidet-Caulet, Voisin, Bertrand, & Fonlupt, 2005). Esta descoberta levou à formulação da hipótese de que a área STSp poderá fazer parte de uma rede que abrange tudo quanto está relacionado com o movimento biológico, com funções supramodais para a integração a um nível superior de estímulos relevantes para a cognição da acção humana (Bidet-Caulet et al., 2005; Blake & Shiffrar, 2006). Atendendo a estes dados, pode hipotetizar-se que o movimento biológico seja não só funcionalmente, como também perceptivamente processado de forma própria. Torna-se, então, relevante estudar a dinâmica dos estímulos visuais e auditivos na percepção de movimento biológico, de modo a compreender o impacto perceptivo desta especialização cerebral.

Nos estudos de percepção bimodal (estimulação visual e auditiva) clássicos, frequentemente o estímulo visual altera as propriedades do estímulo auditivo, influenciando a percepção final. Em tarefas de percepção de movimento, os sujeitos falham frequentemente na discriminação da direcção da fonte sonora, quando acompanhada de movimento visual na direcção oposta. Por sua vez, a identificação do movimento visual não é afectada pela direcção do movimento auditivo (Soto-Faraco, Spence, & Kingstone, 2004). Também em tarefas que simulam o efeito de ventriloquismo, observa-se que os sujeitos tendem a redefinir a localização percebida da fonte sonora de forma a encaixar com a fonte visual, assim como tendem a distorcer as propriedades do som, passando a percebê-lo como a mover-se na mesma direcção do estímulo visual (Soto-Faraco, Spence, Lloyd, & Kingstone, 2004).

Observam-se, por outro lado, situações em que é o estímulo auditivo a influenciar as fontes de informação visual. São, disso exemplo, os estudos que exploram o fenómeno de ventriloquismo temporal, em que um som apresentado em grande proximidade temporal pode alterar as dimensões duracionais percebidas do estímulo visual (Vroomen & Gelder, 2004). Também são notáveis as influências do estímulo auditivo sobre o visual em procedimentos experimentais em que os sujeitos devem bater um dedo de acordo com o ritmo de estímulos visuais e auditivos, sendo que os segundos têm sempre

grande influência (por vezes não percebida) sobre o ritmo reproduzido, em detrimento dos primeiros (Repp, 2003; Repp & Pennel, 2002). Recentemente, Watkins, Shams, Tanako, Haynes, e Rees (2006) observaram ainda que a apresentação de um flash visual breve acompanhada de dois *bips* sonoros é frequentemente percebida incorrectamente como dois flashes visuais.

De acordo com os dados actuais, considera-se que interacções multisensoriais da percepção de movimento se caracterizam pelo domínio, ao invés da interferência, de uma modalidade sensorial sobre a outra (Soto-Faraco, Spence, Lloyd, & Kingstone, 2004). Repp e Pennel (2002) analisaram dados provenientes de vários estudos e concluíram que os estímulos visuais são consistentemente dominantes em tarefas de percepção espacial, deturpando as propriedades dos estímulos auditivos. Observaram, simultaneamente, que os estímulos auditivos são mais fortes em tarefas de percepção temporal.

Os dados até agora disponíveis acerca da multimodalidade na percepção do movimento biológico levam a supor que estes processos podem ser diferentes. Estudos com estímulos biológicos sugerem que as pistas podem, quando incompatíveis, ser integradas, formando um novo percepto. No efeito McGurk (McGurk & Mac Donald, 1976), enquanto os sujeitos ouviam uma voz humana dizer /ba/ e viam uma face humana dizer /ga/, o percepto final era frequentemente /da/.

Do ponto de vista do movimento biológico, pouco se sabe acerca destas interacções. Num estudo em que se avaliava a magnitude de saltos, a condição em que foram emparelhados estímulos visuais e auditivos congruentes produziu melhores resultados que as condições visual e auditiva, separadamente (Effenberg, 2005). Também, um trabalho recente investigou os correlatos comportamentais da integração audiovisual no processamento de pistas de movimento biológico (Brooks, Zwan, Billard, Petreska, Clarke, & Blanke, 2007). Os resultados indicaram a existência de efeitos selectivos de direcção do som na detecção do movimento biológico. Com efeito, quando comparada com a condição de som estacionário, a condição de movimento sonoro na mesma direcção do

movimento visual obteve melhores resultados de detecção. Por outro lado, perante movimento sonoro e visual em direcções opostas, observou-se o efeito inverso, com o aumento significativo dos tempos de detecção. Esta facilitação/inibição bimodal específica para a detecção de movimento audiovisual não foi encontrada em ensaios de controlo, com o PLW invertido.

Alais e Burr (2004) realizaram um estudo metodologicamente semelhante ao de Brooks e colaboradores (2007), com movimento rígido. Neste estudo, os efeitos de facilitação não se verificaram, ou seja, não se encontraram melhorias para além do esperado pela mera combinação estatística dos sinais perante o movimento na mesma direcção. Por sua vez, o som em movimento oposto ao visual permitia resultados semelhantes aos encontrados quando ambos os estímulos se deslocavam na mesma direcção, não se observando o efeito de inibição bimodal.

Analisados em conjunto, estes dados sugerem a existência de um mecanismo específico para a integração de estímulos associados ao movimento humano.

O estudo descrito no presente trabalho teve como objectivo observar os processos envolvidos na percepção audiovisual do movimento biológico. Esperava-se observar pouca predominância sensorial e maior tendência para a integração dos estímulos visual e auditivo. Ou seja, esperava-se que os sujeitos não se centrassem apenas numa pista sensorial, em detrimento da outra. Também se procurava observar na condição audiovisual taxas de acerto mais elevadas que nas condições unimodais visual e auditiva.

De forma a testar estas hipóteses, optou-se por fazer uso de estímulos visuais ambíguos e bi-estáveis, que podem ser percebidos como compatíveis ou como incompatíveis com os estímulos sonoros. Com efeito, esta é uma propriedade possível dos PLW, que foi alvo de estudo por Vanrie, Dekeyser, e Verfaillie (2004). Por consistirem em representações empobrecidas do movimento biológico, os PLW não contêm pistas estruturais ou dinâmicas suficientes para permitir sempre uma identificação correcta da orientação do sujeito em marcha. Assim, um PLW orientado de frente

(0°) para o observador pode ser confundido com um PLW de costas (180°), e o mesmo acontece para as animações semi-laterais, em que orientações de 45° são semelhantes às de 135°. Varie, Dekeyeser, e Verfaillie (2004) constataram que, em estímulos totalmente bi-estáveis (foram equalizados pontos de referência como o alinhamento dos ombros, ancas, joelhos e pés), os sujeitos apresentavam um forte viés perceptivo no sentido de observar os avatares em orientações frontais: em cerca de 80 por cento dos ensaios, o PLW era visto a mover-se em direcção ao observador, sendo que apenas nos restantes 20 por cento dos casos era visto na orientação alternativa (de costas). Na condição de controlo (PLW invertidos), os valores mantiveram-se ao nível do acaso, em 50 por cento dos ensaios era visto de frente e nos restantes era visto de costas. No mesmo estudo, testou-se a robustez deste fenómeno perceptivo. Numa experiência, foram apresentados PLW com uma animação que avançava do fim para o início do ciclo de passada, ou seja, dava a sensação de que o estímulo de movia em afastamento. Nesta experiência, os sujeitos obtiveram taxas de acerto quanto à direcção do movimento do estímulo de 98 por cento, mas em 89 por cento das vezes em que responderam correctamente reportaram que os PLW se mantinham virados de frente, ou seja, frequentemente pareciam andar de costas para onde se dirigiam. Também numa terceira experiência, testou-se a robustez deste viés pela introdução de pistas: numa condição, introduziu-se a oclusão dos pontos luminosos (estes desapareciam sempre que eram escondidos pelo corpo) e na outra condição usou-se a projecção de perspectiva. A proporção de respostas correctas quanto à orientação do estímulo passou a ser de 85 por cento na primeira condição e de 97 por cento na segunda. Conclui-se, assim, que apesar de robusto, o viés visual é sensível a pistas espaciais.

Fazendo uso deste estímulo bi-estável, procurou-se, neste trabalho, observar o impacto de estímulos auditivos congruentes (cuja orientação de movimento era semelhante à do movimento apresentado visualmente) sobre as representações visuais fortemente enviesadas. Deste modo, num primeiro momento, este

trabalho pretendeu obter uma replicação dos dados de Vanrie, Dekeyeser, e Verfaillie (2004), para num segundo momento analisar o efeito das pistas auditivas sobre os estímulos visuais de movimento biológico.

MÉTODO

Sujeitos

Usaram-se 7 participantes, todos dextros, 3 homens e 4 mulheres, com idades compreendidas entre os 23 e os 28 anos. Os sujeitos eram todos ingénuos quanto aos objectivos do estudo e nunca haviam contactado com os estímulos usados. Num momento prévio às experiências, todos os participantes fizeram testes visuais e auditivos, de forma a assegurar que tinham visão normal ou corrigida, assim como uma boa audição, sem discrepâncias de acuidade interaurais.

Estímulos e materiais

Os estímulos visuais utilizados consistiram em PLW de 13 pontos luminosos animados com coordenadas 3D, construídos a partir da base de dados de Vanrie & Verfaillie (2004). Os pontos luminosos eram esferas brancas de 1cm de diâmetro (em média) em movimento, com uma luminância de cerca de 68cd/m² sobre um fundo cinza escuro de aproximadamente 1.53cd/m². Os estímulos eram projectados sobre uma tela a 60Hz, ocupando uma área de projecção de 4.94mx2.10m. Os avatares ocupavam, dos pés à cabeça, um ângulo visual de 5.72° para o observador, que estava sentado a 3m da tela.

A apresentação e manipulação informática de todos os estímulos foi feita a partir de uma aplicação informática baseada em Open GL sobre VR/Net Juggler. Trabalhou-se a com um cluster de três computadores com placa gráfica Nvídia Quadro FX 4500. Cada um dos computadores estava ligado a um canal de imagem, emitido por projectores DLP de 3 chips Christie Mirage S+4k, sendo que apenas se usaram dois canais nesta experiência. Cada canal tinha uma resolução de 1400x1050 pixel. A imagem dos dois canais era projectada na tela,

em superfície contínua, com uma área central de blending das duas projecções.

Durante as apresentações, os PLW moviam-se, como se estivessem a andar em cima de uma passareira rolante (sem translação). Utilizaram-se 18 tipos de estímulos diferentes, correspondentes a 6 orientações de PLW apresentadas nas condições visual, auditiva e audiovisual. Usaram-se as orientações frente (0°), trás (180°), em aproximação para a esquerda e para a direita (45°E e 45°D , respectivamente) e em afastamento para a esquerda e para a direita (135°E e 135°D , respectivamente).

Os estímulos auditivos consistiram em seis tipos de ficheiros de som em que se ouvia um total de oito passos (quatro ciclos de passada). Estes estímulos começavam sempre o ensaio num ponto central com relação à cabeça do sujeito, e posteriormente deslocavam-se em aproximação ou afastamento e, nos ensaios de orientações 45°E , 135°E , 45°D e 135°D , para a esquerda ou para a direita. Estes ficheiros de som foram gerados a partir de software baseado em Open AL e posteriormente gravados, com frequências que variaram entre os 20Hz e os 10.000Hz, concentrando-se a maior parte da informação auditiva apresentada em torno dos 1.000Hz. No contexto experimental, usou-se a placa de som Realtec Intel 8280 IBA, e o som era emitido por colunas colocadas nos pontos extremos da imagem (a 80cm do limite visual) e a 50cm do chão, orientadas para a cabeça dos sujeitos.

Em todos os ensaios audiovisuais, o estímulo sonoro deslocava-se na mesma direcção que o estímulo visual e ambos os estímulos eram apresentados em fase (sincronizados).

Design e procedimentos

Os dezoito tipos de estímulo foram apresentados com cinco ensaios cada um segundo o método do estímulo constate. Cada apresentação teve a duração total de 4s, a mesma usada no estudo de Vanrie, Dekeyser, e Verfaillie (2004), correspondente a quatro ciclos de passada por ensaio. O intervalo entre ensaios teve a duração de 3s.

Durante a experiência, os participantes deviam responder manualmente, apontando

numa tabela octogonal qual a direcção em que o estímulo se movia.

Os participantes foram testados individualmente, numa sala escurecida, em frente à tela onde eram projectados os estímulos. Estavam sentados no correspondente a um ponto intermédio da área de projecção. À sua frente, os participantes tinham uma mesa de superfície negra, com a tabela octogonal. O experimentador estava sentado à esquerda dos indivíduos, ligeiramente recuado, de forma a não integrar a área de visualização dos mesmos.

Os sujeitos eram informados acerca da natureza dos estímulos: um conjunto de pontos luminosos que permitem a impressão de um sujeito humano em marcha, sons de passada, ou ambos em simultâneo. Posteriormente, era apresentada a tabela octogonal. Todos os sujeitos foram instruídos para se centrarem numa cruz branca apresentada durante o período entre ensaios e responderem atempadamente, e foíhes dito que era permitido alterar o sentido das respostas durante os ensaios, se apercebessem de que tinham cometido um erro. Pediu-se que respondessem sempre, mesmo em caso de dúvida. Seguidamente, iniciava-se a sequência experimental.

RESULTADOS

Quanto à distribuição das respostas por orientação, na condição visual, observou-se um forte viés frontal, sendo que em cada cinco ensaios para a frente e cinco ensaios para trás, os sujeitos responderam em média oito vezes para a frente ($\delta=2,06$) e duas vezes para trás ($\delta=2,07$) (cf. Figura 1). Na condição auditiva observou-se uma distribuição da orientação de resposta não enviesada, com igual número médio de respostas frente e trás ($\delta=1,36$).

Na condição audiovisual, este viés foi largamente atenuado, mas não desapareceu. Nesta condição, os sujeitos responderam em média quatro vezes para trás ($\delta=1,51$) e seis vezes para a frente ($\delta=1,46$) (cf. Figura 2).

A redução de viés que se observou na condição audiovisual traduziu-se também num aumento claro do número de acertos.

FIGURA 1

Distribuição das respostas dadas em média pelos sujeitos nos cinco ensaios de cada orientação, na condição visual

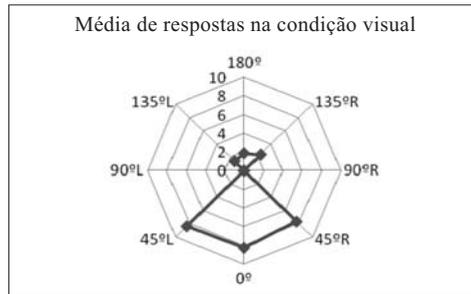


FIGURA 2

Distribuição das respostas dadas em média pelos sujeitos nos cinco ensaios de cada orientação, na condição audiovisual



A condição com maior discrepância quanto aos acertos nas diferentes orientações foi a visual (cf. Figura 3). Nesta condição, observam-se taxas de acerto muito reduzidas nas orientações 135° (0,3) e 180° (0,4), que contrastam com a facilidade dos ensaios de 0° (taxa de acerto de 1), nos quais não se observaram erros. Naturalmente, estas diferenças notórias estão relacionadas com os dados acima apresentados acerca do forte viés visual, sendo possível que, tendo respondido predominantemente para a frente, os sujeitos tenham acertado nos ensaios que efectivamente estavam orientados para a frente e tenham errado mais nos ensaios que estavam orientados para trás. Não se observaram, nesta condição, erros por confusão lateral, ou seja, os erros deveram-se sempre às confusões em profundidade 0°/180°, 45°E/135°E e 45°D/135°D.

Na condição auditiva, observou-se igualmente uma ligeira assimetria entre as taxas de acerto nos ensaios orientados para a frente (0,86) e nos ensaios orientados para trás (0,62), sendo os últimos mais difíceis que os primeiros. Esta pequena diferença poderá ser facilmente explicada pelo facto de, nos ensaios a 180° e a 135°, haver uma clara diminuição da intensidade do estímulo auditivo ao longo da apresentação. Nestes ensaios, é importante referir que 39 dos 46 erros registados se deram por confusão lateral (180°, 135°E e 135°D) e apenas 7 se deveram a confusão entre as orientações para a frente ou para trás. Assim, será relativamente seguro assumir que os estímulos auditivos tinham boas propriedades perceptivas quanto à discriminação de aproximação/afastamento, podendo esperar-se uma acção de desambiguação sobre os estímulos

visuais enviesados. Observou-se igualmente, na condição auditiva, que os ensaios oblíquos obtiveram melhores resultados (0,81) que os ensaios fronto-paralelos (0,51), tendo os melhores resultados surgido nos ensaios de 45° (0,89) e os piores resultados surgido nos ensaios de 180° (0,51). Este dado poderá dever-se ao facto de, nos ensaios oblíquos, haver uma pista adicional devida às diferenças de estimulação interaural, eficaz na diminuição das confusões laterais.

A condição audiovisual obteve resultados globalmente melhores (acerto médio de 0,87), por comparação com as condições unimodais (0,6 na condição visual e 0,76 na condição auditiva). Adicionalmente, foi esta a condição em que se observaram menores discrepâncias entre orientações. Os melhores resultados desta condição surgiram, como esperado, nos ensaios cujos estímulos de moviam de frente para o observador. Nas orientações oblíquas, observaram-se ganhos relevantes quanto às taxas de acerto, mas foi nas orientações fronto-paralelas que se observaram os efeitos mais interessantes. Com efeito, foi nos ensaios a 180° que se observou o efeito de interação entre as pistas visuais e

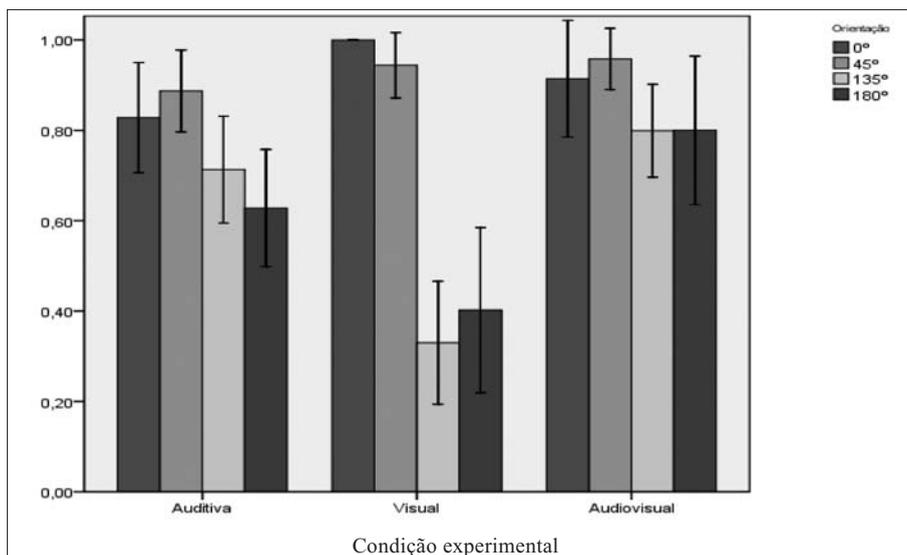
auditivas mais forte, com um aumento muito expressivo nas taxas de acerto: de 0.4 na condição visual e de 0.51 na condição auditiva, obteve-se 0.8 na condição audiovisual. Poder-se-á afirmar, neste caso, que se observou uma diminuição da confusão lateral nos estímulos auditivos por acção dos estímulos visuais, a par da diminuição do viés visual por acção das pistas auditivas.

Mas também nos ensaios de 0° surgiu um efeito de interação audiovisual notório. Nestes ensaios, a condição audiovisual não obteve melhores resultados que a condição unimodal visual, que tinha atingido um valor máximo.

De acordo com o teste de Friedman, a diferença de acertos entre as três condições é estatisticamente significativa ($\chi^2=8,67$, $gl=2$, $p\leq 0.05$). Através da análise dos resultados emparelhados no teste U de Mann-Whitney compararam-se também os dados das condições duas a duas. Este teste confirmou que tanto a condição auditiva como a visual se distinguem da audiovisual ($p\leq 0.01$ e $p\leq 0.05$, respectivamente). Os resultados de ambas as condições unimodais são igualmente diferentes ($p\leq 0.5$).

FIGURA 3

Taxa de respostas certas nas condições visual, auditiva e audiovisual, nas orientações de 0°, 45°, 135° e 180°



DISCUSSÃO

Num primeiro momento, o presente trabalho pretendia uma replicação dos dados observados por Vanrie, Dekeyeser, e Verfaillie (2004) quanto ao forte viés encontrado em estímulos de movimento biológico potencialmente bi-estáveis. Naquele estudo, os autores tinham constatado uma tendência na ordem dos 80 por cento de respostas orientadas de frente para o observador. Contudo, este efeito era quase eliminado quando os investigadores introduziam pistas de perspectiva e tridimensionalidade. Neste estudo, os dados confirmaram em larga medida o esperado, tendo-se observado sempre taxas de viés na ordem dos 80 por cento. Estas taxas, contudo, não foram diminuídas pelas pistas de tridimensionalidade. Numa tentativa explicativa, poder-se-á argumentar que estas pistas continuaram atenuadas, nomeadamente pela dimensão menor dos estímulos visuais apresentados. Com efeito, os ângulos visuais usados originalmente no estudo de Vanrie, Dekeyeser, e Verfaillie (2004) foram de 7.87° dos pés à cabeça do avatar, enquanto os aqui usados foram de 5.72° .

Também os dados de Puce e Perrett (2003) podem ser elucidativos. Como foi descrito na introdução, estes autores encontraram uma activação preferencial de um maior número de células neuronais da área STSp para estímulos orientados para o observador, e um menor número de células activadas perante estímulos de costas. Estes dados podem indicar que o viés encontrado deriva de um processo perceptivo determinado fisiologicamente e de baixo nível, impermeável a factores desambiguadores de menor intensidade. Assim, observa-se que, apesar de sensível a pistas espaciais, o fenómeno de viés perceptivo nas representações bi-estáveis de movimento biológico se mantém quando estas são atenuadas.

A um segundo nível, este trabalho procurava analisar o efeito de pistas auditivas sobre os estímulos visuais de movimento biológico. De acordo com a primeira hipótese enunciada, esperava-se observar um menor efeito de predominância sensorial e uma maior tendência para a integração de estímulos incompatíveis. Os sujeitos não se deveriam centrar apenas numa

pista, em detrimento da outra, à semelhança do observado em variados estudos com movimento rígido, mas deveriam demonstrar atenção a ambas as fontes de informação. Atendendo à perspectiva apresentada por Repp e Pennel (2002), numa tarefa deste tipo (espacial) com movimento rígido, o estímulo visual deveria ser claramente predominante, e os resultados esperados poderiam ser semelhantes àqueles encontrados nos estudos que simulam o efeito de ventriloquismo. Ou seja, seria de esperar que o som adquirisse propriedades perceptivas semelhantes às da imagem, parecendo mover-se no mesmo sentido dos avatares. Apesar dos fortes efeitos de enviesamento frontal, esse efeito nunca se verificou. Todos os dados observados contrariaram esta possibilidade. Se o estímulo visual tivesse adquirido propriedades de referência, o viés visual deveria ter-se mantido na condição audiovisual. O que de facto se observou foi, pelo contrário, uma forte redução deste mesmo viés, indicadora de que não só os estímulos auditivos não pareceram mover-se no mesmo sentido dos estímulos visuais, como foram integrados e agiram sobre aqueles.

Por outro lado, poderia argumentar-se que, pela elevada ambiguidade do PLW bi-estável, os sujeitos tenderiam a optar por usar como pista de referência o estímulo auditivo, mais estável. Observa-se, contudo, que os resultados de viés nunca foram totalmente atenuados, o que teria acontecido caso a pista de referência fosse a auditiva (clara quanto à indicação de movimento em profundidade frente/trás).

Uma das discussões mais activas entre os investigadores que estudam a percepção multimodal de estímulos não biológicos prende-se, precisamente, com o momento que as pistas audiovisuais são combinadas. Mais concretamente, alguns investigadores argumentam que há um momento de processamento comum de baixo nível, que seria responsável por efeitos como a distorção de um input de forma a se assemelhar ao outro, apresentado concomitantemente (como nos efeitos de ventriloquismo visual e temporal). Outros autores argumentam que a elevada variedade encontrada nos resultados desta área poderá dever-se sobretudo ao tipo de tarefa que é colocada e que a combinação dos estímulos só é feita a um nível decisional, de alto nível, levando

os sujeitos a centrarem-se sempre nos estímulos mais apropriados e cuja informação é mais confiável, quanto ao que é pedido (Soto-Faraco, Spence, & Kingstone, 2004; Welsh, Duttonhurt, & Warren, 1986). Dados recentes de fMRI indicam que os processos perceptivos e decisórios na percepção audiovisual de movimento rígido estão intimamente ligados (Soto-Faraco, Spence, Lloyd, & Kingstone, 2004).

Nos dados do presente estudo, não se observam os mesmos fenómenos de escolha de uma pista sensorial. Parece então que a percepção multimodal do movimento biológico pode não seguir as mesmas regras da percepção multimodal de movimento rígido.

A segunda hipótese que se colocava neste trabalho indicava que a condição audiovisual apresentaria melhores resultados que as condições unimodais. Este efeito foi claramente confirmado nos resultados obtidos, sendo as diferenças encontradas entre as condições estatisticamente significativas. Os ensaios em que surgiu a mais forte interacção foram aqueles de orientação a 180. Para este tipo de estímulos, as proporções de acerto eram bastante reduzidas e sofreram um forte aumento na condição audiovisual. É particularmente interessante observar estes resultados neste tipo de ensaios, em que os estímulos visuais e auditivos eram claramente incongruentes, uma vez que o viés fazia o avatar parecer andar de frente e o som indicava um claro movimento de afastamento.

De acordo com o modelo interpretativo de Brooks e colaboradores (2007), quando os estímulos de movimento biológico eram congruentes entre eles, surgia um efeito de facilitação bimodal com claras melhorias perceptivas, mas quando estes estímulos eram incongruentes, havia um efeito de inibição bimodal, com dificuldade destes processos. Os resultados do presente trabalho parecem indicar que os estímulos de ambas as modalidades sensoriais tendem a ser integrados, mesmo quando incompatíveis (quando há viés visual para a frente e o estímulo auditivo se move para trás). Naturalmente, que devido à natureza biestável dos estímulos visuais e às características pouco robustas dos estímulos auditivos, apesar de estes estímulos serem incongruentes separadamente, era sempre possível encontrar

um percepto integrador de ambos os dados sensoriais. Assim, observa-se que, sendo possível, os sujeitos combinam ambas as pistas, que se influenciam mutuamente. Independentemente da facilitação ou inibição da percepção enquanto processo (no estudo supracitado, Brooks e colegas mediram os resultados em tempos de resposta), os resultados multimodais são melhores que as probabilidades de sucesso unimodais combinadas, podendo-se falar numa facilitação perceptiva enquanto resultado.

Apesar de todas as análises desenvolvidas, mantém-se contudo imprevisível a forma como os estímulos audiovisuais interagem face aos valores obtidos nas condições unimodais. A este nível, serão necessários estudos adicionais, com melhor quantificação das propriedades perceptivas, de forma a obter valores que permitam comparações mais claras.

Em suma, pode-se concluir que a percepção audiovisual do movimento biológico obedece a mecanismos de integração multimodal próprios. Todos os nossos dados podem ser explicados por mecanismos de interacção entre as pistas visuais e auditivas, quando apresentadas em simultâneo, o que suporta a ideia da existência de uma área cerebral de processamento comum, a área STSp, que pode ter propriedades de integração supramodal, prévias ao momento de tomada de decisão.

Os resultados observados não esgotam todas as questões acerca da percepção audiovisual do movimento biológico. Serão necessários estudos adicionais, nomeadamente no sentido de conhecer melhor a interacção entre as pistas auditivas e visuais quando incompatíveis ou incongruentes, assim como será pertinente analisar em detalhe quais os valores de igualdade/discrepância necessários para produzir efeitos de facilitação/inibição bimodal específica e qual a sua relação. Igualmente relevante, será aprofundar os dados neurofisiológicos acerca do papel da área STSp nos processos intermodais de percepção do movimento biológico.

REFERÊNCIAS

- Alais, D., & Burr, D. (2004). No direction-specific bimodal facilitation for audiovisual motion detection. *Cognitive Brain Research*, 19, 185-194.

- Beauchamp, M., Lee, K., Haxby, J., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, *34*, 149-159.
- Beintema, J., Oleksiak, A., & Wezel, R. (2006). The influence of biological motion perception on structure-from-motion interpretations at different speeds. *Journal of Vision*, *6*, 712-726.
- Bidet-Caulet, A., Voisin, J., Bertrand, O., & Fonlupt, P. (2005). Listening to a walking human activates the temporal biological motion area. *Neuroimage*, *28*, 132-139.
- Blake, R., & Shiffrar, M. (2006). Perception of human motion. *Annual Review of Psychology*, *58*, 12.1-12.27.
- Bonda, E., Petrides, M., Oery, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience*, *16*, 3737-3744.
- Brooks, A., Zwan, R., Billard, A., Petreska, B., Clarke, S., & Blanke, O. (2007). Auditory motion affects visual biological motion processing. *Neuropsychologia*.
- Dittrich, W., Troscianko, T., Lea, S., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, *25*, 727-738.
- Effenberg, A. (2005). Movement sonification: Effects on perception and action. *IEE Multimedia*, *2*, 53-59.
- Grossman, E., Blake, R., & Kim, C. (2004). Learning to see biological motion: Brain activity parallels behaviour. *Journal of Cognitive Neuroscience*, *16*, 1669-16679.
- Jacobs, A., & Shiffrar, M. (2005). Walking perception by walking observers. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 157-169.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, *14*, 201-211.
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology*, *31*, 210-220.
- McGurk, H., & Mac Donald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746-748.
- Neri, P., Morrone, M., & Burr, D. (1998). Seeing biological motion. *Nature*, *395*, 894-896.
- Perret, D., Rolls, E., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329-342.
- Perret, D., Smith, P., Mistlin, A., Chitty, A., & Head, A. (1985). Visual analysis of body movements by neurons in the temporal cortex of the macaque monkey: A preliminary report. *Behavioral Brain Research*, *16*, 153-170.
- Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. (2005). Specificity of regions processing biological motion. *European Journal of Neuroscience*, *21*, 2864-2875.
- Pollick, F., Kay, J., Heim, K., & Stringer, R. (2005). Gender recognition from point-light walkers. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1247-1265.
- Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London*, *358*, 435-445.
- Repp, B. (2003). Phase attraction in sensorimotor synchronization with auditory sequences: Effect of single and periodic distractors on synchronization accuracy. *Journal of Experimental Psychology*, *29*, 290-309.
- Repp, B., & Pennel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology*, *5*, 1085-1099.
- Richardson, M., & Johnston, L. (2005). Person recognition from dynamic events: The kinematic specification of individual identity in walking style. *Journal of Nonverbal Behavior*, *29*, 25-44.
- Soto-Faraco, S., Spence, C., & Kingstone, A. (2004). Cross-modal dynamic capture: Congruency effects in the perception of motion across sensory modalities. *Journal of Experimental Psychology*, *30*, 330-345.
- Soto-Faraco, S., Spence, C., Lloyd, D., & Kingstone, A. (2004). Moving multisensory research along: Motion perception across sensory modalities. *Current Directions on Psychological Science*, *13*, 29-32.
- Troje, N. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, *2*, 371-387.
- Troje, N., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception and Psychophysics*, *67*, 667-675.
- Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods*, *36*, 625-629.

- Vanrie, J., Dekeyser, M., & Verfaillie, K. (2004). Bistability and biasing effects in the perception of ambiguous point-light walkers. *Perception, 33*, 547-560.
- Vroomen, J., & Gelder, B. (2004). Temporal ventriloquism: Sound modulates the splash-lag effect. *Journal of Experimental Psychology, 3*, 513-518.
- Watkins, S., Shams, L., Tanako, S., Haynes, J., & Rees, G. (2006). Sound alters activity in human v1 in association with illusory visual perception. *Neuroimage, 31*, 1247-1256.
- Welsh, B., Duttonhurt, D., & Warren, H. (1986). Contributions of audition and vision to temporal rate perception. *Perception and Psychophysics, 39*, 294-300.

RESUMO

Pouco se sabe acerca das interações entre estímulos visuais e auditivos na percepção do movimento biológico. Dados recentes indicam que poderá haver uma área cerebral (STSp), que responde a estimulação de movimento biológico, tanto visual quanto auditiva. É provável que os processos intermodais em movimento biológico tendam para uma maior interação e integração de pistas que em movimento rígido. Neste trabalho, procurou-se observar a natureza dessas interações. Para tal,

utilizaram-se estímulos visuais bi-estáveis, com um forte viés frontal. Compararam-se condições visuais, auditivas e audiovisuais. Os dados revelam efeitos de interação entre estímulos. A condição audiovisual apresenta resultados globalmente melhores na redução do viés perceptivo e nas taxas de acerto.

Palavras-chave: Bi-estabilidade, Integração intermodal, Movimento biológico, Percepção multimodal, Viés.

ABSTRACT

There is not enough information about the interactions between visual and auditory stimuli in biological motion perception. Recent data suggests that there might be a brain area (STSp), which responds both to visual and auditory biological motion stimulation. It is also likely that the intermodal processes in biological motion should tend towards greater interaction and integration of stimuli. In this project, we intended to observe the nature of those interactions. To do so, we used strongly biased bistable visual stimuli. We then compared the visual, auditory and audiovisual conditions in two experiments. The results suggest interaction effects between the different stimuli. The audiovisual condition had the best results, both in the bias reduction and in the proportion of correct answers.

Key-words: Bias, Biological motion, Bistability, Crossmodal integration, Multimodal perception.

