



Universidade do Minho

Instituto de Educação e Psicologia

Ana Catarina Macedo Mendonça

**A Bi-Estabilidade nas Representações de
Movimento Humano a Partir de Estímulos
Visuais e Auditivos**

Outubro de 2007



Universidade do Minho

Instituto de Educação e Psicologia

Ana Catarina Macedo Mendonça

**A Bi-Estabilidade nas Representações de
Movimento Humano a Partir de Estímulos
Visuais e Auditivos**

Tese de Mestrado em Psicologia
Área de Especialização em Ciências Cognitivas

Trabalho efectuado sob a orientação do
Professor Doutor Jorge Manuel Ferreira Almeida Santos

Outubro de 2007

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE,
APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO
ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Ana Catarina Macedo Mendonça

AGRADECIMENTOS

Ao Prof. Doutor Jorge A. Santos;

À Dra. Sandra Mouta;

Ao Dr. Bruno Aragão;

Aos colegas do CCG;

Aos amigos e colegas que participaram voluntariamente no estudo;

À Mary.

TÍTULO: A Bi-Estabilidade nas Representações de Movimento Humano a Partir de Estímulos Visuais e Auditivos

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. Sabe-se também que é provável que os processos intermodais em movimento biológico tendam para uma maior interação e integração de estímulos. Neste trabalho procurou-se observar a natureza dessas interações. Para tal, utilizaram-se estímulos visuais bi-estáveis, com um forte viés para as representações orientadas para o observador. Compararam-se condições visuais, auditivas e audiovisuais em duas experiências, uma sem translação dos estímulos visuais e outra com translação. Os dados revelam efeitos claros de interação entre estímulos. A condição audiovisual apresenta resultados globalmente melhores na redução do viés perceptivo e nas taxas de acerto.

Title: The Bi-stability in Human Motion Representations from Visual and Auditory Stimuli

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 first experiment presented visual stimuli without translation and the second experiment introduced the translation.

The results suggest clear 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.

Índice

1. Introdução.....	6
1.1. Percepção Multimodal do Movimento Biológico.....	6
1.2. Objectivos do Estudo.....	12
2. Experiência 1 – Sem Translação	15
2.1. Método.....	15
2.2. Resultados.....	18
2.3. Discussão.....	23
3. Experiência 2 – Com Translação.....	24
3.1. Método.....	24
3.2. Resultados.....	26
3.3. Discussão.....	31
4. Discussão	
Geral.....	32
5. Referências.....	36

1. INTRODUÇÃO

1.1. Percepção Multimodal do Movimento Biológico

O movimento biológico como objecto de estudo científico remonta a meados do século XIX. Os primeiros registos de aquisição de dados de movimento para fins científicos surgiram quando o fotógrafo britânico Muybridge aceitou o desafio de descobrir se, durante o galope e o trote, os cavalos sustentam por alguns instantes todas as patas acima do chão. Para esse propósito, Muybridge usou um conjunto de vinte e quatro sistemas fotográficos, dispostos ao longo de um hipódromo, e obteve a primeira decomposição sistemática do movimento biológico. Mas foi Étienne-Jules Marey, um contemporâneo de Muybridge, quem pela primeira vez se debruçou sobre a parametrização da locomoção humana. De forma a decompor o movimento biológico, ele inventou um sistema fotográfico que lhe permitiu obter imagens que respeitavam os intervalos temporais e espaciais entre posições. Com o intuito de simplificar estas imagens, eliminar formas corporais e focar-se unicamente nas pistas de movimento, Marey colocou bandas marcadores brancos sobre os membros dos seus sujeitos e fotografou-os sucessivamente. Os métodos de aquisição de estímulos de movimentos biológico utilizados hoje em dia são ainda em larga medida similares aos de Marey.

Em 1973, Johansson publicou os primeiros dados experimentais acerca da percepção do movimento biológico. Até então, apenas se tinha estudado a percepção de movimento rígido. Na aquisição dos estímulos, Johansson desenvolveu dois métodos. No primeiro, colocou lâmpadas sobre um assistente vestido num fato preto justo e filmou-o enquanto percorria uma trajectória linear, numa sala escura. Esta metodologia revelou alguns inconvenientes, nomeadamente, o desconforto de todos os fios e as limitações em mudanças de direcção. O segundo método desenvolvido por Johansson tomou mais partido dos sistemas de vídeo. O autor usou bandas reflectoras, em vez de lâmpadas, sobre as principais articulações do corpo. Em alguns casos, as bandas foram usadas como fitas em torno dos membros, de forma a permitir percursos curvilíneos ou mudanças de direcção. Posteriormente, ajustou os níveis de contraste da imagem

para valores máximos, permitindo que apenas fossem exibidos na imagem os pontos das articulações contra um fundo negro homogêneo.

Nas suas experiências, Johansson constatou que este estímulo simplificado de movimento humano, aqui denominado de *point-light walker* (PLW), produzia uma impressão vívida e imediata de se tratar de uma pessoa em andamento. Assim, a iluminação das articulações de uma pessoa em andamento é suficiente para proporcionar uma imagem clara de um humano em movimento, embora este percepto dê 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.

Em larga medida, as evoluções observadas desde Johansson nas representações do movimento ficaram a dever-se a desenvolvimentos tecnológicos. Por exemplo, Cutting (1978) desenvolveu um sistema de animação computacional baseado em síntese artificial. Este sistema tinha o benefício de produzir PLW sem necessitar do recurso a filmagens de sujeitos humanos, uma vez que os estímulos eram gerados a partir de um algoritmo. Por outro lado, enquanto simplificava o processo de aquisição, esta técnica colocava novos problemas, como a necessidade de criar um novo algoritmo para cada acção nova e a perda de naturalidade dos estímulos (Dekeyser, Verfaillie, & Vanrie, 2002).

Mais recentemente, surgiu um conjunto variado de técnicas para a aquisição de movimento humano. Algumas técnicas fazem uso de infravermelhos, emitidos ou reflectidos por marcadores colocados nos sujeitos e captados por câmaras, outras recorrem a marcadores magnéticos, e permitem a obtenção de coordenadas 3D do movimento que podem ser posteriormente animadas computacionalmente (Blake & Shiffrar, 2006). De forma a superar constrangimentos técnicos e temporais na aquisição destes dados, vários investigadores criaram as suas próprias bases de dados com filmes ou coordenadas 3D de várias acções humanas, encontrando-se algumas disponíveis à comunidade científica (ver, por exemplo, Vanrie & Verfaillie, 2004).

A acção humana não é processada da mesma forma que o movimento rígido. A sensibilidade ao movimento biológico aumenta rapidamente com o número de pontos iluminados, 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 simples. A 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). Estas propriedades não são, contudo, ilimitadas. Shiffrar e Pinto (2002) enunciaram 3 limites à robustez da percepção do movimento humano:

- 1) A orientação é crítica. Se a acção é apresentada invertida de cima para baixo, os sinais de movimento parecem ser interpretados pelo sistema de percepção de movimento rígido do observador;
- 2) Se as características temporais da acção humana são incompatíveis com a dinâmica do movimento humano, a acção é interpretada como um objecto.
- 3) Dados de imagiologia sugerem que o sistema motor pode estar envolvido na análise de movimentos possíveis, mas não impossíveis. Isto indica que nem todas as acções são tratadas de igual forma pelo sistema visual.

O estudo da percepção do movimento biológico a partir do uso das metodologias descritas tornou, por sua vez, 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 (Troje, 2002; Pollick et al., 2005), a sua identidade (Troje, Westhoff, & Lavrov, 2005; Richardson & Johnston, 2005; Loula et al., 2005; Jacobs & Shiffrar, 2005) e até estados emocionais (Dittrich et al., 1996).

A robustez deste tipo de estímulo e as suas propriedades são melhor compreendidas à luz dos dados neurofisiológicos entretanto acumulados. 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 et al., 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 et al., 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 et al, 2002; Peuskens et al, 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.

À semelhança da área STS, descobriu-se recentemente uma área cerebral que responde selectivamente à observação de acções humanas, a área F5 do córtex pré-motor. Os neurónios desta área cerebral são activados tanto quando o primata observa, como quando executa determinadas acções (Murata et al., 1997), pelo que têm sido apelidados de “mirror neurons” e supõe-se que terão como função a preparação para a acção, ou a aprendizagem da acção pela observação. Surpreendentemente, observou-se que esta área também é activada quando se ouvem sons relacionados com a actividade. (Kohler et al, 2002).

Mas também 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 et al., 2005). Esta descoberta levou à formulação da hipótese de que a área STSp poderá fazer parte de uma rede que integra tudo quanto está relacionado com o movimento biológico, que pode ter 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, será de esperar que a percepção audiovisual do movimento biológico tenha propriedades distintas daquelas observadas em processos perceptivos semelhantes, com movimento rígido. Contudo, ainda não se encontram dados científicos suficientes para confirmar esta hipótese. Torna-se, então, relevante estudar a dinâmica dos estímulos visuais e auditivos no movimento biológico, de modo a comparar com os dados já conhecidos do movimento rígido.

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 et al., 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 o 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 & Pennel, 2002; Repp, 2003). Recentemente, Watkins e colaboradores (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 científicos 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 et al., 2004). Repp e Pennel (2002) integraram 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 anteriores com estímulos biológicos indicaram que as pistas podem, quando incompatíveis, ser integradas, formando um novo percepto. No efeito McGurk (McGurk & MacDonald, 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, A., 2005). Também, um trabalho recente investigou os correlatos comportamentais da integração audiovisual no processamento de pistas de movimento biológico (Brooks et al., 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) tinham desenvolvido previamente um estudo metodologicamente semelhante ao de Brooks e colaboradores (2007), com movimento rígido. Neste estudo, os efeitos de facilitação não se tinham verificado, não se tendo encontrado melhorias para além do esperado pela mera

combinação estatística dos sinais perante o movimento na mesma direcção. Ao invés, tinha-se observado que 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.

1.2. Objectivos do Estudo

No seguimento do enquadramento teórico acima apresentado, definiu-se como objectivo para o presente estudo observar os processos envolvidos na percepção audiovisual do movimento biológico. Assumiram-se, como hipóteses:

- 1) Deve observar-se menor predominância sensorial e maior tendência para a integração dos estímulos visual e auditivo. Ou seja, espera-se que os sujeitos não se centrem apenas numa pista, em detrimento da outra.
- 2) Deve observar-se um efeito de facilitação bimodal, i.e., a condição audiovisual deve apresentar melhores resultados que as condições visual e auditiva.

De forma a testar estas hipóteses, optou-se por fazer uso de estímulos visuais ambivalentes 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 recentemente 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° (ver figura 1).

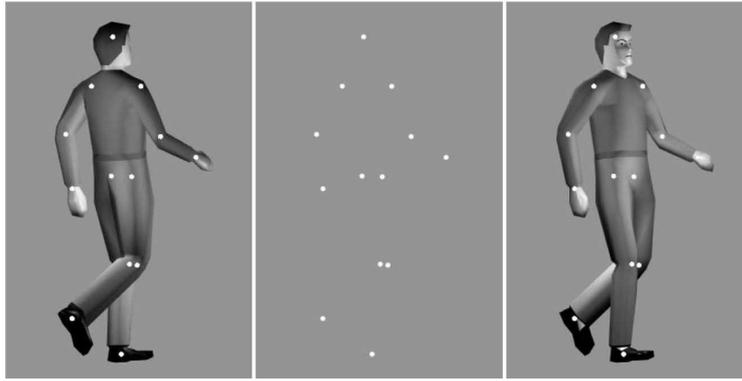


Figura 1: Ilustração de Vanrie, Dekeyser, & Verfaillie (2004):

Os estímulos em afastamento para a direita (135°D – ver representação da esquerda), são semelhantes aos estímulos em aproximação para a direita (45°D – ver representação da direita), quanto à configuração e ao padrão de movimento dos pontos luminosos. Qualquer uma destas interpretações do PLW é possível e estável.

No estudo acima referido, os autores verificaram 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 e na condição de controlo (PLW invertidos) os valores foram apenas de 50 por cento para orientações frontais (resposta ao acaso). 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 0.85 na primeira

condição e de 0.97 na segunda. Conclui-se, assim, que apesar de robusto, o viés visual é sensível a pistas espaciais.

Fazendo uso deste artefacto perceptivo, procurou-se, no presente estudo, observar o impacto de estímulos auditivos congruentes sobre estas representações visuais fortemente enviesadas. Deste modo, num primeiro momento, este trabalho pretendeu obter uma replicação dos dados de Vanrie, Dekeyser e Verfaillie (2004), para num segundo momento analisar o efeito das pistas auditivas sobre os estímulos visuais de movimento biológico.

De acordo com os objectivos e metodologias definidas, este projecto apresentou então quatro fases distintas:

- *1ª Fase - Definição dos estímulos visuais:*

Durante esta fase, foram gerados e testados vários estímulos visuais. Ao longo de vários pré-testes, ajustou-se a dimensão do avatar, a sua distância ao observador, as luminâncias e contrastes. O objectivo foi conseguir estímulos com movimento em de frente e de costas para o sujeito (0° e 180°) e em aproximação e afastamento para a esquerda e para a direita (45°E, 45°D, 135°E, 135°D), que os sujeitos conseguissem identificar correctamente entre 60 a 80 por cento das vezes. Pretendeu-se assim criar um conjunto de situações suficientemente ambíguas para permitirem a formação de perceptos alternativos, mantendo o elevado viés frontal. Decorreu desta fase a necessidade de criar duas experiências, uma com e outra sem translação dos PLW.

2ª Fase – Definição dos Estímulos Auditivos:

Durante esta fase foram gerados e testados vários tipos de ficheiros de som. Procurou-se assegurar a boa espacialização destes estímulos e definiram-se distâncias interaurais, a distância à fonte sonora e a intensidade do som. Escolheram-se os ficheiros com as melhores propriedades perceptivas, que permitiam taxas de acerto em pré-teste entre os 70 e os 80 por cento.

3ª Fase – Experiência 1, Sem Translação:

Nesta fase experimental, procurou-se criar um plano e contexto experimental semelhante ao de Vanrie, Dekeyser e Verfaillie (2004). Pretendeu-se replicar os valores de viés daquele estudo para a condição visual e comparar estes valores com os da condição auditiva e audiovisual. Uma vez que na

primeira fase se tinham encontrado valores de enviesamento superiores aos esperados, nomeadamente para as orientações 0° e 180° dos PLW, optou-se por fazer uma segunda experiência, menos ambígua do ponto de vista visual.

4ª Fase – Experiência 2, Com Translação:

Esta fase experimental foi em tudo semelhante à anterior, tendo-se apenas introduzido o movimento real dos avatares no espaço.

São abaixo apresentadas as descrições detalhadas da 3ª Fase e da 4ª Fase de estudo. Todos os sujeitos participaram em ambas as experiências, e foram contrabalanceados entre elas, de forma a evitar efeitos de ordem, de aprendizagem e de fadiga. As condições visual, auditiva e audiovisual foram testadas num plano intrasujeito, pelo que as taxas de viés e de acerto puderam ser analisadas sujeito a sujeito.

2. EXPERIÊNCIA 1 – SEM TRANSLAÇÃO

2.1. Método

2.1.1. Sujeitos

Todos os participantes foram seleccionados de acordo com uma amostragem não aleatória, dado que não se esperavam efeitos associados às variáveis sexo/educação/estatuto socioeconómico. Uma vez que se conhecem efeitos de idade sobre as competências de acuidade perceptiva, esta variável foi controlada mantendo-se, pelo que todos os sujeitos tinham idades compreendidas entre os 23 e os 28 anos de idade. Ao todo, usaram-se 7 sujeitos (colegas e estudantes), todos dextros, 3 homens e 4 mulheres.

Num momento prévio às experiências, todos os sujeitos 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.

2.1.2. Estímulos e Material

Os estímulos visuais desta condição experimental consistiram em PLW de 13 pontos luminosos animados com coordenadas 3D, construídos a partir da base de dados de Vanrie & Verfaillie (2004) (disponível em www.psychonomic.org/archive/). Os pontos luminosos eram esferas brancas de 1 cm de diâmetro (em média) em movimento, com uma luminância de cerca de 68 cd/m² sobre um fundo cinza escuro de aproximadamente 1.53 cd/m². Os estímulos eram projectados sobre uma tela a 60 Hz, ocupando uma área de projecção de 4.94 m x 2.10 m. Os avatares mediam, dos pés à cabeça, 30 cm na tela, o que corresponde a um ângulo visual de 6.36° para o observador, que estava sentado a 3 m da tela.

A apresentação e manipulação informática de todos os estímulos foi feita a partir do Biomose, uma aplicação informática baseada em Open GL. Trabalhou-se a com um cluster de três computadores a partir do sistema VR/Net Juggler. 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 e usou-se uma placa gráfica Nvidia Quadro FX 4500. 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 passadeira rolante (sem translação), e foram colocados no mundo virtual a uma distância de 20 m do observador.

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 lateral (45°E e 45°D) e em afastamento lateral (135°E e 135°D). Não se apresentaram orientações puramente laterais (90°), uma vez que não acrescentavam resultados relevantes para a experiência e a tornavam bastante mais morosa e cansativa para os sujeitos.

Os estímulos auditivos consistiram em seis tipos de ficheiros de som, cada um com a duração de 4 s, 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 20 Hz e os 10.000 Hz, concentrando-se a maior parte da informação auditiva apresentada em torno dos 1.000 Hz. 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 80 cm do limite visual) e a 50 cm 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).

2.1.3. Design Experimental

Os dezoito tipos de estímulo foram apresentados em cinco ensaios cada um, em ordem pseudo aleatória entre eles. Cada apresentação teve a duração total de 4 s, a mesma usada no estudo de Vanrie, Dekeyeser e Verfaillie (2004), o correspondente a quatro ciclos de passada por ensaio. O intervalo entre ensaios teve a duração de 3 s, consistindo a duração total da experiência em cerca de 11 minutos.

Durante a experiência, os sujeitos deviam responder manualmente, apontando numa tabela octogonal (ver fig. 2) qual a direcção em que o estímulo se movia.

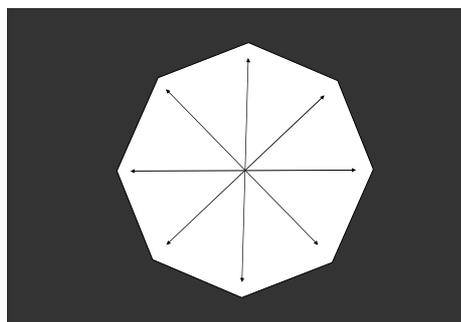


Figura 2 – Tabela de Resposta

Quando o sujeito percebia o estímulo num movimento de aproximação, deveria apontar na seta descendente correspondente, discriminando a lateralidade. As respostas em afastamento deveriam ser assinaladas nas setas

ascendentes. As setas laterais, apesar de representadas, não deveriam ser usadas, uma vez que nenhum ensaio continha estímulos em movimento naquela direcção.

Todas as respostas dos sujeitos foram anotadas manualmente pelo experimentador.

2.1.4. Procedimentos

Todos os sujeitos foram testados individualmente, numa grande 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 sujeitos tinham uma mesa de superfície negra, onde constava a tabela de respostas. 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 de respostas, com indicação da orientação a que correspondia cada seta e explicação da inexistência de ensaios laterais. Todos os sujeitos foram instruídos para se centrarem na cruz branca apresentada durante o período entre ensaios e responderem atempadamente, e foi-lhes dito que era permitido alterar o sentido das respostas durante os ensaios, se se apercebessem de que tinham cometido um erro. Pediu-se que apresentassem sempre resposta, mesmo em caso de dúvida. Seguidamente, iniciava-se a sequência experimental.

2.2. Resultados

Ao longo da experiência, registaram-se respostas para todos os ensaios, não se tendo observado omissões. Três sujeitos já tinham participado na experiência 2 previamente, os restantes quatro eram inexperientes. De acordo com o teste U de Mann-Whitney, a diferença entre ambos os grupos quanto às taxas de acerto é significativa ($U= 0.00$, $n_1=4$, $n_2=3$, $p\leq 0.05$). Como se poderá observar na Tabela 1, contudo, esta diferença reflecte sobretudo proporções de acerto globalmente superiores, sendo que os resultados foram similares entre

ambos os grupos quanto à evolução ao longo das condições visual, auditiva e audiovisual.

Tabela 1: Média das proporções de acerto, dos sujeitos sem experiência e com experiência, nas condições visual, auditiva e audiovisual.

	Sem Experiência	Com Experiência
Visual	0,6	0,73
Auditiva	0,71	0,87
Audiovisual	0,84	0,91

Assim, optou-se por analisar os dados gerais de ambos os grupos em conjunto, ainda que mais abaixo se observem os resultados sujeito a sujeito.

Quanto à distribuição das respostas por orientação, observou-se uma boa homogeneidade na condição auditiva. Na condição visual, houve um forte viés, sendo que em cada cinco ensaios para a frente e cinco ensaios para trás, os sujeitos responderam em média cerca de oito vezes para a frente e duas vezes para trás (Figura 3).

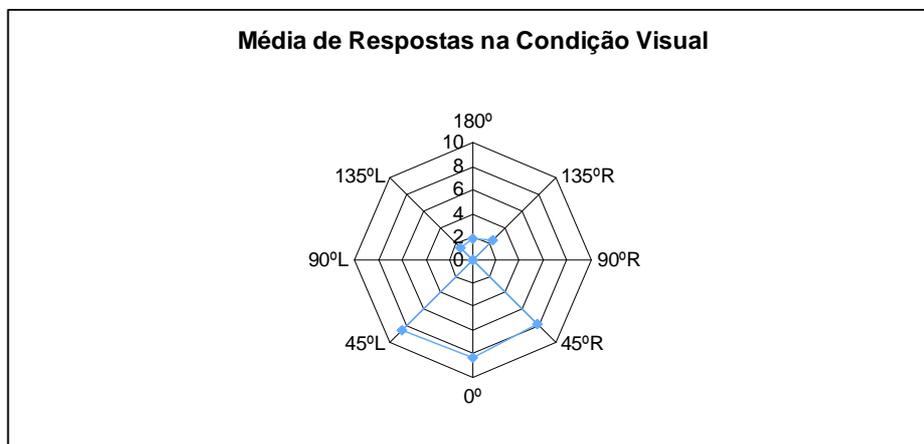


Figura 3: Distribuição das respostas dadas em média pelos sujeitos nos cinco ensaios de cada orientação, na condição visual.

Na condição audiovisual, este viés foi largamente atenuado, mas não desapareceu. Nesta condição, os sujeitos responderam em média cerca de quatro vezes para trás e seis vezes para a frente (Figura 4).



Figura 4: Distribuição das respostas dadas em média pelos sujeitos nos cinco ensaios de cada orientação, na condição audiovisual.

A redução de viés que se observou na condição audiovisual traduziu-se também num aumento claro do número de acertos. Uma vez que não se observaram diferenças relevantes entre as taxas de acerto nas orientações 45°E e 45°D, assim como nas orientações 135°E e 135°D, estas foram agrupadas para efeitos de análise de resultados. Deste modo, passamos a observar os dados condição a condição.

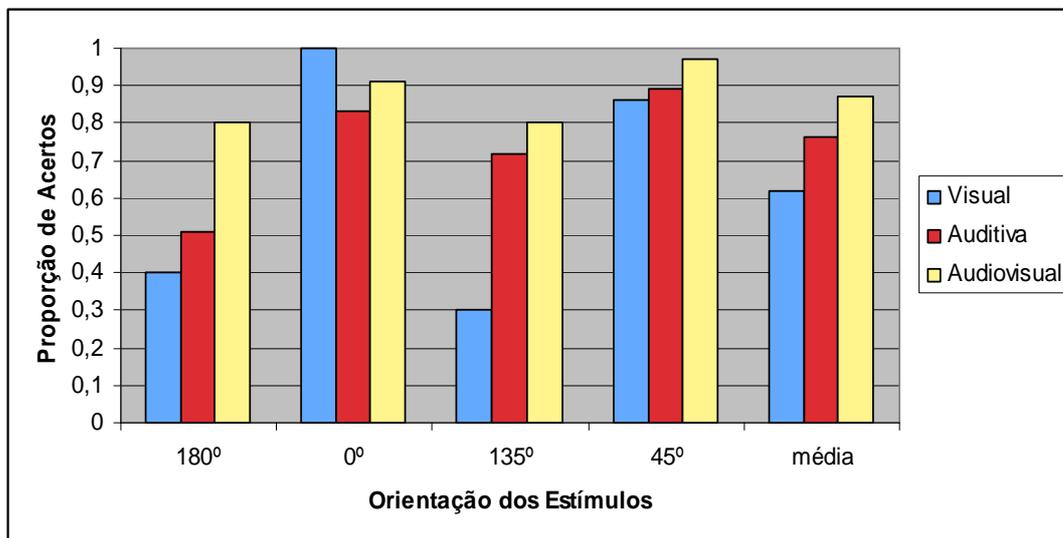


Figura 5: Proporção de acertos nas condições visual, auditiva e audiovisual, nos ensaios fronto-paralelos e oblíquos. Média das proporções de acerto das três condições.

A condição com maior discrepância quanto aos acertos nas diferentes orientações é a visual. Nesta condição, observam-se probabilidades de acerto muito reduzidas nas orientações 135° e 180°, que contrastam com a facilidade

dos ensaios de 0°, nos quais não se observaram erros. Naturalmente, estas diferenças notórias estão relacionadas com os dados acima apresentados acerca do forte viés visual, sendo natural 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 proporções de acerto nos ensaios orientados para a frente e nos ensaios orientados para trás, 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 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 que os ensaios fronto-paralelos, tendo os melhores resultados surgido nos ensaios em aproximação lateral (45°) e os piores resultados surgido nos ensaios em afastamento (180°). 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, por comparação com as condições unimodais. Adicionalmente, foi esta a condição em que se observaram menores discrepâncias entre orientações e menores diferenças entre sujeitos. 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 proporções 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 interacção entre as pistas visuais e auditivas mais forte, com um aumento muito

expressivo nas taxas de acerto: de uma proporção de acerto de 0.4 na condição visual e de 0.51 na condição auditiva, obteve-se uma proporção de acerto de 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 0º surgiu um efeito de interacção audiovisual notório. Nestes ensaios, a condição audiovisual não obteve melhores resultados que a condição unimodal visual, que tinha atingido um valor máximo. Este dado interessante revela efeitos de interacção bimodal que levam a excluir a hipótese de os sujeitos tenderem a centrar-se na pista mais clara. Com efeito, mesmo quando a pista visual não apresentava qualquer ambiguidade, quando considerada em conjunto com a pista auditiva, menos fácil, não manteve as suas propriedades perceptivas. Os efeitos de interacção bimodal parecem, assim, acontecer em todas as orientações, mas não obedecendo a mecanismos simples de combinação de probabilidades.

De acordo com o teste de Friedman, para três grupos sem distribuição normal num plano intrasujeito, a diferença de resultados entre as três condições é estatisticamente significativa ($\chi^2=86.00$, $gl=1$, $p\leq 0.01$), pelo que as suas distribuições não são estatisticamente semelhantes e esta dissemelhança não se deve ao acaso.

De forma a melhor analisar estes efeitos, convém observar os resultados em maior detalhe, sujeito a sujeito. Na Tabela 2, são apresentadas as proporções de acerto de cada sujeito em cada condição, assim como a proporção média de acertos totais e a média apenas dos acertos nas condições unimodais, para efeito de comparação com a condição bimodal.

Tabela 2: Proporções de acerto dos sete sujeitos ao longo das três condições experimentais.

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Sujeito 6	Sujeito 7
Auditiva	0,83	0,87	0,73	0,77	0,60	0,90	0,67
Visual	0,87	0,60	0,67	0,47	0,60	0,63	0,57
Média (a,v)	0,85	0,74	0,70	0,62	0,60	0,77	0,62
Audiovisual	0,93	0,83	0,87	0,93	0,70	0,97	0,87
Média (a,v,av)	0,88	0,77	0,76	0,72	0,63	0,83	0,70

Os sujeitos 1, 2 e 6 já tinham realizado a experiência 2 previamente. Como se pode observar, estes sujeitos obtiveram, de facto, melhores resultados globais que os restantes. Contudo, todos os sujeitos apresentaram, na condição audiovisual, resultados superiores aos observados na condição visual. Assim, pode dizer-se que parece ter havido um efeito de bimodalidade sobre os PLW enviesados.

Também, os resultados da condição audiovisual não traduziram uma simples média das proporções de acerto das condições unimodais, mas tiveram sempre algum ganho quanto a esta. Os sujeitos 1,3,5 e 7 obtiveram proporções de acerto nas condições visuais que foram semelhantes às das auditivas. Todos estes sujeitos apresentaram, na condição audiovisual, um resultado um pouco mais elevado que os resultados unimodais.

O sujeito 2 obteve resultados surpreendentes, com acertos superiores na condição auditiva, por comparação com a condição audiovisual. Este sujeito leva-nos a crer que será incorrecto assumir que há sempre resultados superiores em ensaios bimodais, sendo mais correcto afirmar que os resultados da condição audiovisual são superiores à simples média de probabilidades das condições unimodais.

O sujeito 4 apresentou um perfil de forte viés na condição visual. Apesar de os resultados deste sujeito não terem sido particularmente elevados na condição auditiva, observou-se um forte efeito de interacção bimodal, com um ganho elevado na condição audiovisual. Também o sujeito 6 obteve resultados claramente melhores na condição auditiva que na condição visual, tendo na condição audiovisual atingido valores quase máximos de acerto.

2.3. Discussão

Quanto ao objectivo primário, de replicação dos dados de viés visual obtidos por Vanrie, Dekeyser e Verfaillie (2004), observou-se, com efeito, um valor próximo do encontrado por aqueles autores na condição visual. Este viés foi largamente reduzido, mas não eliminado, na condição audiovisual.

De acordo com as hipóteses enunciadas, era esperado encontrar resultados indicadores de uma tendência para integrar os estímulos visuais e auditivos,

mesmo quando estes pudessem assumir características incompatíveis, ao contrário da centração no estímulo mais claro observada nos estudos de movimento rígido. A análise do efeito dos estímulos auditivos sobre o viés visual permite confirmar estas expectativas. Com efeito, o forte viés foi largamente reduzido na condição bimodal, perante pistas auditivas de afastamento (incompatíveis com o percepto visual). Ainda assim, não se observou uma redução total do viés, que seria indicadora da possível centração no estímulo auditivo, por ser o mais claro.

Era igualmente esperado encontrar, na condição audiovisual, resultados melhores que nas condições unimodais, compatíveis com um efeito de facilitação bimodal da percepção. Quando comparadas as três condições experimentais, observou-se efectivamente uma diferença significativa entre os grupos. A condição audiovisual obteve menores variações de resultados nos diferentes ensaios, assim como valores menos discrepantes entre os vários sujeitos que as restantes condições. Isto poderá levar a concluir que estes ensaios eram mais estáveis e claros, do ponto de vista perceptivo. Também se observou, em média, um aumento claro das taxas de acerto na condição bimodal e todos os sujeitos obtiveram melhores resultados nesta condição que na condição visual.

Mas parece incorrecto assumir que a condição audiovisual obtém sempre melhores resultados. O efeito da interacção bimodal assumiu outros contornos nos ensaios cujos estímulos se apresentavam a 0°: os ensaios visuais obtiveram dados indicadores de um estímulo claro e não ambíguo (não houve erros) e os ensaios audiovisuais não replicaram este valor, talvez por acção dos estímulos sonoros, menos fáceis. Este efeito surgiu também na observação dos dados individuais, em que o sujeito 2 obteve bons resultados na condição auditiva e menos acertos na condição visual, não tendo em ensaios bimodais obtido melhores resultados que nos ensaios de som. O resultado deste e dos restantes sujeitos não traduziu, contudo, uma simples média das probabilidades de acerto das duas condições unimodais, mas um valor claramente superior a esta, o que indica um processo de interacção intermodal, compatível com um efeito de facilitação perceptiva.

3. EXPERIÊNCIA 2 – COM TRANSLAÇÃO

3.1. Métodos

3.1.1. Sujeitos

Os sujeitos desta experiência foram os mesmos da experiência anterior. Três sujeitos efectuaram esta experiência em primeiro lugar, os restantes quatro tinham participado na experiência 1 previamente. O intervalo entre as experiências foi, para todos os sujeitos, de cerca de um minuto, não se tendo permitido que abandonassem o seu lugar e evitou-se feedback ou diálogo durante esse período.

3.1.2. Estímulos

Todos os estímulos para esta experiência foram similares aos da experiência 1. Os PLW foram apresentados nas seis orientações acima descritas, mas com translação efectiva, a uma velocidade de 1 m/s. Assim, ao invés de se movimentarem sem de facto abandonarem o ponto inicial de projecção, nesta experiência os avatares iniciavam os ensaios num ponto visual e auditivo central com relação à cabeça do sujeito e posteriormente deslocavam-se. Nos ensaios com movimento lateral, os PLW deslocavam-se em direcção às extremidades da tela. Nos ensaios em aproximação, os avatares expandiam no tamanho dos pontos, assim como na distância entre eles, tornando-se gradualmente maiores. Nos ensaios em afastamento, acontecia o efeito inverso.

Para diminuir a facilidade destes ensaios (em experiências piloto tinham-se obtido cem por cento de acertos na condição visual), os objectos visuais foram colocados a uma distância de 30 m do observador no ponto de partida. Assim, os avatares mediam, dos pés à cabeça, no momento inicial de cada ensaio, 21 cm na tela, ocupando um ângulo visual de 4.45°, que posteriormente aumentava ou diminuía, consoante o tipo de ensaio. Os pontos brancos tinham um diâmetro médio inicial de 8 mm na tela. As cores e luminâncias não foram alteradas nesta experiência.

Os estímulos auditivos desta experiência foram os mesmos da experiência anterior e conservaram os mesmos valores de intensidade e frequência.

3.1.3. Design Experimental

O design desta experiência foi em tudo semelhante ao da experiência anterior. Manteve-se o mesmo número de estímulos, com cinco repetições cada, e com as mesmas durações, pelo que esta experiência manteve a mesma duração total da anterior. Toda a sequência de ensaios foi pseudo aleatorizada computacionalmente.

3.1.4. Procedimentos

Mantiveram-se, nesta experiência, os procedimentos da experiência 1.

Aos sujeitos que já tinham realizado previamente a experiência 1, foi apenas explicado que os estímulos agora se moviam efectivamente no espaço e que deveriam manter o mesmo protocolo de resposta.

3.2. Resultados

À semelhança do que aconteceu na experiência 1, registaram-se respostas para todos os sujeitos e não se observaram omissões. Quatro dos sujeitos eram já experientes, os restantes três efectuaram esta experiência em primeiro lugar. Contudo, ao contrário do que sucedeu na experiência 1, não se observaram diferenças significativas entre ambos os grupos quanto às taxas de acerto ($U=5,50$, $n_1=3$, $n_2=4$, $p>0,05$), ou seja, não se observaram diferenças relevantes quanto ao posicionamento dos valores médios obtidos por cada grupo.

Como se poderá constatar, pela análise da Figura 6, os dois grupos foram, contudo, diferentes quanto à distribuição dos seus resultados. Os sujeitos que já tinham realizado a experiência 1 anteriormente apresentaram resultados mais heterogéneos e uma maior diferença entre os valores máximo e mínimo, enquanto que os sujeitos não experientes apresentaram valores mais concentrados em torno do valor médio e com uma dispersão de sentido negativo.

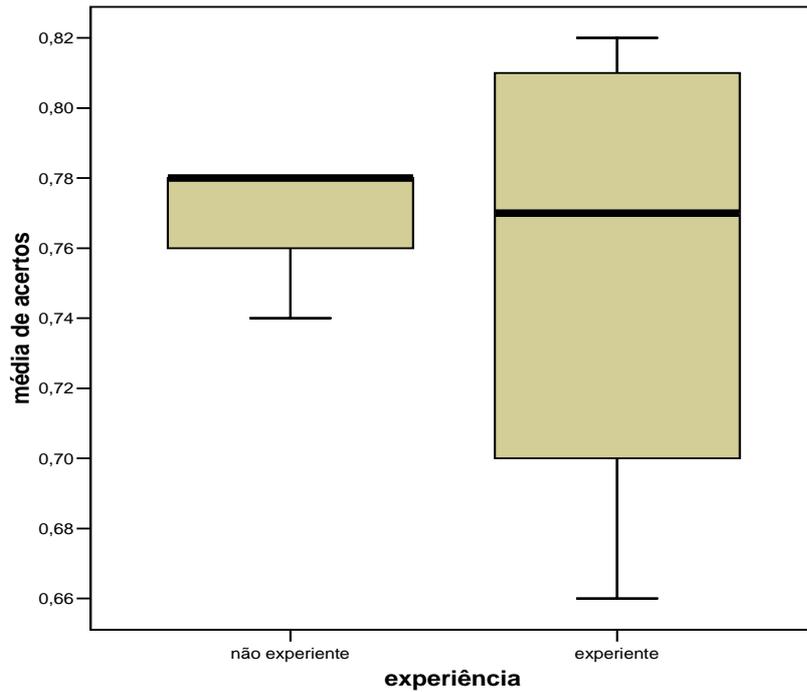


Figura 6: Diagrama de extremos e quartis acerca da distribuição de resultados dos sujeitos experientes e não experientes.

Será igualmente interessante constatar que, ao contrário do que se tinha observado na experiência anterior, aqui os sujeitos experientes não obtiveram melhores resultados, tendo até piorado na condição visual, ainda que pareça ter havido alguma aprendizagem auditiva.

Tabela 3: Média das probabilidades de acerto, dos sujeitos sem experiência e com experiência, nas condições visual, auditiva e audiovisual.

	Sem Experiência	Com Experiência
Visual	0,69	0,53
Auditiva	0,69	0,80
Audiovisual	0,90	0,84

Quanto aos valores de viés e à distribuição de respostas pelas várias orientações, valores muito semelhantes aos observados anteriormente, com um forte viés visual que não foi atenuado pelas pistas de expansão introduzidas (Figura 7), o que de certa forma veio contrariar os dados obtidos por Vanrie,

Dekeyser e Verfaillie (2004). Nos ensaios do lado esquerdo, o viés médio atingiu valores próximos de nove respostas para a frente e uma para trás em cada cinco ensaios para a frente e cinco para trás. Os ensaios fronto-paralelos e do lado direito mantiveram o rácio de oito respostas para a frente por duas para trás.

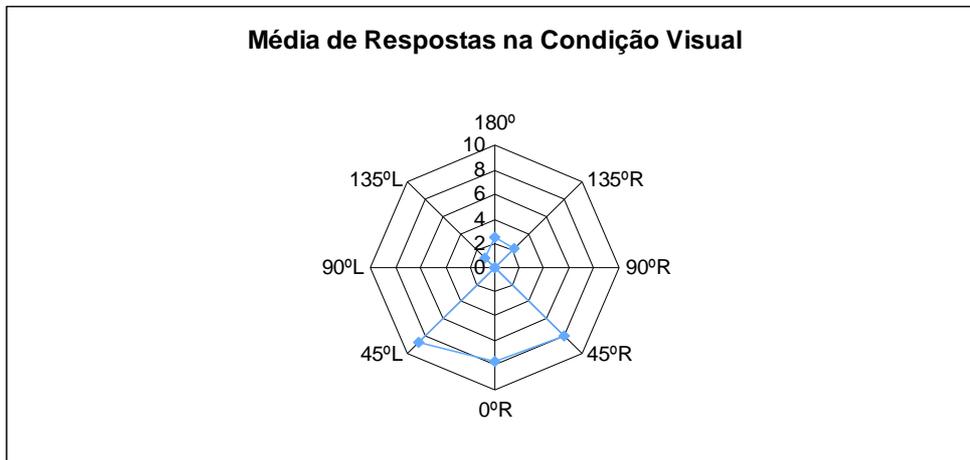


Figura 7: Distribuição das respostas dadas em média pelos sujeitos nos cinco ensaios de cada orientação, na condição visual.

A condição audiovisual obteve dados muito semelhantes aos anteriormente observados, que representaram uma diminuição clara, mas não total, do enviesamento visual (Figura 8).

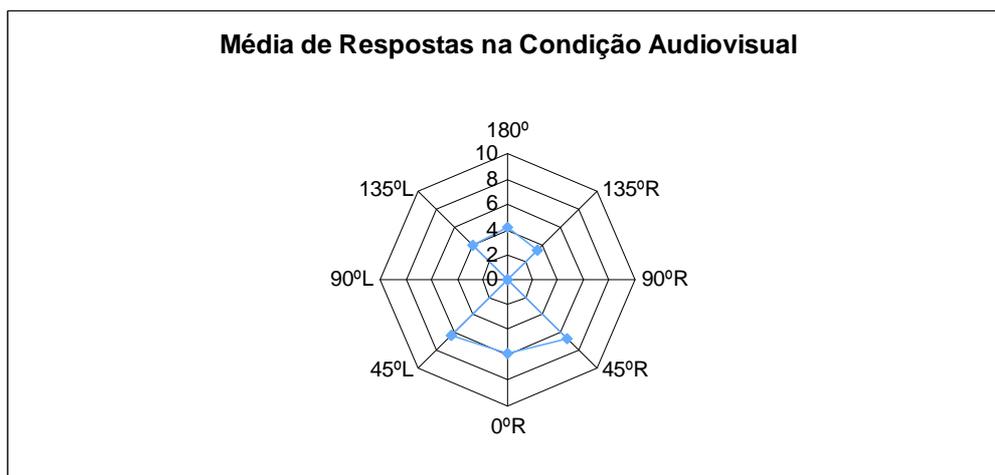


Figura 8: Distribuição das respostas dadas em média pelos sujeitos nos cinco ensaios de cada orientação, na condição audiovisual.

No que respeita às taxas de acerto, os valores médios observados indicam uma melhoria global dos resultados na condição audiovisual (Figura 6) e são muito semelhantes aos observados na experiência 1.

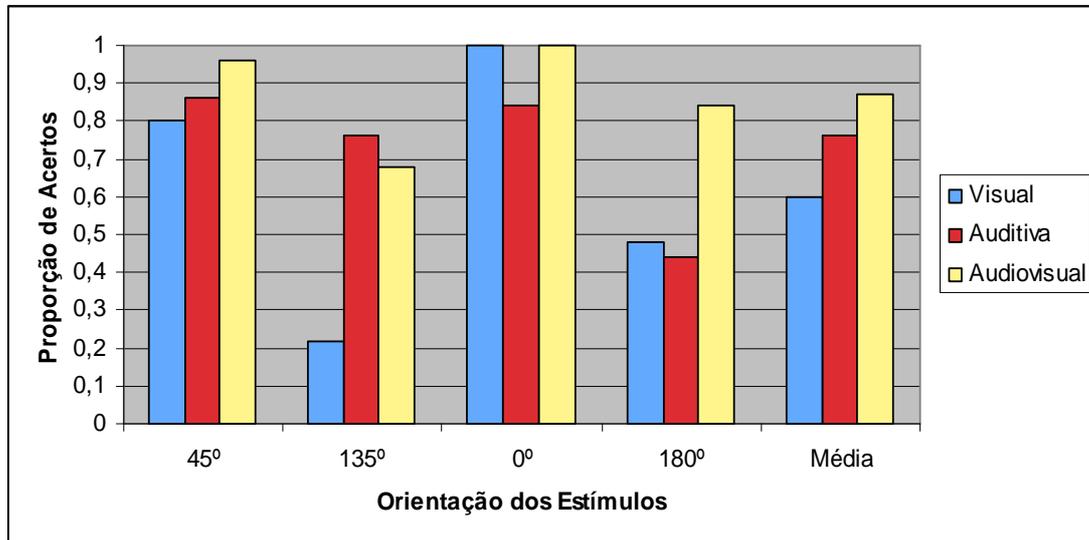


Figura 9: Proporção de acertos nas condições visual, auditiva e audiovisual nos ensaios fronto-paralelos e oblíquos. Média das proporções de acerto das três condições.

Uma vez mais, nesta experiência, observaram-se os valores mais elevados de acerto nos ensaios orientados para a frente. A condição visual foi aquela que obteve valores mais discrepantes entre os eixos frente/trás, tal como tinha sido observado na experiência 1, o que uma vez mais traduz o forte viés destes estímulos, em nada diminuído pela presença de pistas de translação. Uma vez mais, todos os erros observados nesta condição deveram-se a confusões frente/trás (devidas ao carácter bi-estável dos estímulos) e nunca a confusões laterais.

Na condição auditiva, houve uma notória semelhança entre os valores desta experiência e os observados na experiência 1, dado desejável, uma vez que se mantiveram todas as características destes ensaios entre ambas. É de referir que se observou um valor claramente mais baixo que os restantes, o dos ensaios com orientação 180°, dado que pode ser justificado, como acima assinalado, pela forte diminuição da intensidade do estímulo ao longo dos ensaios, a par da ausência de pistas binaurais de discriminação de lateralidade. A confusão lateral constituiu, uma vez mais, a maior fonte de erros nesta condição.

A condição audiovisual obteve, em média, resultados melhores que as condições unimodais. Nos ensaios orientados a 45°, obtiveram-se acertos superiores à condição visual e auditiva, atingindo-se quase a totalidade sucesso. Nos ensaios a 135°, que apresentavam uma forte discrepância entre os valores de acerto visual e auditivo, os resultados da condição audiovisual não foram superiores aos da condição auditiva, mas foram claramente superiores à média dos dois valores de acerto unimodais.

Os ensaios de orientação a 0° obtiveram, nesta experiência, o máximo de sucesso, à semelhança do que se tinha observado nos ensaios visuais. Assim, à semelhança do que aconteceu na experiência 1, observou-se um ganho claro face à média de proporções de acerto nas condições unimodais, tendo esse ganho sido suficiente para atingir o valor máximo de acertos, ao contrário do que se tinha observado anteriormente. Nos ensaios de orientação a 180°, replicaram-se os dados da experiência 1. Apesar de o estímulo visual se tornar gradualmente mais pequeno ao longo do ensaio, as taxas de acerto na condição visual mantiveram-se muito baixas, e de valor próximo ao das taxas de acerto na condição auditiva. Os ensaios audiovisuais revelaram, nesta orientação, uma evolução expressiva face aos valores médios unimodais: obteve-se uma taxa de acerto de 0.48 na condição visual, 0.44 na condição auditiva e 0.84 na condição audiovisual.

A diferença de resultados entre as três condições foi estatisticamente significativa pelo teste de Friedman ($\chi^2=78.00$, $gl=1$, $p\leq 0.01$).

Nos valores de acerto individuais de cada sujeito condição a condição, observaram-se dados semelhantes. De modo a melhor comparar o comportamento de cada sujeito entre as experiências, volta-se a apresentar, abaixo, a Tabela 2, com os dados relativos à experiência 1.

Uma vez mais, observou-se que todos os sujeitos obtiveram, na condição audiovisual, resultados superiores aos obtidos na condição visual, resultados estes que traduziram a acção dos estímulos auditivos sobre os estímulos visuais. Os sujeitos 1, 2 e 6 eram inexperientes, ao passo que os restantes já tinham participado anteriormente na experiência 1.

Tabela 2 (replicada): Proporções de acerto dos sete sujeitos na experiência 1.

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Sujeito 6	Sujeito 7
Auditiva	0,83	0,87	0,73	0,77	0,60	0,90	0,67
Visual	0,87	0,60	0,67	0,47	0,60	0,63	0,57
Média (a,v)	0,85	0,74	0,70	0,62	0,60	0,77	0,62
Audiovisual	0,93	0,83	0,87	0,93	0,70	0,97	0,87
Média (a,v,av)	0,88	0,77	0,76	0,72	0,63	0,83	0,70

Tabela 4: Proporções de acerto dos sete sujeitos na experiência 2.

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Sujeito 6	Sujeito 7
Auditiva	0,57	0,80	0,80	0,97	0,77	0,73	0,77
Visual	0,83	0,60	0,57	0,50	0,50	0,70	0,57
Média (a,v)	0,70	0,70	0,69	0,74	0,64	0,72	0,67
Audiovisual	0,90	0,90	0,83	0,97	0,67	0,90	0,90
Média (a,v,av)	0,77	0,77	0,73	0,81	0,65	0,78	0,75

Da experiência 2 para a 1, o sujeito 1 e o sujeito 6 fizeram alguma aprendizagem auditiva, mas esta não se reflectiu nos resultados da condição audiovisual, que se mantiveram igualmente elevados. O sujeito 2 manteve o mesmo perfil de acertos. Quanto aos sujeitos experientes, tanto o sujeito 4, como o 5 e o 7 fizeram aprendizagens auditivas, mas mantiveram as taxas de acerto audiovisual semelhantes. O sujeito 3 manteve o seu perfil.

Todos os sujeitos mantiveram os resultados esperados, de um ganho claro de sucesso na condição audiovisual face aos valores médios das condições unimodais. O sujeito 4 voltou a distinguir-se, pelo forte efeito da pista auditiva sobre o elevado viés visual, tendo atingido valores quase máximos de acerto tanto na condição visual, quanto na audiovisual. O sujeito 5 obteve um resultado melhor na condição auditiva que na condição visual. Este efeito já se tinha verificado na experiência 1, com o sujeito 2, e traduz o efeito interactivo entre as pistas.

3.3. Discussão

Na experiência 2 observou-se um valor de viés visual bastante elevado. Este dado tornou-se surpreendente por, ao longo dos ensaios, se terem introduzido pistas visíveis de aumento/diminuição no tamanho dos pontos e no espaço entre eles, o que no estudo de Vanrie, Dekeyser e Verfaillie (2004) tinha alterado os valores de viés. De um modo geral, este dado vem apenas reforçar a robustez do fenómeno perceptivo da bi-estabilidade enviesada, e tornar mais relevante o papel desempenhado pelo estímulo auditivo na sua desambiguação.

Entre ambas as experiências não se observaram efeitos contraditórios ou diferentes. Uma vez mais, confirmou-se a hipótese de um efeito de interacção entre os estímulos visuais e auditivos, com uma melhoria significativa dos resultados na condição audiovisual. Esta melhoria não se traduz na escolha de uma modalidade sensorial mais clara que se torna pista de referência, o que se pode deduzir pela redução, mas não eliminação, dos valores de viés, ou pelos casos em que a melhoria é relativa à média, mas não absoluta, face aos valores unimodais.

Por outro lado, os efeitos de interacção entre as pistas visuais e auditivas foram bi-direccionais, ou seja, tanto se observaram melhorias quando eram as pistas auditivas as mais claras, como quando eram as pistas visuais que obtinham melhor acerto. Assim, parece que tanto a presença da pista auditiva melhora as propriedades perceptivas do estímulo visual, como a pista visual pode melhorar as propriedades do estímulo sonoro. Exemplo disto foi o sujeito 1, que obteve sempre melhores resultados na condição visual que na auditiva e que obteve ganhos de acerto semelhantes aos restantes sujeitos, que obtinham melhores resultados na condição auditiva.

Os valores na condição audiovisual mantêm-se, contudo, imprevisíveis face aos valores das condições unimodais. Não parece haver uma relação clara entre os valores de acerto visuais e auditivos e os valores de ganho na condição audiovisual. Parece, ainda assim, que quanto mais discrepantes são os resultados unimodais, mais imprevisível será esse ganho, podendo nesses casos não se superar o melhor valor unimodal.

Como conclusão segura quanto ao efeito da interação das pistas sensoriais, observa-se que há uma integração dos dados prévia ao momento de decisão, ao invés de uma escolha quanto à pista a assumir como referência. O resultado desta integração caracteriza-se, por sua vez, num aumento face à simples média de proporções unimodais. As implicações deste efeito serão discutidas em maior detalhe mais abaixo.

4. DISCUSSÃO GERAL

Num primeiro momento, o presente trabalho pretendia uma replicação dos dados observados por Vanrie, Dekeyser e Verfaillie (2004) quanto ao forte viés encontrado em estímulos de movimento biológico potencialmente biestáveis. 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, assim como se mantiveram, na experiência 2, com a introdução de movimento de translação. Numa tentativa explicativa, poder-se-á argumentar que estas pistas continuaram atenuadas, nomeadamente pela dimensão diminuída dos estímulos visuais apresentados. Com efeito, os ângulos visuais usados originalmente no estudo de Vanrie, Dekeyser e Verfaillie (2004) foram de 7.87° dos pés à cabeça do avatar, enquanto que os aqui usados foram de 6.36° na experiência 1 e de 4.45° na experiência 2.

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. Então, os sujeitos não se deveriam centrar apenas numa pista, em detrimento da outra, distinguindo assim a percepção bimodal de movimento biológico do movimento rígido. 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). Também na análise de resultados sujeito a sujeito, não se observaram na condição audiovisual as mesmas taxas de acerto que nas condições visual e auditiva. É igualmente relevante observar que estes resultados também não traduzem uma simples flutuação entre pistas assumidas como referência (não se assemelham a uma média de probabilidades), mas ao invés parecem indicar uma integração da decisão prévia ao momento de decisão.

Uma das discussões mais activas entre académicos 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 (Welsh, 1986; Soto-Faraco, Kingstone & Spence, 2004). Dados recentes de fMRI indicam que os processos perceptivos e decisionais na percepção audiovisual de movimento rígido estão intimamente ligados (Soto-Faraco et al., 2004). Os resultados acima descritos quanto aos processos de percepção audiovisual do movimento biológico são indicadores de que, neste tipo específico de estímulos, as pistas serão combinadas previamente, podendo-se estar perante um processo de baixo nível. De acordo com os dados colhidos, parece então que a percepção multimodal do movimento biológico não parece seguir as mesmas regras da percepção multimodal de movimento rígido.

A segunda hipótese enunciada neste estudo indicava que se deveria observar um efeito de facilitação bimodal, ou seja, que a condição audiovisual deveria apresentar 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°. Em ambas as experiências, 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 dificultação destes processos. Os resultados de ambas as experiências 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 bi-estável 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. Esta potenciação de resultados acontece sempre, mas deve ser observada face à combinação de proporções unimodais, e não face a estas proporções isoladamente. Adicionalmente, este efeito de potenciação é bi-direccional, ou seja há melhorias de acerto, quer quando é a pista auditiva, quer quando é a pista visual, que obtém melhores resultados na condição unimodal.

Apesar de todas as análises desenvolvidas, mantém-se contudo imprevisível a forma como os estímulos audiovisuais interagem face aos valores obtidos nas condições unimodais. Parece haver uma relação entre a discrepância de acertos visuais e auditivos e o valor de ganho na condição audiovisual, nomeadamente, parece que quando são mais discrepantes, o resultado se torna mais imprevisível, podendo o ganho não ser suficiente para superar o melhor valor unimodal. 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 se distingue da percepção audiovisual de outro tipo de movimento, uma vez que os processos observados não encontram correlatos com aqueles reportados quanto ao segundo tipo de percepção multimodal.

Todos os dados colhidos são indicadores de mecanismos de interação 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, com 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 interação 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*.
- Cutting, J. (1978). A program to generate synthetic walkers as dynamic point-light displays. *Behavior Research Methods, Instruments, & Computers*, 1, 91-94.
- Dekeyser, M., Verfaillie, K., & Vanrie, J. (2002). Creating stimuli for the study of biological-motion perception. *Behavior Research Methods, Instruments, & Computers*, 34, 375-382.
- 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.
- Kohler, E., Keysers, C., Umiltà, M., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297, 846-848.
- 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.

- Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology*, 78, 2226-2230.
- 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., et al. (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., & Poppel, 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.
- Shiffrar, M., & Pinto, J. (2002) The visual analysis of bodily motion. In W. Prinz & B. Hommel (Eds.), *Common Mechanisms in Perception and action*, Oxford: University Press, 381-399.

- 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.
- 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.
- 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., Dekeyser, M., & Verfaillie, K. (2004). Bistability and biasing effects in the perception of ambiguous point-light walkers. *Perception*, 33, 547-560.
- Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods*, 36, 625-629.
- 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.