

A palavra na leitura de imagens: a cidade que o anjo vê e os címbalos que o ladrão não vê

Adriana Baptista

Escola Superior de Educação do Porto, IPP
Centro de Linguística da Universidade de Lisboa

Ainda que seja comumente aceite que a leitura de imagens pressupõe modalidades de processamento que ultrapassam a percepção visual e que quando dizemos leitura não nos referimos a uma mera actividade de reconhecimento (para a qual bastaria eventualmente dominar uma gramática visual constituída por algumas regras de representação mimética do real, intuitivamente adquiridas), não é ainda muito claro exactamente de que falamos quando falamos em ler imagens.

Se por literacia visual se entende consensualmente a capacidade para compreender, produzir, manusear e fruir significados veiculados através de imagens picturais, já o conceito restrito de leitura, se aplicado a imagens, nem sempre mostra ser consensual. Apesar disso, actividades ditas de leitura de imagens são cada vez mais actividades propostas em contexto escolar (fazem mesmo parte das orientações curriculares dos programas de Português e Educação Visual no Ensino Básico), ainda que nem sempre se defina com rigor como se desenvolve essa competência de leitura ou mesmo o que se pretende que o aluno saiba e faça. Ler uma imagem pode ser equivalente a percepcioná-la atenta e demoradamente, a identificar os elementos que a compõem, a ser capaz de a descrever ou, tão-só, a ser capaz de a rotular através de uma expressão ou frase. Legendar imagens é uma das actividades para a qual os alunos são frequentemente solicitados, como se fosse evidente que quem lê, legenda. Ou seja, ler uma imagem parece estar contaminado pela dimensão da decifração verbal da leitura exigindo-se, assim, que, a cada imagem percepcionada, seja aposta uma etiqueta verbal (oral ou escrita) mais ou menos extensa, (de carácter designativo, descritivo, explicativo e/ou narrativo) entendida a transdução para a linguagem verbal como a consequência inevitável da percepção visual.

Parece, pois, que o entendimento que se faz de “ler uma imagem” não está muito longe do conceito de decifração, apesar de a leitura de imagens incluir tarefas que não são de todo mecânicas, como: identificar o que uma imagem representa, isolar um item, de entre todos os que compõem a imagem, para ser o foco

visual relevante, ou ignorar alguns itens individuais presentes para processar um conjunto morfológicamente coerente e sintacticamente significativo, inferir de que contexto temporal e espacial foi retirado aquela espécie de *frame* que congelou um determinado momento, adivinhar o que aconteceria (aconteceu) depois daquele momento, compreender os objectivos daquela representação particular e processar a sua decifração com o apoio da linguagem verbal. Está, assim, por definir o que se ensina quando se ensina a ler uma imagem, se é ou não possível a decifração do código visual, quais as estratégias cognitivas para aceder à compreensão de uma imagem e qual o papel da palavra (presente ou ausente da área de percepção) na leitura de imagens.

Podemos hoje dizer, na linha do que propuseram Bauret (1990), Thibault-Laulan (1973), Saint-Martin (1987), Dondis (1976) e Oliveira (2006) que, para além da descodificação visual dos elementos representados, a leitura de imagens pressupõe a competência para a interpretação de elementos semânticos (através dos quais seria possível fazer a categorização dos elementos denotativamente representados e dos seus valores conotativos relevantes para a estrutura retórica da imagem) e de elementos pragmáticos (que permitem a identificação dos objectivos de reprodução da imagem).

Mas apesar de Goldsmith (1984), na pegada de Colin Cherry, distinguir na leitura de imagens *readability* and *understanding* e defender que nas imagens não há compreensão sem decifração, há, provavelmente, no caso da leitura de imagens, a necessidade de discutir se podemos falar, de facto, em decifração. Provavelmente, a transposição do termo decifração, referenciado na literatura científica para a leitura de textos verbais, é abusiva quando este passa a ser usado para a leitura de imagens dado que primeira actividade desenvolvida perante a percepção de uma imagem é antes uma actividade de identificação e não de decifração. A decifração, como sabemos, normalmente alvo de aprendizagem sistemática, pressupõe o domínio de um código de sinais e das suas regras combinatórias, de convencionalidade inequívoca em cada língua, enquanto que a identificação dos elementos picturais nas imagens, mesmo que estas, por não se servirem de códigos arbitrários e convencionais, possam ser lidas de modo diferente por diferentes observadores, é, frequentemente, apoiada pelo mimetismo representativo. Assim, a identificação de um qualquer item representado numa imagem (Biederman (1987) defende que estes necessitam apenas que alguns traços estruturais do referente sejam percebidos para que a sua identificação ocorra com sucesso) é fortemente suportada pelos índices de *figuratividade* das imagens (tal como a define Moles, 1981) e pode, eventualmente, ser dificultada pelos índices de *complexidade*, entendidos estes como índices de imprevisibilidade para um determinado contexto visual (Moles, 1981).

Todavia, a identificação de uma determinada imagem não depende exclusivamente dos índices de figuratividade dos elementos que a compõem. Mesmo em imagens com graus de complexidade muito baixos, nem sempre a identificação se processa de forma inequívoca. Tomemos como exemplo uma imagem que representa uma figura humana (sem contexto espacial definido) com uma bata branca e um estetoscópio. Cada um destes elementos, por si só, é um ícone e um grau de figuratividade elevado favorecerá, com certeza, a sua identificação. No entanto, identificar uma figura humana, uma bata e um estetoscópio não é o mesmo que identificar um médico. Para que se identifique a representação de um médico na imagem, alguns dos ícones terão de ser compreendidos como símbolos de uma profissão e não como meros objectos. Quem vê terá, pois, de dominar na imagem, para além dos componentes lexicais, uma estrutura frásica, atribuindo a cada um dos seus constituintes uma função sintáctica. Esta estrutura frásica será diferente daquela que contém a mesma figura humana de bata branca com uma lamela e um microscópio, por exemplo, ou com um protector facial de rede e um favo de mel na mão. Aí, ainda que em nenhum dos casos estejamos perante situações de alotopia, a figuratividade da bata branca (um elemento provavelmente pregnante e central) não é o factor relevante. Na identificação de cada uma das figuras interessa o valor relativo que se dá a cada um dos elementos que a compõe e à sua inter-relação com os outros elementos.

Acreditamos que ambas as actividades que acabámos de descrever e que podem ocorrer de forma sucessiva ou quase simultaneamente são modalidades diferentes de identificação (parcelar e/ou global) dos itens representados na imagem. Todavia, para que a identificação da figura como “médico” ocorra, será necessário que alguns itens sejam interpretados e não apenas identificados. Ou seja, na leitura de imagens, não há compreensão sem identificação, mas algumas identificações podem necessitar da compreensão do valor simbólico dos elementos previamente identificados.

Assumir que a identificação pode incluir diferentes modalidades é assumir a complexidade de uma actividade que, à partida, poderia ser considerada simplesmente perceptiva e apenas dependente do grau de figuratividade da imagem. De facto, cada um dos elementos representados numa imagem pode ser percepcionado, identificado, relativizado, desvalorizado, sobrevalorizado ou mesmo não percepcionado. Tal não seria possível numa actividade de decifração de texto verbal que se processa com base no reconhecimento sistemático de um código apresentado numa estrutura sequencial.

A identificação de um item representado numa imagem pode também ser dificultada quando o item (na sua globalidade ou num dos elementos que o compõem) não é conhecido do sujeito que o percepciona. Logo, o desconhecimento

do elemento representado, i.e. a inexistência de uma relação experiencial com o referido elemento ou com as estratégias para a sua representação (cf. Nodelman, 2000), pode ser suficiente para que este, ainda que rigorosamente representado, possa eventualmente não ser identificado.

Sabendo que usamos designações verbais para o que conhecemos e que o mapa neurológico, proposto por Levelt e Indefrey (2000), identifica para a nomeação de imagens e para a produção de palavras uma mesma área cortical, coloca-se desde logo uma questão: é ou não o processo de identificação de imagens suportado por um processo de nomeação e/ou designação? Ou seja, dito de outro modo, pode a identificação de um item percebido visualmente realizar-se de forma independente da sua nomeação?

Parece-nos, na linha do que discutimos para a relação referencial da imagem com a legenda (Baptista, 2009), que, quando o processo de identificação de uma imagem se restringe ao reconhecimento do objecto representado, é coadjuvado por processos de nomeação (onde a palavra serve apenas para consubstanciar a representação do objecto para o leitor da imagem), mas quando a identificação pressupõe a compreensão da relação dos objectos percebidos e identificados, este processo é coadjuvado por estratégias de designação (que evidenciam a relação da palavra com o mundo).

Hoje em dia, muitas imagens, talvez uma esmagadora maioria, apresentam índices de complexidade significativos dado que optam por estratégias de representação que privilegiam o desvio (em graus diferentes e com intenções comunicativas também distintas) face à representação mimética do real. Esse desvio, que pode ir de uma simples hiperbolização de uma característica (como no caso da caricatura), à sínquise gráfica de dois elementos de universos semânticos diferentes (como, por exemplo, na publicidade ou na ilustração infantil, o caso da antropomorfização de uma régua) à metonimização de um elemento gráfico (como no caso da utilização de uma mancha vermelha para a representação da dor, do pânico ou da guerra), pode ser um factor de expressividade gráfica capaz de favorecer a identificação do item representado ao salientar uma das suas características. No entanto, o mesmo desvio também pode ser um factor de complexidade semântica que obstaculiza o processo de identificação do item representado, quando, dada a sua forma ou a sua relação imprevisível com o real experimentado, não é reconhecido pelo observador, incapaz de estabelecer uma relação entre a dimensão retórica da representação e o item em questão. Ou seja, a dificuldade na interpretação de um desvio representativo pode ser um factor que dificulta a identificação do item representado.

Considerámos, pois, o grau e a qualidade do desvio factores informativos importante em cada imagem. Por isso, ler uma imagem, exige sempre perceber o

grau de desvio que a mesma exhibe. Aliás, este será sempre o factor que distinguirá o real, do real representado e que tornará significativa uma dada representação e não outra. Ler uma imagem é, pois, na sua essência, identificar o elemento representado, apesar do eventual desvio face ao real que este possa exhibir, mas também, e simultaneamente, compreender o valor expositivo, narrativo e estético desse mesmo desvio num dado contexto de apresentação.

No caso da leitura de textos verbais, ainda que saibamos que esta compreende a decifração e a compreensão, aceitamos que possa haver decifração de um texto (enquanto processo mecânico de identificação gráfica de um código e sua correspondência fónica) sem que o leitor compreenda o significado literal e/ou não literal do texto que decifrou. Ora, nas imagens, a decifração não existe enquanto fase mecânica da leitura, nem propedêutica da compreensão. Se bem que a compreensão sem a identificação dos elementos que constituem as imagens não seja possível, é preciso ter em conta que a identificação dos elementos constitutivos da imagem pode ocorrer simultaneamente com a compreensão da imagem ou sem que a compreensão dos seus sentidos se verifique. A leitura, porém, exige sempre algo para além da identificação dos itens representados. Todavia, erradamente, nas imagens – tal como, aliás, para alguns, a decifração nos textos verbais – a identificação dos elementos representados é, frequentemente, tida como sendo leitura.

Partiremos, aqui, do princípio que as gramáticas da linguagem verbal e pictural são obviamente muito distintas e que a utilização de uma metalinguagem teórica emprestada da linguística arrastará sempre consigo um enfoque analítico específico, mas aceitaremos que as imagens são uma forma muito particular de linguagem e que, apesar de não poderem ser decifradas, podem ser lidas. Mais concretamente, defenderemos (de acordo com Noth, 1995) que as imagens são capazes de veicular conteúdos discursivos (dicênticos e argumentativos, tal como a linguagem verbal, quando referenciadas a um contexto de ocorrência) e que por isso se constituem enquanto textos disponíveis para ser lidos (na acepção da semiótica cultural, defendida por Sonesson (1998), onde tudo o que pertence a um sistema particular de representação pode ser considerado um texto).

Consideramos, pois, que a leitura de uma imagem é uma actividade complexa (de identificação e compreensão) capaz fornecer ao leitor informações não só sobre o que está representado, mas também sobre os objectivos pragmáticos e estéticos da sua representação. E defendemos que uma parte significativa dessas informações pode ser obtida através da análise do desvio representativo, elemento apenas aparentemente inócuo. É, aliás, a identificação dos valores deste desvio face ao real que nos permite distinguir desenho, pintura e fotografia, habilidades gráficas de incapacidades representativas e atribuir a cada um dos traços que

compõem a dimensão representativa da imagem um valor significativo. Assim, para lermos uma imagem, precisamos também de saber gerir as informações que a mesma nos fornece sobre as dimensões técnicas utilizadas na representação.

Ainda que a dupla articulação dos elementos que constituem a imagem seja muito particular e distinta da que se observa na linguagem verbal (Biederman (1987), por exemplo, identificou 36 *geons* – *geometrical ions* – como unidades básicas da linguagem visual, contra os 55 fonemas existentes em todas as línguas do mundo, e descobriu que o reconhecimento visual de um objecto conhecido pode ser feito mediante apenas três *geons*, enquanto um objecto desconhecido, obscuro ou fora de contexto exigirá um número significativamente maior de *geons* para ser identificado), a competência para a identificação dos elementos estruturantes da linguagem visual (ao nível lexical como a cor, a linha ou o ponto, ao nível morfológico, como os padrões ou a perspectiva e, ao nível sintáctico, como a escala ou o movimento) pode ser adquirida e desenvolvida desde tenra idade. Esta competência permite que os sujeitos sejam capazes de perceber que se encontram a perceber imagens gráficas e não a realidade e sejam também capazes de perceber facilmente, na maioria das representações gráficas, o grau de similitude e o grau de desvio entre o objecto representado e o real de molde a que a sua identificação ou a identificação dos seus desvios seja possível.

Queremos ainda referir que, para que a leitura da estrutura discursiva de uma imagem seja possível, não basta, pois, identificar os seus elementos, é necessário introduzir na sua estrutura de simultaneidade uma ordem de concatenação desses elementos. Esta ordem de concatenação coerente e coesa dos elementos que compõem uma imagem, que a capacita para ser um texto, passa a existir nas imagens, no momento em que são observadas. O olhar do leitor, comandado pela necessidade de sentido perante a simultaneidade do que observa, organiza os elementos e a sua importância e estrutura padrões de percepção diferentes, conforme a experiência dos observadores e os objectivos de observação (veja-se Buswell, 1935 e Henderson, 2004), comandados pela necessidade de sentido, que organizam a percepção identificando relações retóricas entre os elementos representados (cf Baptista et al., 2011). Uma vez que sabemos que durante a observação de uma imagem, as fixações iniciais efectuadas pelo observador podem determinar a direcção das que se seguem e a compreensão dos dados processados visualmente pode alterar a direcção das sacadas (cf. Becker e Jurgens, 1979), podemos dizer que, ao ver, o observador organiza sintacticamente um tempo gráfico e, ao fazê-lo traz o tempo e a ordem para dentro da imagem.

Até que ponto esta ordem é marcada por uma estratégia de textualização verbal do que é visto, concretizando-se a identificação dos elementos da imagens através de uma espécie de legendagem mental que sobrepõe duas linguagens e

duas gramáticas? Lester (2006) na terceira premissa, de entre as três que enuncia para a sua teoria sintáctica da comunicação visual, postula que: "Images are remembered by thinking about them with words." e, ao fazê-lo, coloca irremediavelmente a necessidade da palavra na leitura da imagem, uma vez que a leitura não é apenas percepção e a memória é activada em alguns dos processos de identificação e em todos os processos de compreensão.

Assim, para que a leitura da imagem se processe, torna-se necessário que o contexto da imagem, a nível da gestão morfológica e sintáctica do espaço e do tempo gráficos, seja significativo para quem a observa. Estes dois eixos (que, na linguagem verbal, se prendem, um, com a descrição e outro, com a narração) são fundamentais para a compreensão da imagem, ainda que esse espaço e tempo (observado e intuído) sejam sempre descontextualizados, ou seja, ainda que as imagens, apesar de exibirem de forma simultânea uma panóplia eventualmente significativa e coerentemente estruturada de elementos gráficos que ajudam a construir o cenário e um acontecimento (decorrido ou iminente), sejam sempre, no seu todo, um recorte, uma janela emoldurada das representações possíveis do real, uma frase espacial do mundo, artificialmente inserida num contexto de apresentação, e não, como abusivamente dizemos, a representação do mundo.

Sempre que percebemos uma imagem, se identificamos correctamente os seus elementos, a forma como se articulam entre si e o espaço gráfico que estes ajudam a construir, identificamos o contexto a que pertencem, ainda que dele vislumbremos uma elíptica fatia, mas também o contexto em que a imagem é exibida. Se virmos um veleiro dentro de uma garrafa, um homem sob a chuva afaçado a enrolar, em esforço, as velas do veleiro, um veleiro com o casco seco meio pintado, completaremos diferentemente o contexto onde se enquadra cada um dos veleiros e a identificação dos diferentes contextos ajudará a desencadear inferências sobre o antes e o depois do fragmento representado e a promover leituras diferentes dos objectivos pragmáticos e/ou estéticos da imagem. Por vezes, o balanço entre estes dois contextos (o de onde saiu a imagem e aquele em que a imagem foi inserida) é suficientemente significativo para compreendermos os objectivos de apresentação da imagem. Por vezes, é insuficiente e a imagem reclama um título, uma legenda, um texto anexo.

De tudo o que ficou dito decorre que a actividade de leitura de imagens mantém com a palavra uma relação complexa: primeiro porque esta contribui para a nomeação dos elementos percebidos, depois porque permite que os mesmos sejam designados, depois porque, apesar da ausência de decurso temporal na imagem, a palavra ajuda a textualizar o espaço das imagens fixas em estruturas narrativas. É, aliás, surpreendente vermos que, de acordo com O'Connor et al. (1999), quando solicitados a descrever verbalmente um conjunto de imagens

(fazendo uma legenda ou listando palavras ou frases que as descrevessem), os observadores usam descritores adjectivais apenas numa percentagem de 20%, enquanto os descritores narrativos ocorrem em cerca de 80%.

Como se relacionará, pois, a palavra com a gestão sintáctica do tempo na leitura de imagens fixas? Uma classificação útil e operatória dos diferentes tipos de imagens pode ser a proposta por Mitchell (1986, 1994) que distingue imagens perceptuais, verbais, mentais, gráficas e ópticas. De uma maneira muito geral, poderemos dizer que esta classificação distingue imagens captadas fisiologicamente ou mecanicamente (as perceptuais e as ópticas), de imagens produzidas (as gráficas), de imagens imaginadas (as verbais e as mentais). Nesta tipologia, faltam ainda as imagens digitais, uma categoria muito complexa de imagens virtuais, transformáveis tecnologicamente, que se deixa perceber em estádios diferentes de desenvolvimento, passíveis de se deixar armazenar. Uma vez impressas adquirem o estatuto de imagens gráficas.

Gostaríamos de identificar algumas características para opor estes diferentes tipos de imagens. Consideramos que as imagens gráficas e as ópticas têm uma dimensão material e um contexto de apresentação e que a mesma imagem pode ser vista e lida por diferentes pessoas. Estas congelam um determinado momento de uma forma fixa, através de uma determinada técnica de representação ou captação e é esse mesmo momento que pode ser percebido. As imagens verbais e neutras só podem ser vistas pelo sujeito que as constrói e podem ser permanentemente modificadas; são desencadeadas por textos verbais, pelo contacto sensorial com o mundo e pela imaginação e são imateriais e fluidas. Eventualmente, podem ser transmitidas a outrem, descritas com a ajuda da palavra ou de imagens gráficas. As imagens perceptuais dependem do órgão da visão, mas as mentais e verbais podem ser *vistas* de olhos fechados. Johnson-Laird (1983: 165) define as imagens mentais como o correlato perceptual dos modelos mentais que são normalmente construídos a partir de experiências concretas e articulando informações de diferentes *media* e, com Posner e Raichle (2001) ficámos a saber que a formação de imagens mentais activa áreas extra-estriadas que também são activadas na percepção visual, misturando-se assim o processamento de contributos sensoriais com o processamento de padrões visuais puramente mentais.

Teoricamente, tudo o que origina imagens visuais perceptuais, pode ser registado mentalmente, graficamente ou opticamente. As possibilidades de registo das imagens ópticas são grandes e permitem visionar o que não é perceptível pelo olho humano. O registo óptico é susceptível de gerar imagens incompreensíveis. Todas as imagens gráficas podem ser recordadas como imagens mentais, todas as imagens gráficas e ou ópticas podem gerar novas imagens mentais, todas as

imagens mentais podem ser transformadas em imagens gráficas. As imagens mentais não podem ser registadas nem percebidas.

Defendemos que o leitor que percebe imagens acede às suas informações denotativas e conotativas, organizando sintacticamente o tempo mental ou gráfico (intuindo um antes e um depois do momento representado) e organizando morfológicamente o espaço mental ou gráfico, estabelecendo relações de isotopia ou alotopia significativas entre os elementos representados) elaborando, pois, uma estrutura narrativa e/ou descritiva o que autoriza frequentemente a textualização de narrativas quando se vêem imagens.

Neste texto em que reflectimos sobre algumas particularidades da leitura de imagens, de entre os diferentes tipos de imagens com que coabitamos quotidianamente cotejaremos apenas dois casos, que nos ocuparam em diferentes momentos de um percurso de investigação¹, nos quais a leitura nos pareceu muito interessante e muito complexa: o caso da leitura da imagem mentais de um espaço urbano, activada para elaborar uma instrução de percurso e o caso da leitura de imagens gráficas retoricamente elaboradas do discurso publicitário onde se pretende que uma narrativa alegórica ajude a identificar as características de um produto.

No início do filme *As Asas do Desejo* (1987), de Wim Wenders, a cidade de Berlim e os que a habitam são frequentemente apresentados através de um ponto de vista fantástico. O espectador é levado a questionar-se sobre quais são os olhos que vêem o que ele próprio está a ver e como pode a personagem que empresta o ponto de vista à imagem estar dentro do campo de visão das outras personagens, como se fosse uma delas. Por vezes, o ponto de vista de uma dada sequência usa planos picados e é quase uma metáfora do olhar onisciente de Deus, ou melhor dizendo, neste caso particular, do olhar perscrutador e silencioso de duas personagens sub-repticiamente presentes e sombrias que descobriremos, mais tarde, serem os anjos Daniel e Cassiel. Estes, apesar de incorpóreos e invisíveis para as demais personagens da diegese (excepto para as crianças), aparecem no ecrã, para que os vejamos, adossados às paredes, disfarçados nos pináculos das catedrais, por um instante, instáveis e alados, no abrupto beiral de um edifício.

1. Um deles, entre 1989 e 1991, durante a elaboração da tese de Mestrado em Linguística Portuguesa Descritiva/Sociolinguística, na FLUL, quando, sob a orientação da Professora Isabel Hub Faria, reflecti sobre a verbalização das representações mentais do espaço nas instruções de percurso, e outro quando, em 2008, desenvolvi um breve estudo sobre uma imagem publicitária dos *Alarmes Cobra*, para ser apresentado na Lição do concurso a professora coordenadora da ESE/IPP, arguida pela Professora Isabel Hub Faria.

Frequentemente, sentimo-los (não os vemos) a pairar sobre a cidade, como se a sobrevoassem, atentos ao perigo que todos os pensamentos humanos, tornados audíveis, encerram. De facto, a maior parte das vezes pensamos que os podemos ver (ou seja, que são visíveis) porque a imagem que aparece no ecrã, que em tudo se assemelha às que percebemos visualmente, corresponde a uma percepção que não pode ser a nossa e não se identifica com a de mais ninguém visível no cenário. De facto, só os imaginamos. Vamos percebendo que são personagens bizarras, capazes de ver de um ponto de vista não humano, de ouvir vozes, de pensar com palavras não proferidas e não audíveis e de sentir a dor alheia. Assim, do mesmo modo que sabemos que vêem e ouvem o que os outros não vêem nem ouvem, sabemos que ninguém os pode ver e não nos surpreendemos que a película nos mostre duas personagens que, misturadas no meio de todos os restantes humanos, não sejam visíveis. Coabitam o espaço e o tempo diegéticos, mas de forma alotópica. A impressão de que estão sempre presentes é tal que, mesmo quando não estão no nosso campo de visão, sentimos as suas asas a roçar as nossas cabeças. Pensamo-las como anjos zeladores da imaterialidade dos pensamentos e das frustrações. A sua missão parece ser a de estar presentes, quais guardas pretorianos no sofrimento dos homens.

No início deste filme, enquanto espectadores somos, também, ouvintes de um estranho discurso verbal. O ambiente sonoro no filme está sobrepovoado de palavras e frases ditas que não vemos ser articuladas pelas personagens. Vamos, pois, tomando consciência de que o nosso acesso a estas vozes mentais é mediado por um meta-ouvido. A forma como o ponto de vista sonoro se harmoniza com a deslocação no espaço cénico das duas personagens (sempre que os anjos se aproximam dos sujeitos ouvimos as suas frases, quando se afastam, deixamos de as ouvir) leva-nos a pensar que ouvimos apenas o que elas ouvem, ou seja, ouvimos pelo angelical ponto de vista sonoro de Daniel e Cassiel. Mas, esta espécie de som rumorejado suscita-nos uma imensa estranheza, dado que deixa a nu uma sintaxe algo desarticulada, omissões, sobreposições, repetições e silêncios e torna evidente que, na ausência da familiar linearidade verbal, ainda que familiarizados com o léxico, os nossos índices de incompreensão semântica são muito superiores ao normal. Somos capazes de perceber que o que ouvimos não são verdadeiros sons físicos, mas uma espécie de voz mental que dá corpo às representações mentais proposicionais que fazem parte dos pensamentos humanos com uma mansa profusão de palavras minimamente articuladas, mas não totalmente disponíveis para ser compreendidas porque provavelmente coexistiriam melhor se emparelhadas com uma pluralidade de imagens mentais. Através dos ouvidos dos anjos ouvimos o que as personagens pensam, mas Wim Wenders, apesar de cineasta, não nos mostra o que vêem enquanto pensam.

Esta pequena particularidade, só por si, constrói no filme um ambiente de maravilhamento, onde vemos um mundo real e ouvimos um discurso interior que não se compatibilizam entre si porque os detentores dos pensamentos a cujas palavras e frases acedemos, pensam uma realidade que nada tem a ver com o espaço onde se movimentam e que não é a que vemos.

O filme só adopta um ponto de vista (visual e sonoro) verdadeiramente compatível com as capacidades humanas quando o anjo Daniel questiona e repudia a sua condição de anjo, abandona o espaço etéreo que habitava e aterra no solo, logo secundado da armadura dourada que o atinge na frente e o fere. Liberto da sombria e muda dimensão do preto e branco, deixa de ser anjo e ganha, como pretendia, cor, sangue, sensações e voz. Fica, portanto, capaz de falar.

Surpreendentemente, as primeiras frases que profere na sua quase infante condição humana são uma instrução de percurso. Questionado sobre onde fica a praça de Postdam, apressa-se a ser útil. A instrução que dá é, no entanto, desconcertante. Sem memórias de passeante na cidade, recupera um mapa mental, uma espécie de planta da cidade, própria dos tempos em que, enquanto anjo, a sobrevoava e profere uma instrução indecifrável para quem da cidade não conhece a vista aérea, mas apenas um emaranhado de ruas onde põe os pés e onde os edifícios, que não são apenas linhas traçadas no solo, funcionam como barreiras, mas também como *landmarks* visuais. Uma cidade onde as ruas não se cruzam apenas para construir um reticulado da simultaneidade, antes se sucedem, se percorridas, e exigem que um passeante flexione à esquerda e à direita para serem percebidas. Ver uma cidade de cima é como ver pensamentos. Para que esta se torne visível e compreensível para a inteligibilidade humana é preciso percorrela para que o discurso sobre ela discorra e não plane.

Esta simbiose entre a experiência do mundo, a construção de mapas mentais e a sua verbalização coloca a relação entre a imagem e palavra verbal numa dimensão complexa que expõe uma situação em que um sujeito lê a sua própria imagem mental da cidade, construída com base na sua experiência de passeante alado e a traduz em palavras que não a deixam ver, antes a obliteram porque não se aproximam em nada da experiência de um passeante pedestre. O estudo sobre instruções de percurso que a instrução do anjo Daniel desencadeou veio mostrar (cf. Baptista, 1991) que a verbalização de imagens mentais do reticulado urbano usa uma geometria projectiva onde o sujeito que constrói os eixos estruturantes da imagem se projecta com eles na instrução, executando verbalmente acções sobre a sua imagem mental do espaço consubstanciadas em verbos de movimento como *virar, cortar, contornar, subir, descer, seguir*. Como então, concluímos (Baptista, 1991: 124): “Foram eles [os verbos], portanto, os obreiros da tradução da tridimensionalidade, através da integração do elemento temporal que, sequenciali-

zando as etapas, permitiu a desmontagem de um modelo sem a sua desestruturação de modo a que pudesse seguidamente ser, de novo, arquitectado pelo ouvinte. Não resistimos a referir o tempo como a arquitectura intelectualizada do espaço tridimensional.”

Este tempo de que falávamos, em 1991, ao analisar uma série de instruções de percurso, fornecidas por polícias, arquitectos e taxistas, e de que os verbos são os obreiros, é o tempo estruturante do próprio texto instrucional da instrução de percurso.

Vejamos agora o caso da imagem publicitária aos alarmes Cobra, cuja leitura testámos com e sem a presença do texto designativo que figura no canto inferior esquerdo da imagem.



FIG. 1 IMAGEM PUBLICITÁRIA, DA AGENCIA NAGA DDB DA MALÁSIA, AOS ALARMES COBRA PARA CARROS, 2007

A cena retratada decorre num parque de estacionamento, onde se vêem dois carros e dois personagens orientais. Um magro, semi-ajoelhado, tenta assaltar um dos carros, um BMW, servindo-se provavelmente de uma gazua para abrir a porta cuidadosamente; tem um gorro verde escuro na cabeça, uma carteira-estojo à cintura, próximo dos joelhos algumas ferramentas e olha concentradamente a fechadura da porta. O outro personagem, colocado ligeiramente atrás e à esquerda do pretenso assaltante, é obeso e veste o único ponto de cor pregnante da imagem, uma túnica vermelha, tal como a cabeça da cobra que dá corpo ao símbolo da marca de alarmes que acompanha o texto do anúncio. Este personagem usa turbante, está estático, fitando de esguelha, mas fixamente os movimentos do pretenso assaltante e segura, nos dois braços abertos, dois címbalos.

O movimento compreendido na cena é, pois, mínimo já que um dos personagens não executa qualquer acção, mas a compreensão deste quadro como uma narrativa depende exactamente do receio do seu movimento iminente. No momento em que o carro for assaltado, o ruído do alarme, personificado neste músico, soará, assim que este junte os braços, batendo os címbalos, quais mandíbulas de uma cobra a morder-nos.

Na leitura da imagem verificou-se que a presença do texto designativo era fundamental para a compreensão cabal da imagem e dos seus valores retóricos (cf. Baptista, 2008), mas mais importante do que o papel da palavra na compreensão da imagem, gostaríamos de registar o papel da palavra na identificação dos itens representados iconicamente. Ou seja, quando o texto está presente e a compreensão da imagem se processa (apesar de apenas um sujeito no estudo referir o nome “Cobra alarmes” na descrição da imagem), o tocador de címbalos recebe designações como “alarme” e “senhor-alarme” e os címbalos são também designados como “alarme a disparar”; por seu lado, o homem ajoelhado é um “ladrão” ou “assaltante”. Ou seja, a leitura da imagem processa-se comandada pela presença da designação do alarme e todos os elementos alotópicos são enquadrados numa leitura significativa e coerente, onde tudo se canaliza para introduzir na imagem aquilo que lá não está de facto, mas que foi designado verbalmente: o alarme. Assim, o mapa lexical usado na descrição da imagem organiza-se ao redor do tema *assalto* e *alarme* e os personagens e objectos não são nomeados, mas designados dentro da retórica narrativa do assalto. A designação *Alarmes Cobra* é suficiente para congelar a acção e topicalizar o suspense. Apesar da imagem ter um franco valor narrativo, a cena é descrita apenas do ponto de vista iminente, sendo referido sistematicamente e com expressividade que o alarme está “prestes a disparar”.

Escolhemos estes dois casos para ilustrar o papel da palavra na leitura de imagens porque nos parecem antípodas. Nas instruções de percurso partimos de uma imagem mental sem discurso narrativo para a construção de um percurso, ou seja, através da palavra, lemos a imagem temporalizando fortemente o espaço; na alegoria narrativa da publicidade aos *Alarmes Cobra* partimos de uma narrativa visual, onde uma pluralidade de acções decorrem ou estão prestes a decorrer para designar um produto, ou seja, lemos uma imagem, que apesar de fixa tem um enorme cinetismo, congelando uma acção, através da estratégia verbal de lhe apor uma simples designação. No primeiro caso, o texto lineariza a simultaneidade da imagem, ou seja, a palavra gere sintacticamente o tempo gráfico da imagem; no segundo caso, a palavra designa, ou seja, organiza morfológicamente o espaço gráfico, dando sentido a todas as alotopias e avaliando o grau de desvio face ao mundo como retórico e significativo.

Ambos os exemplos nos parecem provar que a leitura de imagens sobrevive numa estreita relação com a palavra, que a palavra na leitura de imagens não corresponde a uma mera decifração e que influencia a leitura quer ao nível da identificação, quer ao nível da compreensão.

REFERÊNCIAS BIBLIOGRÁFICAS

- BAPTISTA, A. (1991). Da representação mental do espaço à sua verbalização. Tese Mestrado. FLUL.
- (2008). O legível e o visível: as relações da palavra com a imagem. Texto não publicado da Lição na candidatura ao concurso de Provas Públicas para professor coordenador na ESE/IPP.
- (2009). Interações Texto/Imagem: o caso particular da legenda de fotografia. Fundação Calouste Gulbenkian/Fundação para a Ciência e Tecnologia.
- BAPTISTA, A.; FARIA, I. e LUEGI, P. (2011). *Hybrid processing of rhetorically elaborated texts*. In CICOM Congress Communication, Cognition and Media Proceedings. Braga: Aletheia.
- BAURET, G. (1990). De l'esquisse d'une théorie à la dernière aventure d'une pensée. Les cahiers de la photographie, 25, 7-13.
- BECKER, W. E. e JURGENS, R. (1979). Analysis of the saccadic system by means of double step stimuli. In *Vision Research*, 19: 967-983.
- BIEDERMAN, I. (1987). Recognition by components. A theory of Human image understanding. *Psychological Review*. Vol. 94, n.º 2, 115-147.
- BUSWELL, G. (1935). *How people look at pictures. A study of the psychology of perception in Art*. Chicago: university of Chicago Press.
- DONDIS, D. A. (1976). *La sintaxis de la imagen*. Barcelona: Gustavo Gili.
- GOLDSMITH, Evelyn (1984). *Research into illustration: an approach and a review*. Cambridge: Cambridge University Press.
- JOHNSON-LAIRD, P. N. (1983). *Words and images in the age of technology*. Cambridge: University Press.
- (1986). *Mental Models*. Cambridge. MA: Harvard University Press.
- HENDERSON, J. M. e FERRREIRA, F. (Eds.). *The interface of language, vision and Action: Eye movements and the visual world*. New York: Psychology Press.
- LANGER, S. K. (1960). *Philosophy in a new key*. New York: Charles Scribner's Sons.
- LESTER, P. M. (2006). *Syntactic Theory of Visual Communication*. disponível em <http://commfaculty.fullerton.edu/lester/wrtings/viscomtheory.html>.
- LEVELT, J. M. e INDEFREY, P. (2000). *The speaking Mind/Brain: Where do Spoken Words*

- come from? In A. Marantz, Y. Miyashita & W. O'Neil (Edits). *Image, Language and Brain* (pp. 77-93). Cambridge. The MIT Press.
- MITCHELL, W. T. (1986). *Iconology: Images, text, ideology*. Chicago: University of Chicago Press.
- (1994). *Picture theory, Essays on verbal and visual representation*. Chicago: University of Chicago Press.
- MOLES, Abraham (1981). *L'image, communication fonctionnelle*. Paris: Casterman.
- NODELMAN, P. (2000). *The Implied Viewer: Some Speculations about What Children's Picture Books Invite Readers to do and to be*. In *International Journal of the Centre for Research and Education in the Arts, CREArTA*, Sydney: University of Technology. 1.1 (June 2000): 23-43.
- NOTH, W. (1995). Can pictures lie? The semiotic review of Books, 6, (2). Disponível em: <http://chass.utoronto.ca/epc/srb/srb/pictures.html>.
- OLIVEIRA, Sara (2006). *Texto Visual e leitura critica: o dito, o omitido e o sugerido*. In *Linguagem e Ensino*, vol. 9, n.º 1, 2006 (15-39).
- O'CONNOR, B. C; O'CONNOR, M. K. e ABBAS (2001). Users reaction as access mechanism. An exploration based on captions for images. *Journal of the American Society for information Science*, 50(8), pp. 681-697.
- PAPAFRAGOU et al. (2008). *Does language guide event perception? Evidence from eye movements*. *Cognition* 108 (2008) 155-184.
- POSNER, M. e Raichle, M. (2001). *Imagens da Mente*. Porto: Porto Editora.
- SAINT-MARTIN, F. (1987). *Sémiologie du langage visuel*. Québec: Presse Universitaire du Québec.
- SONESSON, Göran. (1998a)). *The concept of text in cultural semiotics: Text as interpretation*. disponível em <http://www.arthist.lu.se/kultsem/sonesson/TextTartu1.html>.
- THIBAUT-LAULAN, A. M. (1973). *Image et langage*. In B. Poitier (Ed.). *Le langage*. Paris: Centre d'Étude de Promotion de la Lecture.

