

Université de Montréal

Une silhouette naturelle est-elle fréquemment classée dans plusieurs catégories de base?

par
Colin Boudrias-Fournier

Département de psychologie
Faculté des arts et des sciences

Mémoire présenté à la faculté des études supérieures
en vue de l'obtention du grade de maître (M.Sc.)
en psychologie

Septembre, 2013

© Colin Boudrias-Fournier 2013

Résumé

Les silhouettes ambiguës, comme celle du lapin/canard (Jastrow, 1899), ont été étudiées selon plusieurs approches. Toutefois, les figures prises en exemples dans la large majorité des études sont généralement les mêmes. Cette redondance des images ambiguës utilisées pousse à croire qu'elles sont peut-être assez rares. Certaines observations anecdotiques suggèrent cependant qu'elles seraient au contraire relativement fréquentes. C'est ce que cherche à déterminer cette expérience. Nous avons utilisé des modèles tridimensionnels d'animaux projetés de façon aléatoire afin d'en extraire les silhouettes dont la complexité périmétrique a ensuite été modifiée par lissage. Treize sujets ont dû indiquer ce qu'ils percevaient dans l'image. Nous démontrons qu'une silhouette est classée en moyenne dans 1.9079 catégories de base. Nous avons également démontré qu'une diminution de la complexité périmétrique rend d'abord une silhouette plus ambiguë pour éventuellement atteindre un sommet (équivalent à environ six fois la complexité périmétrique d'un disque) à la suite duquel l'ambiguïté chute.

Mots clés : Images ambiguës, Silhouettes, Complexité périmétrique, Illusion, Ambiguïté

Abstract

Ambiguous silhouettes such as the duck/rabbit (Jastrow, 1899) have been studied by several approaches. However, the figures taken as examples in the vast majority of studies are generally the same. This redundancy of the ambiguous images used in literature implies they may be quite rare. On the other hand, anecdotal evidence suggests that they might be relatively frequent. This is what this experiment is trying to establish. We used three-dimensional models of animals from a random point of view to extract silhouettes whose perimetric complexity was subsequently modified by smoothing. Thirteen subjects were asked to indicate what they saw in the image. We show that silhouettes are classified on average with 1.9079 based categories. We also established that a decrease in the perimetric complexity initially makes a more ambiguous figure but that this effect eventually reaches a peak (at a perimetric complexity of approximately 6 times that of a disk) after which ambiguity drops.

Mots clés : Ambiguous images, Silhouettes, Perimetric complexity, Illusions, Ambiguity

TABLE DES MATIÈRES

<u>Résumé.....</u>	<u>iii</u>
<u>Abstract.....</u>	<u>iv</u>
<u>Remerciements.....</u>	<u>ix</u>
<u>INTRODUCTION GÉNÉRALE.....</u>	<u>1</u>
<u>ARTICLE.....</u>	<u>17</u>
<u>Références.....</u>	<u>34</u>
<u>Figures.....</u>	<u>38</u>
<u>Appendice.....</u>	<u>43</u>
<u>DISCUSSION GÉNÉRALE.....</u>	<u>45</u>
<u>Références.....</u>	<u>57</u>

LISTE DES FIGURES

Figure 1. (A) Illusion de la jeune femme-vieille dame (Boring, 1930) (B) Visage-rat (C) Lapin canard (Jastrow 1899) (D) Chef-chien (Rock 1956).....	4
Figure 2. Quelques exemples des élastiques en forme d'animaux de Yumiko Ohashi et Masanori Haneda.....	10
Figure 3. La figure AWK de Kihlstrom (2006) évoquant à la fois une baleine et un kangourou.....	11
Figure 4. Modèle tridimensionnel d'un Jaguar vu de profil, de face, en plongée et en contre-plongée.....	14
Figure 5. Exemple de silhouette ambiguë, un modèle d'abeille avec une complexité périmétrique de 20,03	48

FIGURES DE LA SECTION ARTICLE

- Figure 1. Quelques exemples des élastiques en forme d'animaux de Yumiko Ohashi et Masanori Haneda.....38
- Figure 2. La figure AWK de Kihlstrom (2006) évoquant à la fois une baleine et un kangourou.....39
- Figure 3. Stimuli créés d'après le même modèle tridimensionnel d'éléphant. Les trois silhouettes ont des niveaux de complexité périmétrique de respectivement 59, 46, et 30. La figure a est présentée aux coordonnées sphériques 335.04, -367.49, 51.93, la figure b figure à 147.38 -154.87 451.99.....40
- Figure 4. Graphique représentant le nombre de modes moyens pour quatre niveaux de lissage (Respectivement 3, 6, 9 et 14).41
- Figure 5. Nombre moyen de réponses par sujet selon le niveau de complexité périmétrique de la silhouette.42

Remerciements

Je tiens avant tout à remercier mon directeur de recherche, Frédéric Gosselin, pour le partage de son expertise en recherche, mais également pour sa curiosité singulière. C'est grâce à celle-ci, ayant mené au questionnement initial sur la rareté des silhouettes ambiguës, que j'ai eu le privilège de vivre de première main les aléas de la recherche exploratoire. Son aide fut plus qu'essentielle afin de contourner certains problèmes rencontrés au cours de cette recherche. J'aimerais également remercier Xavier Morin-Duschesne pour sa patience et son aide constante en programmation informatique. L'appui et le support des autres étudiants du laboratoire, particulièrement Nicolas Dupuis-Roy, Catherine Éthier-Macher, et Alexandre Couët-Garand fut également très apprécié. Finalement, j'aimerais remercier ma mère Michèle Boudrias pour ses heures de travail à corriger ma grammaire douteuse. Qui eut cru qu'elle dusse un jour payer si cher les choix d'éducation alternative faits au début de mon parcours académique.

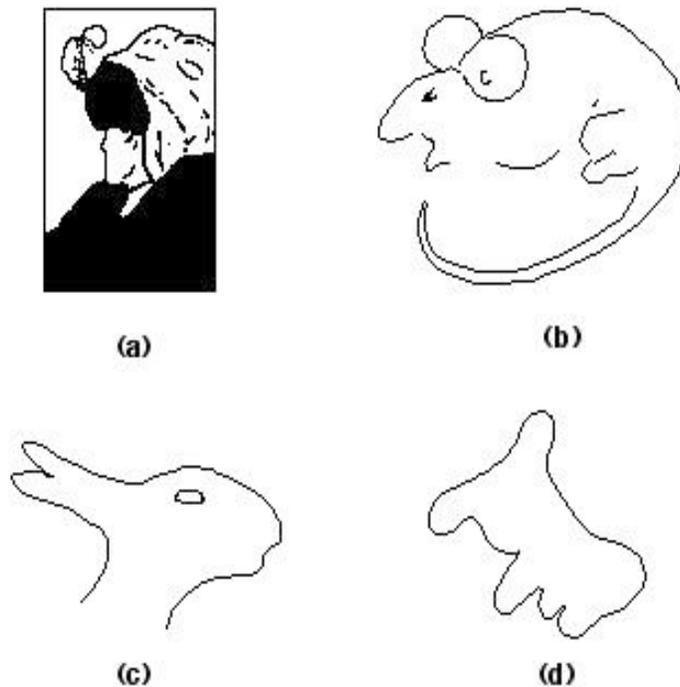
INTRODUCTION GÉNÉRALE

L'environnement qui nous entoure est visuellement équivoque. En effet, un nombre infini de scènes tridimensionnelles sont compatibles avec notre image rétinienne en deux dimensions. La perception visuelle que nous expérimentons à chaque instant est donc le fruit d'une inférence de notre système visuel d'après notre connaissance du monde (Helmholtz, 2001; Gregory, 2002; Marr, 1982; Mamassian & Goutcher, 2001). En général, cette inférence se fait sans le moindre accroc; notre système visuel traite toutes les interprétations possibles afin de rejeter celles qui sont non réalistes et d'arriver finalement à une seule interprétation plausible. Il arrive en revanche, dans certains cas particuliers, que notre connaissance du monde ne réussisse pas à dissiper l'ambiguïté. L'inférence de notre système visuel mène alors à plus d'une interprétation plausible.

De telles images, dites ambiguës, sont connues depuis longtemps. Un sous-groupe d'entre elles, les figures fond-forme, piquaient déjà la curiosité en l'an 100, après avoir été relevées sur d'anciens ornements arabes (Piccolino & Wade, 2006).

Cependant, bien peu d'études se penchent sur la fréquence ou la probabilité de percevoir des images ambiguës dans la vie courante. En effet, la grande majorité des articles portant sur le sujet réutilisent constamment la même poignée de quatre ou cinq figures. Les figures ambiguës généralement utilisées sont bien connues. Les exemples typiques sont la jeune femme et la vieille dame (Boring, 1930), l'illusion chien et chef cuisinier (Rock, 1956) ou encore l'illusion lapin/canard (Jastrow, 1899). Ces images sont présentées à la figure 1.

Figure 1 : Exemples d'images ambiguës soit: Jeune femme-vieille dame (a), Visage-rat (b), Lapin canard (c), Chef-chien (d).



Cette redondance dans le choix des stimuli porte à croire que ces types d'images sont rares et qu'il serait donc difficile d'en créer de nouvelles. Pourtant, Kihlstrom (2006) a relevé un exemple d'une silhouette rudimentaire de baleine qui peut à la fois être perçue comme un kangourou. D'autre part, nous avons constaté en observant des bracelets en caoutchouc représentant la silhouette d'animaux, que plusieurs d'entre eux pouvaient être perçus comme représentant un autre animal. Nous reviendrons sur ces exemples mais, dans les deux cas, il semble que l'ambiguïté soit involontaire et fortuite.

À la lumière de ces observations anecdotiques, il semblerait qu'au contraire, ce genre d'illusions soit assez courant et que, dans un contexte de présentation approprié,

un grand nombre d'images puissent être perçues comme telle. C'est ce que l'étude centrale présentée dans ce mémoire vise à déterminer.

Ce mémoire portera, dans un premier temps, sur la revue des observations empiriques et des approches communément utilisées dans l'étude des images ambiguës, pour ensuite se pencher sur le positionnement de la problématique. La présentation de l'article de ce mémoire sera suivie d'une discussion sur les résultats rapportés par celui-ci. Les limites de cette étude ainsi que des suggestions pour des recherches subséquentes seront finalement présentés.

1. Les figures bistables

La grande majorité des études sur les images ambiguës portent sur les stimuli bistables, par exemple les illusions fond-forme comme le vase de Rubin (1921). Il s'agit de figures pouvant être interprétées de deux façons, mais ces deux perceptions ne peuvent être simultanées. En effet, le passage vers le second point de vue annule la perception du premier. Un grand nombre d'études se sont penchées sur les éléments qui influencent les probabilités de percevoir l'une ou l'autre des images.

Il semble que ce qui détermine le changement de perception dans les images bistables soit lié à un processus cognitif plutôt que physiologique. En effet, pour les figures qui découlent d'une rivalité binoculaire ou d'une illusion à perspectives variables, une adaptation a lieu selon le temps de présentation d'un stimulus (Mamassian & Goutcher, 2005; Sundareswara & Schrater, 2008), selon sa constance ou son intermittence (Pearson & Brascamp, 2008) et selon l'angle de présentation (Rock,

1974). Ces constats sont autant d'indices d'un processus cognitif. Il semble également que les anomalies psychiatriques (Hunt & Guilford, 1933) de même que les lésions cérébrales (Bisiach, Ricci et al. 1999) aient un impact sur la perception de ce type de figures.

De façon plus générale, il semble que la culture du participant (Bagby, 1957) ainsi que le contexte culturel au moment de la présentation influencent la perception de l'ensemble des stimuli ambigus. En effet, l'illusion lapin-canard semble être perçue plus fréquemment comme étant un lapin durant la période de Pâques (Brugger & Brugger, 1993). Finalement, l'arrière-plan du stimulus influence également la perception de celui-ci. Le mouvement à l'arrière d'une image (Kramer & Yantis, 1997) de même que des formes présentées en fond peuvent suggérer une interprétation plutôt qu'une autre (Sundareswara & Schrater, 2008).

2. Perception ambiguë des objets tridimensionnels

Tout comme les images bidimensionnelles, les objets tridimensionnels sont très propices à créer des perceptions ambiguës. Comme nous l'avons abordé plus tôt, leur nature tridimensionnelle force notre système visuel à faire des inférences puisque ces objets ont le potentiel de générer une multitude d'images distinctes. À l'inverse, des objets tridimensionnels distincts peuvent, dans certains contextes, créer des images rétinienne très similaires (Kersten, Mamassian et al. 2004).

Ainsi, l'ambiguïté survient lorsque plusieurs objets peuvent avoir créé la même image perçue. Lorsque c'est le cas, le système visuel est forcé de choisir une interprétation (Kernster, 2004). Les théories visuelles stipulent qu'afin de contourner ce

caractère équivoque, le système visuel fait certaines suppositions basées sur notre connaissance de notre environnement (Hegde & Kersten, 2010). Par exemple, la perception des formes solides est constamment biaisée vers la convexité plutôt que la concavité, un biais robuste à travers une gamme de formes, grosseurs et tâches (Langer & Bühlhoff 2001). Le système visuel suppose également que la source lumineuse provient généralement d'en haut (Mamassian & Goutcher, 2001).

2.1 Le biais *Viewing from above*

Dans le même ordre d'idées, le système visuel fait aussi preuve du biais *viewing from above*. Il s'agit de la tendance du système visuel à percevoir les objets comme étant vus en plongée plutôt qu'en contre-plongée (Mamassian & Landy, 1998). Ce biais explique peut-être le fait que le cube de Necker soit généralement perçu comme vu en plongée, entre 60 et 100% des essais (Sundareswara & Schrater, 2008; Troje & McAdam, 2010).

2.2 Le point de vue canonique

De façon similaire, le point de vue canonique (Palmer, Rosch, & Chase, 1980) représente la perspective la « meilleure » ou la « plus typique » d'un objet. Cette perspective qui permet une reconnaissance plus rapide et facile d'un objet est corrélée au point de vue imaginé par un sujet lorsque qu'on lui demande de se représenter l'objet en question. Il semblerait qu'une vue en oblique, aux trois-quarts, soit généralement canonique (Blanz, Tarr, & Bühlhoff, 1999; Verfaillie & Boutsen, 1995). De plus, pour la

majorité des objets, les vues préférées étaient soit de face ou en plongée. La seule exception était pour des objets représentant des aéronefs (avions, hélicoptère, etc. Objets fréquemment perçus du dessous) pour lesquels on ne retrouvait aucune préférence entre les vues en plongée ou en contre-plongée de l'objet. Finalement, on ne retrouve aucun point de vue canonique avec des objets inexistantes (*nonsense objects*). Ces constatations suggèrent que la familiarité de la prise de vue pourrait être liée au point de vue canonique (Blanz, et al, 1999).

3. Indices utilisés pour la perception d'un objet tridimensionnel

Comment le système visuel nous permet-il d'inférer qu'un objet tridimensionnel est le même malgré les grandes variations de l'image qu'il produit sur la rétine? Le point de vue, l'éclairage, l'occlusion et le fond sont des indices particulièrement importants à la reconnaissance. L'ambiguïté apparaît lorsque les variations dans les propriétés d'une scène changent l'image perçue d'un objet. L'absence de certains indices lors d'une tâche de reconnaissance réduit les probabilités de reconnaissance, de la même façon que le ferait du bruit gaussien dans les théories classiques de la détection d'un signal (Kersten, Mamassian et al. 2004). La privation de certains indices de reconnaissance donne donc une plus grande probabilité aux images d'être ambiguës. Un exemple de représentation d'objets tridimensionnels auquel certains indices essentiels de reconnaissance ont été retirés est la silhouette.

4. Les silhouettes

En effet, les silhouettes sont des figures qui semblent être très enclines à l'ambiguïté. Elles représentent des objets tridimensionnels dans lesquels les indices primordiaux à la reconnaissance comme l'éclairage, la texture (Kersten, Mamassian et al. 2004) de même que l'occlusion (qui permet normalement de savoir quelle partie de l'objet en couvre une autre (Troje & McAdam, 2010)) ne sont pas présents. Les études faites sur ce type d'images ambiguës se concentrent généralement sur l'ambiguïté dans la direction du mouvement de l'objet. Un exemple s'étant répandu très rapidement sur internet est la ballerine de Nubuyuki Kayahara (2003 <http://www.procreo.jp/labo/silhouette.swf>) qui est perçue comme tournant dans une direction dans un peu plus de la moitié des cas (Munger, 2008; Troje & McAdam, 2010).

Cependant, peu d'études se concentrent sur l'ambiguïté au niveau de la catégorie de base de l'objet perçu sans une silhouette, c'est-à-dire une confusion au niveau de l'objet présenté plutôt que dans le mouvement de celui-ci ou dans ses attributs.

5. Exemples d'images fortuitement ambiguës

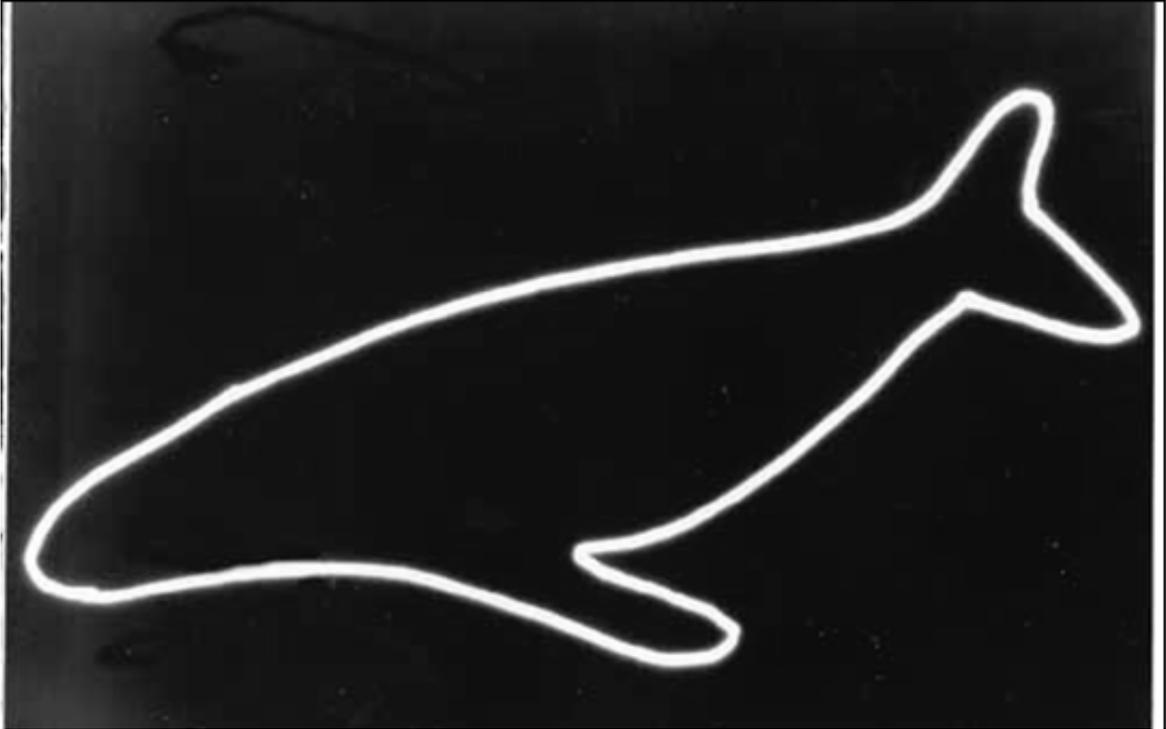
Deux designers, Yumiko Ohashi et Masanori Haneda, ont commercialisé des élastiques qui ressemblent à différents animaux de Zoo (voir Figure 2). En les observant durant un certain temps et sous différents points de vue, nous avons constaté qu'ils peuvent tous être interprétés de plusieurs façons : « la girafe » évoque une baleine, le « kangourou » un phoque, « l'autruche » un escargot, etc. Cette ambiguïté est probablement causée par l'aspect assez rudimentaire des élastiques représentant simplement la silhouette des animaux. Fait intéressant, ces ambiguïtés ne semblent pas avoir été remarquées par les concepteurs, mais créent néanmoins des résultats intrigants.

Tout aussi fortuitement que nous, Kihlstrom (2006) a découvert qu'une image de la silhouette d'une « baleine » dans sa chambre à coucher évoquait un kangourou (voir Figure 3). Il a nommé le stimulus « AWK » (acronyme de « Arizona Whale-Kangourou »). Ces deux exemples portent à croire que les images ambiguës sont fréquentes lors de la présentation de silhouettes peu détaillées.

Figure 2: Quelques exemples des élastiques en forme d'animaux de Yumiko Ohashi et Masanori Haneda <http://www.plus-d.com/arb/index.html>



Figure 3: La figure AWK de Kihlstrom (2006), évoquant à la fois une baleine et un kangourou.



6. Problématique

Il semble donc légitime de se questionner quant à la fréquence d'images ambiguës, plus spécifiquement les silhouettes, dans notre environnement. L'article de ce mémoire est une étude exploratoire visant à recueillir des données descriptives. Nous chercherons à déterminer (1) la fréquence des images ambiguës et (2) si le niveau de complexité d'une image a un impact sur l'ambiguïté de celui-ci.

6.1 Premièrement, en nous basant sur nos propres observations des élastiques en forme d'animaux, de même que celle du AWK de Kihlstrom (2006), il nous semble

raisonnable de croire que les images ambiguës sont assez fréquentes. Nous chercherons donc à démontrer que, dans une banque suffisamment grande de silhouettes d'animaux, un certain nombre d'entre elles seront ambiguës.

6.2 Ensuite, nous nous attendons à ce que plus une image est lissée ou simple, plus ses probabilités d'être perçue comme étant ambiguë augmentent et ce jusqu'à un certain point, au delà duquel l'ambiguïté diminue. Nous postulons qu'après ce seuil, l'image est trop dégradée pour que l'on puisse y percevoir autre chose qu'une tache informe. Par exemple, une image plus grossière comportant un moins grand nombre de géons (des formes tridimensionnelles de base représentés par Biederman (1987) qui, une fois agencées des différentes façons peuvent, selon lui, représenter tous les objets existants) pourra être interprétée de plusieurs façons plausibles.

6.3 Finalement, en nous basant sur le biais du *point of view from above* (Mamassian & Landy, 1998) de même que sur le point de vue canonique (Palmer & al. 1980), il semble raisonnable de prévoir que, sous certains angles, un objet soit plus difficile à reconnaître. En effet, puisque le système visuel semble avoir un biais pour une perception plus familière des objets, il est possible que des stimuli observés sous des angles moins fréquents dans la nature, qui vont à l'encontre de ce genre de points de vue, soient plus ambigus. Par exemple, un objet perçu du dessous sera probablement reconnu moins facilement qu'un autre perçu d'un point de vue typiquement canonique et sera peut-être interprété comme étant autre chose.

Méthode

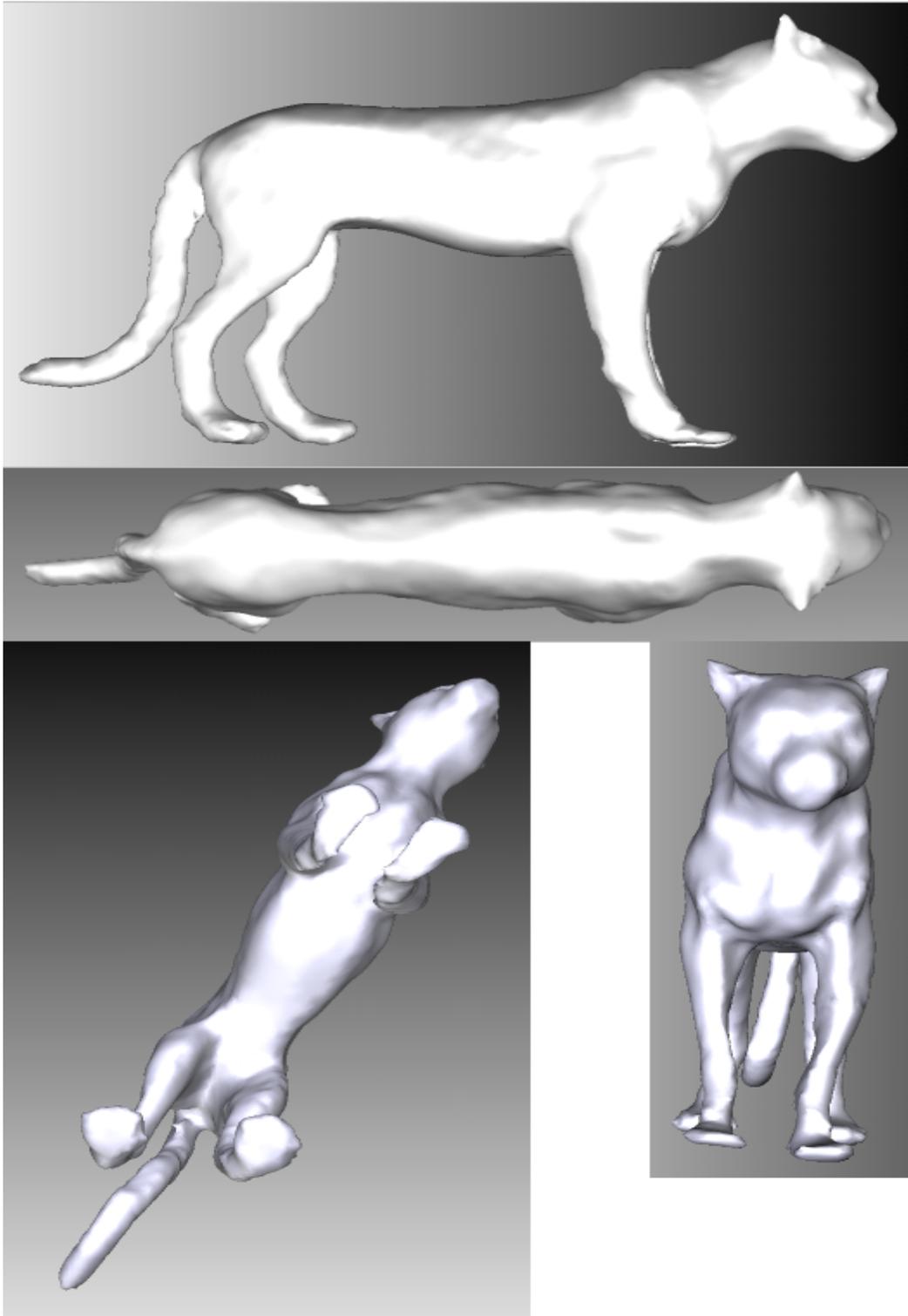
Une banque de 57 modèles tridimensionnels (voir Figure 4) représentant différents animaux a été collectée. Il s'agit d'animaux communs : des mammifères, oiseaux, reptiles, animaux marins et insectes de différentes régions du globe. La banque contient les animaux suivants:

mammifères terrestres : âne, autruche, bison, bouc, chat, chauve-souris, cheval, chèvre, chien (bulldog, chihuahua) , cochon, dromadaire, éléphant, girafe, gorille, hippopotame, jaguar, kangourou, lapin, loup, mouton, ours, porc, rat, raton-laveur, rhinocéros, smilodon (tigre à dents de sabre), taureau, tigre, vache.

oiseaux: aigle (ailes ouvertes), autruche, canard (ailes fermées, ailes ouvertes), cigogne, flamand rose, hibou, moineau (ailes ouvertes et ailes fermées), perruche, pingouin, poule, toucan.

Reptiles/ amphibiens/ animaux marins/autres: abeille, alligator, araignée, caméléon, dauphin, fourmi, grenouille, lézard, papillon, poisson, requin à tête de marteau, requin blanc, serpent, tortue, tortue marine

Figure 4 : Modèle tridimensionnel d'un Jaguar, vu de profil, de face, en plongée et en contreplongé



De ces modèles, un programme a été créé utilisant MATLAB (Mathworks Inc.). Il permet de présenter la silhouette du modèle tridimensionnel sous un point de vue aléatoire.

Cette silhouette était ensuite modifiée selon un filtre gaussien avec un niveau de déviation variable. Le résultat est une image semblable à l'originale, mais plus compacte, ayant une superficie réduite. Ce processus simplifie la figure, réduit son niveau de détails et la rapproche de plus en plus d'un cercle. Il s'agit d'une technique déjà utilisé par Attneave & Arnoult en 1956 pour altérer la complexité d'une image. Le niveau de lissage (Σ) est également déterminé aléatoirement en instaurant cependant un plafond à un certain niveau pour éviter une dégradation excessive. Le programme calcule ensuite le niveau de complexité périmétrique (*Perimetric complexity*, Attneave & Arnoult, 1956) de la figure, i.e. périmètre² / surface. Ce calcul permet d'indiquer la complexité d'une forme sans tenir compte de sa grosseur, son emplacement ou son orientation. Ce rapport est fortement lié à la complexité perçue consciemment par des sujets exposés à des formes (Attneave, 1957) et permet de déterminer le niveau de détail du stimulus. Par exemple, la figure la moins complexe, un disque, possède une complexité périmétrique de 4π (i.e., $(4\pi^2r^2) / (\pi r^2) = 4\pi$). Nous rapporterons la complexité périmétrique comme un facteur de celle d'un cercle suivant la proposition de Watson (2012) : périmètre² / (4π *surface).

Contribution à l'article

À titre de premier auteur de l'article suivant, j'ai contribué à la réalisation expérimentale d'une idée originale de Frédéric Gosselin en opérant la cueillette des données, une partie des analyses ainsi que la rédaction d'un premier jet du manuscrit. De plus, j'ai participé avec Frédéric Gosselin à l'élaboration du protocole expérimental. Xavier Morin Duchesne a écrit le programme expérimental en java script, plusieurs fonctions de production de stimuli, de traitement de données et d'analyse en Matlab.

ARTICLE

Running head: LA FRÉQUENCE DES IMAGES AMBIGUËS

Une silhouette naturelle est-elle fréquemment classée dans plusieurs
catégories de base?

Colin Boudrias-Fournier

Département de psychologie, Université de Montréal

Frédéric Gosselin

Département de psychologie, Université de Montréal

Xavier Morin Duchesne

Département de psychologie, Université de Montréal

Résumé

Les silhouettes ambiguës, comme celle du lapin/canard ont été étudiées selon plusieurs approches. Toutefois, les figures prises en exemple dans la large majorité des études sont généralement les mêmes. Cette redondance des images ambiguës utilisées pousse à croire qu'elles sont peut-être assez rares. Certaines observations anecdotiques suggèrent cependant qu'elles seraient, au contraire, relativement fréquentes. C'est ce que cherche à déterminer cette expérience. Nous avons utilisé des modèles tridimensionnels d'animaux projetés de façon aléatoire afin d'en extraire les silhouettes dont la complexité périmétrique a ensuite été modifiée par lissage. Treize sujets ont dû indiquer ce qu'ils percevaient dans l'image. Nous montrons qu'une silhouette est classée dans 1.91 catégories de base en moyenne. Nous montrons aussi qu'une diminution de la complexité périmétrique rend d'abord une silhouette plus ambiguë pour atteindre un sommet (équivalent à environ six fois la complexité périmétrique d'un disque) puis moins ambiguë.

Notre environnement visuel est très équivoque. En fait, un nombre infini de scènes tridimensionnelles pourraient être compatibles avec une image en deux dimensions projetée sur notre rétine. Notre perception visuelle est le résultat d'une inférence de notre système visuel basée sur notre connaissance du monde et de notre environnement (Helmholtz, 2001; Gregory, 2002; Marr, 1982; Mamassian & Goutcher, 2001). Généralement, cette inférence se fait sans le moindre accroc. Notre système visuel traite toutes les interprétations possibles afin de rejeter celles qui ne sont pas réalistes et de convenir d'une seule interprétation plausible (Kersten, Mamassian et al. 2004). Cependant, il arrive que notre système visuel n'aboutisse pas à une interprétation unique. Ce genre d'images ambiguës, pouvant être interprétées de plusieurs façons, est bien connu. La plupart d'entre elles, comme l'illusion de la jeune femme/vieille dame (Boring, 1930), l'illusion du cuisinier et du chien (Rock, 1956) et l'illusion lapin/canard (Jastrow, 1899) sont fréquemment et presque exclusivement utilisées dans la littérature.

Cette redondance dans le choix des stimuli porte à croire que ces types d'images sont rares et qu'il serait donc difficile d'en créer de nouvelles. Pourtant, Kihlstrom (2006) a relevé un exemple d'une silhouette rudimentaire de baleine qui peut aussi être perçue comme un kangourou (voir figure 1). D'autre part, nous avons constaté en observant des élastiques en caoutchouc représentant la silhouette d'animaux, que plusieurs d'entre eux pouvaient être perçus comme représentant un autre animal (voir figure 2). Dans les deux cas, il semble que l'ambiguïté soit involontaire et fortuite.

Insérer la Figure 1 ici

Insérer la Figure 2 ici

Ces exemples suggèrent que les images simples représentant le contour d'un objet et contenant peu de détails (comme des silhouettes par exemple) pourraient être fréquemment ambiguës. Le but de cette étude est de vérifier si, dans un contexte de présentation particulier, les images ambiguës peuvent être retrouvées fréquemment.

À la lumière des observations mentionnées plus haut, nous supposons que les images ambiguës sont en fait plus fréquentes que ce qui est généralement supposé. La rareté des silhouettes ambiguës que l'on retrouve dans la littérature refléterait plutôt un fossé entre notre perception (capable d'identifier aisément l'ambiguïté) et notre imagination (difficilement capable d'identifier l'ambiguïté).

Dans le but de répondre à cette question, cette étude exploratoire a été conçue afin de recueillir des données descriptives sur l'impact du niveau de détail d'une silhouette sur l'ambiguïté de la perception qu'elle produit.

Nous formulons l'hypothèse que l'ambiguïté des silhouettes croîtra en fonction du niveau de lissage de la silhouette pour éventuellement atteindre un pic puis diminuer, toujours en fonction du niveau de lissage. Nous nous attendons à ce qu'un objet plus complexe, plus détaillé soit plus distinctif. Cependant, une silhouette fortement lissée finira éventuellement par ressembler à une tache. La silhouette risque alors d'inspirer un grand nombre de réponses très variable, comme le ferait du bruit, et de faire chuter l'ambiguïté.

Expérience

Méthode

Sujets: Treize sujets (cinq hommes et huit femmes) ont été recrutés.

Stimuli: Une banque de 57 modèles tridimensionnels d'animaux et d'insectes a été rassemblée. Certains modèles provenaient de la collection de *Princeton Benchmark models* (Shilane, Min, Kazhdan and Funkhouser, 2004), alors que d'autres appartiennent au domaine public. Les 57 animaux sont des animaux communs parmi lesquels on retrouve des oiseaux, des mammifères, des reptiles et des insectes de différentes régions du globe.

Un programme, écrit pour le logiciel Matlab (Mathworks, Inc.), créait des silhouettes à partir de points de vue aléatoires de ces modèles tridimensionnels. L'image produite était ensuite lissée. Ce procédé consiste à dégrader une image à l'aide d'un filtre normal, ce qui donne une réplique embrouillée de l'original. Le contraste est ensuite

augmenté afin d'éliminer les teintes de gris. Les contours de l'image résultante sont ainsi aplanis, perdent de leurs détails (voir figure 3). Un processus similaire a déjà été utilisé par Attneave (Attneave & Arnoult, 1956). Le niveau de lissage était déterminé de façon aléatoire selon un écart type variant entre 7 et 25 pixels dans le but d'éliminer les silhouettes qui auraient été trop aisées ou trop difficiles à reconnaître. Cinq-cent silhouettes ont été créées par modèles, pour un total de 28 500 stimuli. Les stimuli avaient une surface standardisée, chaque silhouette contenait 50 000 pixels noirs présentés sur un fond blanc.

Le programme calculait ensuite le niveau de complexité périmétrique (Attneave & Arnoult, 1956) de la figure. Ce calcul quantifie la complexité d'une forme et donne un résultat invariable selon sa grosseur ou son orientation. La complexité périmétrique est mesurée avec la formule P^2/A , où le carré du périmètre est divisé par la surface de la forme. Nous rapportons ici la complexité périmétrique en fonction de celle d'un disque (i.e., $P^2/(4\pi A)$), l'objet le plus simple, suivant une proposition faite par Watson (2012). La complexité périmétrique d'une forme est corrélée avec la complexité perçue de la même forme par des sujets (Attneave, 1957). Nous nous sommes assurés que la complexité périmétrique était distribuée de manière similaire pour tous les modèles tridimensionnels.

Insérer la Figure 3 ici

Équipement: L'expérience a été programmée avec java script et était hébergée sur une page web créée dans ce but. Les sujets effectuaient l'expérience chez eux sur leur propre ordinateur.

Procédure: Les participants devaient tout d'abord se rendre sur une page web bilingue expliquant l'expérience. Il y était spécifié que des silhouettes représentant des créatures, des objets ou des scènes leur seraient présentées. Il était également indiqué, que certaines silhouettes paraîtraient ambiguës et qu'ils ne devaient pas se limiter dans ce qu'ils cherchaient à y percevoir. Les participants devaient se connecter et s'identifier en utilisant leur adresse courriel et un mot de passe. Ils pouvaient quitter l'expérience à tout moment et la reprendre ensuite à leur discrétion. Aucun nombre maximal ou minimal d'essais n'était spécifié. Le nombre d'essais effectués varie entre 10 et 1302, avec une moyenne de 508,69 essais par sujet et un écart type de 465,89.

Chaque essai se faisait en deux étapes: premièrement, une silhouette était présentée au sujet et celui-ci devait appuyer sur la touche *enter* lorsqu'il y reconnaissait un objet. Une seconde fenêtre apparaissait alors, dans laquelle le participant devait inscrire ce qu'il avait vu.

Préparation pour analyses. Avant les analyses, toutes les réponses ont été traitées à l'aide d'un programme créé pour reconnaître les mots similaires. Le but était de s'assurer que les mots équivalents seraient traités comme tels. Cette fonction fouillait toutes les réponses et demandait à l'expérimentateur quelles réponses devaient être considérées comme équivalentes. Le lexique ainsi créé a pu être révérifié et modifié à la main subséquemment. Toutes les réponses vides ou toutes les réponses désignant des

objets pouvant prendre différentes formes tel que tache, bâton, sculpture d'art moderne, etc. ont été placées dans la catégorie « je ne sais pas ».

Finalement, nous tenions compte du niveau de spécificité des réponses. Nous tentions autant que possible de ramener les réponses à leur catégorie de base (Rosch, Mervis, et al. 1976) comme par exemple: oiseau, chat, chien. Il y avait cependant deux exceptions. Les termes très génériques étaient toujours considérés comme équivalents (mammifère = animal). De plus, les objets qui étaient indissociables du fait que la texture est une caractéristique absente des silhouettes, étaient également considérés similaires. Par exemple, tous les félins, puisque les motifs de leur pelage étaient occultés, ont été considérés comme des réponses équivalentes (chat = puma = léopard).

Résultats : Les réponses données pour chacun des 57 modèles étaient ensuite divisées en quartiles selon leur niveau de complexité périmétrique. L'intervalle entre les seuils minimal et maximal de chacun des groupes se devait d'être variable afin de s'assurer que chaque groupe contenait le même nombre d'essais. Le nombre de réponses distinctes dans chacun des groupes divisé par 57 (le nombre total de modèles) servant d'indicateur de la quantité de réponses distinctes pour chacun des groupes.

Cependant, une simple mesure des réponses distinctes ne représente pas une juste mesure de l'ambiguïté puisqu'une silhouette suggérant une pléthore d'interprétations, à la façon d'un test de Rorschach (1921), se rapproche plus du bruit que d'une image ambiguë telle qu'elles sont généralement définies dans la littérature.

Nous avons élaboré un critère statistique afin de déterminer si une seule

catégorie (verdict de silhouette claire) ou plusieurs catégories (verdict de silhouette ambiguë) reviennent un nombre suffisant de fois pour que l'on puisse rejeter l'hypothèse nulle (verdict de bruit visuel) avec une probabilité inférieure à un seuil acceptable.

Nous savons que nos sujets connaissent au moins 541 catégories de base d'animaux puisqu'il s'agit du nombre de réponses différentes faites durant l'expérience. Nous savons également que chaque modèle 3D a été vu 29 fois en moyenne par niveau de complexité périmétrique. Nous avons calculé à l'aide d'une simulation par ordinateur de type Monte Carlo (voir appendice) la probabilité d'avoir exactement un mode de N réponses identiques par hasard (hypothèse nulle) : pour $N=1$, cette probabilité est égale à 0.39; pour $N=2$, elle est égale à 0.02; et pour $N=3$, elle est égale à 0.0005.

Autrement dit, il est improbable d'observer par hasard exactement deux réponses identiques. Même en tenant compte de la fréquence lexicale des catégories animales, cette conclusion demeure vraie. Nous avons montré, plus généralement, qu'il est improbable d'avoir un ou plusieurs mode de deux éléments ou plus. Nous avons donc fixé notre *seuil d'ambiguïté* à deux réponses identiques. La moyenne de réponses identiques par animal pour chacun des groupes était respectivement de 1.61, 2.21, 2.03 et 1.79 (voir figure 4).

Insérez la Figure 4 ici

Finalement, nous avons envisagé qu'il serait possible que certains sujets répondent la même catégorie de base à chaque fois qu'un modèle 3d particulier leur est présenté. Un mode (une catégorie de base statistiquement significative) pourrait alors être seulement dû à un sujet particulier. La pluralité des catégories de base données pour un même modèle proviendrait, quant à elle, non pas de l'ambiguïté de celui-ci, mais simplement de la diversité des sujets, chacun d'eux répondant une catégorie de base. Afin de vérifier cette hypothèse, nous avons calculé le nombre de catégories de base utilisées en moyenne par sujets (voir figure 5). Les sujets ont produit un nombre équivalent de réponses pour chaque groupe de complexité.

Insérez la Figure 5 ici

Discussion générale

Tout d'abord, les silhouettes semblent étonnamment ambiguës. Nous constatons effectivement qu'en moyenne, peu importe leur complexité périmétrique, les silhouettes ont 1.91 modes en moyenne. Nous constatons également que même les silhouettes les plus complexes, pour lesquelles nous nous attendions à ce que la reconnaissance soit la

plus aisée, sont en moyenne désignées par plus d'une catégorie de base, en l'occurrence 1.8.

Ensuite, cette expérience a aidé à confirmer la relation prédite entre le niveau de complexité périmétrique d'une silhouette et le niveau d'ambiguïté. Nous avons effectivement pu constater que la quantité de modes distincts donnés par animal atteint un sommet pour ensuite chuter. Ceci appuie l'hypothèse voulant que l'ambiguïté augmente avec le lissage d'une silhouette jusqu'à un certain seuil au-delà duquel l'ambiguïté fait place au bruit.

Cette tendance est particulièrement marquée avec quatre groupes de complexité (voir figure 4), où l'ambiguïté est basse pour les silhouettes très peu complexes (1.6 modes en moyenne pour une complexité périmétrique de 3, une ambiguïté moins élevée que les silhouettes les plus complexes), plafonne à une complexité d'environ 6 avec en moyenne 2.37 modes, pour ensuite chuter à 1.63 modes avec les silhouettes ayant une complexité de 14.

Il demeure cependant important de nuancer les résultats obtenus. Nous avons soulevé des cas de figures où la mesure de l'ambiguïté des images a pu être surestimée ou sous-estimée.

Par exemple, le fait que l'ensemble des modèles tridimensionnels choisis soient des animaux peut potentiellement avoir entraîné une sous-estimation du niveau d'ambiguïté. En effet, il est possible que suite à la présentation d'un certain nombre de stimuli très peu lissés, les sujets aient réalisé la récurrence du thème de la faune. Il est donc possible que lorsqu'exposés à des stimuli plus ambigus, les participants se soient

forcés d'y percevoir encore une fois un animal, limitant ainsi leurs propres réponses. Le fait que 49.07% des réponses données sont des animaux suggère cette possibilité. Il serait intéressant d'élargir l'éventail des modèles tridimensionnels à des objets plus hétérogènes en évitant de se limiter à un thème particulier.

A l'opposé, il est également possible que certaines réponses enregistrées comme appartenant à des modes distincts aient en fait été perçues de la même façon par les participants, surestimant ainsi le niveau d'ambiguïté des images. Par exemple, pour une silhouette similaire provenant du même modèle tridimensionnel, un sujet peut avoir répondu qu'il s'agit d'un avion alors qu'un autre a répondu un oiseau. Si les deux participants ont vu au même endroit les ailes, la tête (la cabine de pilotage) et la queue (l'arrière de l'appareil), ils ont perçu la même silhouette de la même façon, l'ont simplement interprétée différemment. Ce genre de situation peut s'être produit également avec certains quadrupèdes. Ayant une allure générale similaire, il est possible qu'un sujet répondant puma ait vu la même silhouette qu'un sujet ayant répondu chien. Il serait intéressant de s'assurer que les réponses considérées comme distinctes sont réellement des perceptions distinctes des silhouettes.

Finalement, un autre biais potentiel de nos résultats pourrait être le manque de redondance dans la présentation de nos stimuli. Dû au très grand nombre de stimuli produits (28 500) très peu de silhouettes identiques ont été perçues par plus d'un sujet. Les analyses ont été effectuées en regroupant, pour un même modèle, les stimuli ayant un niveau de complexité périmétrique semblable. Or, il serait intéressant de présenter exactement les mêmes stimuli à divers sujets, afin d'obtenir une vérification plus solide

de la constance des réponses inter-sujet. Ceci pourrait être effectué en faisant un plus grand nombre d'essais, en continuant à utiliser la méthodologie de cet article ou en sélectionnant les silhouettes les plus ambiguës, restreignant ainsi le nombre de stimuli pour ensuite les présenter à de nouveaux sujets.

Malgré les améliorations méthodologiques possibles, force est de constater que les silhouettes ambiguës semblent plus fréquentes dans notre environnement que ce qui est généralement supposé. Nous pouvons nous questionner sur la cause de leur nombre aussi limité dans la littérature. S'agit-il d'un manque d'intérêt pour le sujet en sciences de la perception? Serais-ce parce que la majorité des images auxquelles nous sommes exposés dans notre quotidien offrent généralement les indices facilitant la reconnaissance (tels que l'occlusion et l'ombrage), les silhouettes les plus ambiguës restant relativement peu commune? Ou serait-ce plutôt un perfectionnement de notre système visuel, nous rendant rigide aux différentes interprétations possibles d'une image, peut-être dans le but de ne pas y accorder notre attention et ainsi risquer d'augmenter inutilement au quotidien la confusion que celles-ci inspirent?

Références

- Attneave, F. (1957). Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology* **53**(4): 221-227.
- Attneave, F. and M. D. Arnoult (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin* **53**(6): 452-471.
- Bagby, J. W. (1957). A cross-cultural study of perceptual dominance in binocular rivalry. *Journal of Abnormal and Social Psychology* **54**: 331-334.
- Biederman, I. (1987). Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review* **94**(2): 115-147.
- Bisiach, E., R. Ricci, et al. (1999). Unilateral neglect and disambiguation of the Necker cube. *Brain* **122**: 131-140.
- Blanz, V., Tarr, M. J., & Bülthoff, H. H. (1999). What object attributes determine canonical views? *Perception*, **28**(5), 575-599.
- Boring, E. G. (1930). A New Ambiguous Figure. *The American Journal of Psychology* **42**: 444.
- Gregory, R. L. (2002) *Perceptions as hypotheses* in A. Noe, & E. Thompson, (Ed). *Vision and mind: Selected readings in the philosophy of perception*. (pp. 111-133). Cambridge, MA, US: MIT Press; US.
- Hegde, J. and D. Kersten (2010). A Link between Visual Disambiguation and Visual Memory. *Journal of Neuroscience* **30**(45): 15124-15133.

- Helmholtz, H. V. (1988). *Concerning the perceptions in general* in Benjamin, L. T. Jr. (Ed). *A history of psychology: Original sources and contemporary research* (pp. 97-100). New York, NY, England: McGraw-Hill Book Company.
- Hunt, J. M. and J. P. Guilford (1933). Fluctuation of an Ambiguous Figure in Dementia Præcox and in Manic-depressive Patients. *Journal of Abnormal and Social Psychology* **27**.
- Jastrow, J. (1899). The Mind's Eye. *Popular science monthly* **54**: 299-312.
- Kersten, D., P. Mamassian, et al. (2004). Object Perception as Bayesian Inference. *Annual Review of Psychology* **55**(1): 271-304.
- Kihlstrom, J. F. (2006, March). A new reversible figure and an old one. Paper presented at the annual meeting of the Society of Experimental Psychologists, La Jolla, Ca., March 25, 2006.
- Kramer, P. and S. Yantis (1997). Perceptual grouping in space and time: Evidence from the Ternus display. *Perception & Psychophysics* **59**: 87-99.
- Langer, M. S. and H. H. Bülthoff (2001). A prior for global convexity in local shape-from-shading. *Perception* **30**: 403-410.
- Mamassian, P. and R. Goutcher (2001). Prior knowledge on the illumination position. *Cognition* **81**: B1-B9.
- Mamassian, P. and R. Goutcher (2005). Temporal dynamics in bistable perception. *Journal of Vision* **5**(4)(7): 361-375.

- Mamassian, P. and M. S. Landy (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research* **38**: 2817-2832.
- Marr, D. (1982). Vision, A computational investigation into the human representation and processing of visual information. *WH San Francisco: Freeman and Company*.
- Munger, D. (2008). Casual Fridays: TK-421, why can't you spin that woman in reverse?.
- Palmer, S., Rosch, E., & Chase, P. (1980). Canonical perspective and the perception of objects. *Attention and Performance IX* (pp. 135-151). Hillsdale, NJ: Lawrence Erlbaum: J. Long, A. Baddeley.
- Pearson, J. and J. Brascamp (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Sciences* **12**(9): 334-341.
- Piccolino, M. and N. J. Wade (2006). Guest editorial essay: Flagging early examples of ambiguity I. *Perception & Psychophysics* **35**(7): 861-864.
- Rock, I. (1956). The orientation of forms on the retina and in the environment. *American Journal of Psychology* **69**: 513-528.
- Rock, I. (1974). The perception of disoriented figures. *Scientific American* **230**(1): 78-85.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382-352.
- Rubin, E. (1921). Visuell wahrgenommene Figuren. Copenhagen: Gyldendalske.

Shilane, P., Min, P., Kazhdan, M., & Funkhouser, T. (2004). [The Princeton Shape Benchmark](#) Shape Modeling International, Genova, Italy, June 2004

Sundareswara, R. and P. R. Schrater (2008). Perceptual multistability predicted by search model for Bayesian decisions. *Journal of Vision* **8**(5): 12-12.

Troje, N. F. and M. McAdam (2010). The viewing-from-above bias and the silhouette illusion. *i-Perception* **1**(3): 143-148.

Verfaillie, K., & Boutsen, L. (1995). A corpus of 714 full-color images of depth-rotated objects. *Perception & Psychophysics*, *57*(7), 925-961.

Watson, A. B. Perimetric Complexity of Binary Digital Images, *The Mathematica Journal*, 2012.

Figures

Figure 1: Quelques exemples des élastiques en forme d'animaux de Yumiko Ohashi et Masanori Haneda
<http://www.plus-d.com/arb/index.html>



Figure 2: La figure AWK de Kihlstrom (2006), évoquant à la fois une baleine et un kangourou.

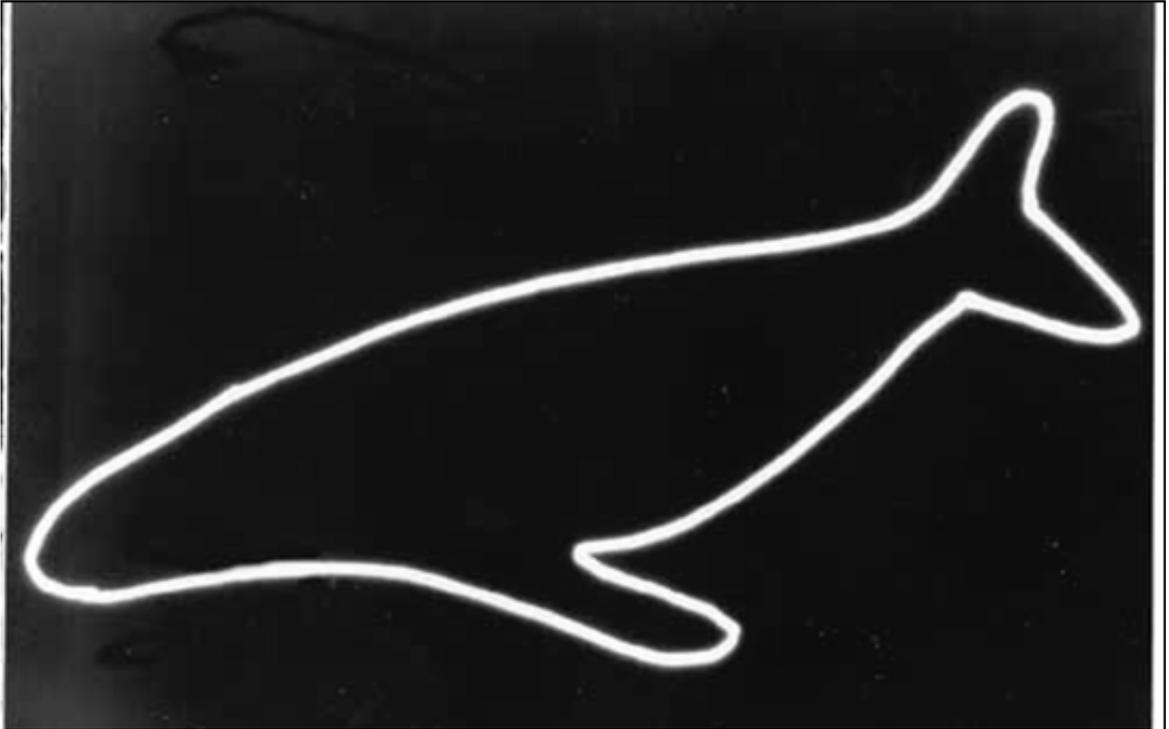
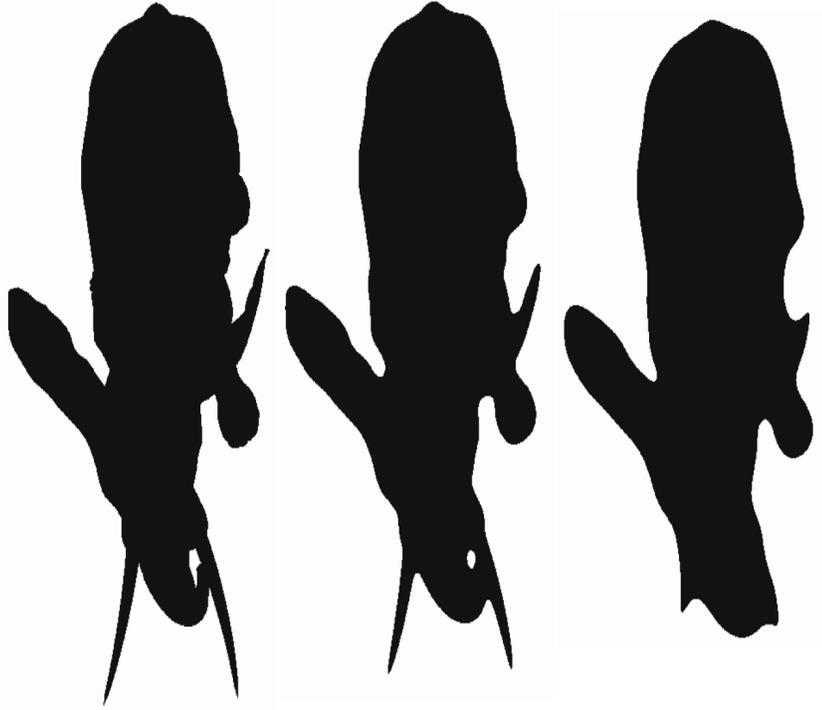


Figure 3 : Stimuli créés d'après le même modèle tridimensionnel d'éléphant. Les trois silhouettes ont des niveaux de complexité périmétrique de respectivement 59, 46, et 30. La figure a est présentée aux coordonnées sphériques 335.04 -367.49 51.93, la figure b figure à 147.38 -154.87 451.99.



(a)



(b)

Figure 4 : Graphique représentant le nombre de modes moyens pour quatre niveaux de complexité (Respectivement 3, 6, 9 et 14).

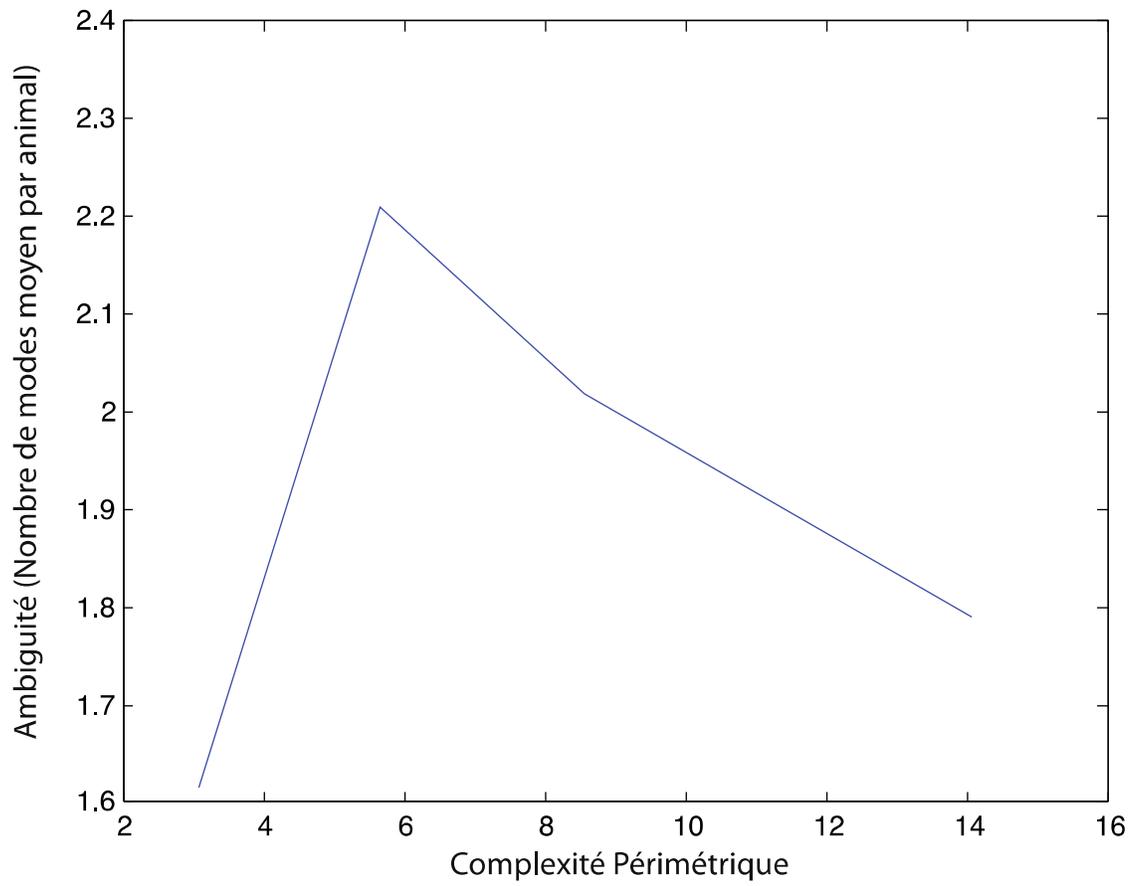
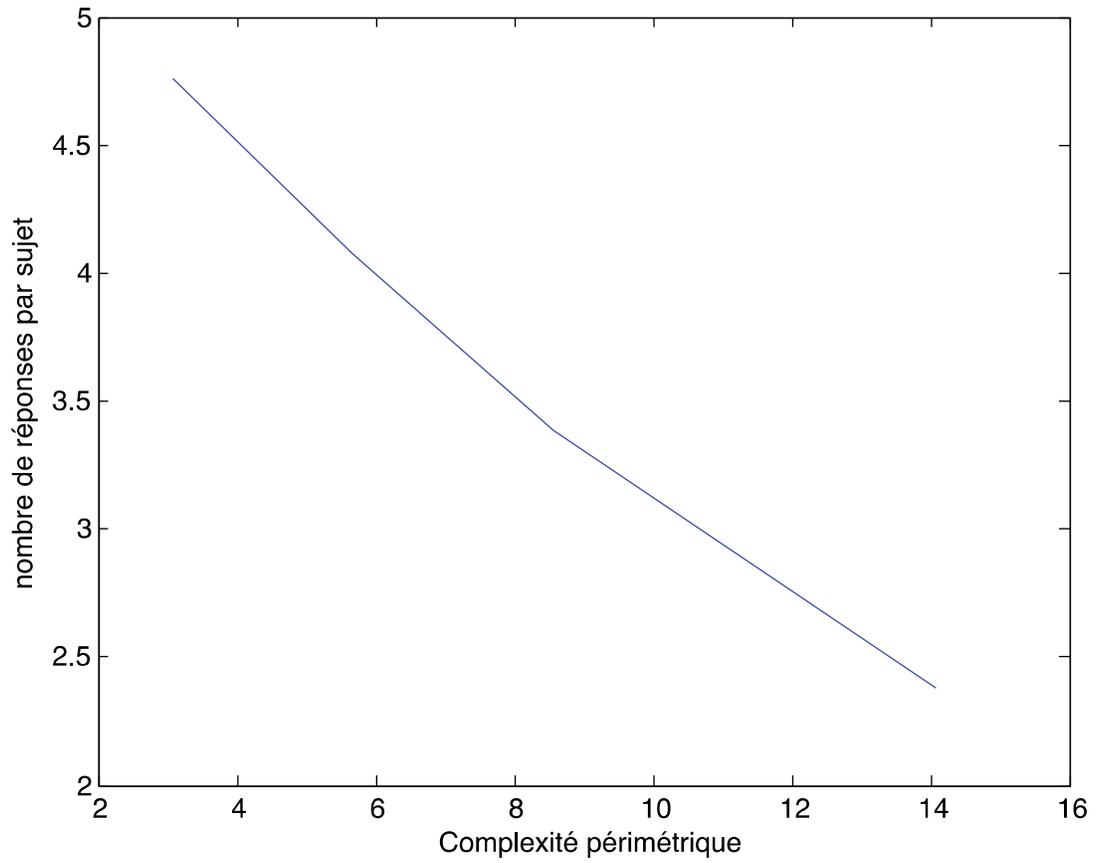


Figure 5: Nombre moyen de réponses par sujet selon le niveau de complexité périmétrique de la silhouette.



Appendice

Modèle statistique utilisé dans les analyses, élaboré au moyen de simulations de type Monte Carlo:

```
W = 541;           % Total number of categories used during testing
N = 29;           % nb of trials per 3d model per level of complexity(4)
ps = rand(1,W);   % change for lexical frequencies of animal categories
%ps = ones(1,W);

NN = 100000;      % infinity

ps = ps/sum(ps);  % made to sum to 1
ps = cumsum(ps);
ps2 = zeros(1,W+1);
ps2(2:end) = ps;
% simulation
%temp = ceil(rand(NN,N)*W);

temp = rand(NN,N);
temp2 = histc(temp', ps2);
result = temp2(1:end-1,:);

% some stats
% prob that exactly word will be repeated at least once in one draw by
chance :
% i.e., a sign of clarity
for criteria = 1:N,
    p(criteria) = sum(sum(result>criteria, 1)==1)/NN;
end
figure, plot(p)

% proba that more than one word will be repeated by chance : i.e., a sign
% of ambiguity
for criteria = 1:N,
    p(criteria) = sum(sum(result>criteria, 1)>1)/NN;
end
figure, plot(p)
```

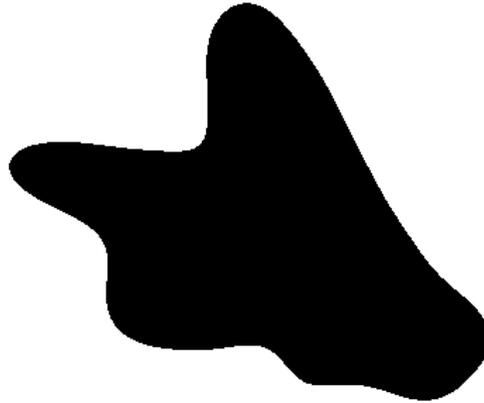

DISCUSSION GÉNÉRALE

Les apports de l'article de ce mémoire sont, dans l'état actuel de celui-ci, assez superficiels. Certains constats ont tout de même permis de soulever des questionnements ainsi que différentes pistes pour de futures études.

Premièrement, nous avons découvert que les silhouettes naturelles sont fréquemment ambiguës. Effectivement, pour tous les niveaux de complexité confondus, les silhouettes étaient en moyenne catégorisées dans 1.9079 modes. Nous avons également pu constater que même les silhouettes très complexes, celles pour lesquelles nous nous attendions à ce que la reconnaissance soit assez facile, présentent en moyenne plus d'un mode (1.8 mode pour les silhouettes ayant une complexité d'environ 14). Ceci dépasse nos attentes puisqu'il semblerait que même avec un niveau de détail élevé, les silhouettes sont des images ayant une grande propension à l'ambiguïté.

La figure 5 représente un modèle d'abeille ayant été lissé et ayant une complexité périmétrique de 20,03. La réponse donnée par un sujet pour cette image spécifique était « Tête de chat ». Divers autres niveaux de base ont été donnés pour ce modèle à ce niveau de lissage. Bonnet, fantôme, volant de badminton et guêpe sont des réponses obtenues une fois. Main pouce vers le haut et raie ont été répétés deux fois. Papillon et tête de chat répétés trois fois. Cœur, à quant à lui répété à quatre reprises.

Figure 5: Exemple de silhouette ambiguë, un modèle d'abeille avec une complexité périmétrique de 20.03



Ensuite, nous avons pu déterminer que le niveau de lissage, ou de dégradation, d'une silhouette influe sur l'ambiguïté de celle-ci. En fait, à mesure que le lissage augmente, l'ambiguïté fait de même pour ensuite chuter après avoir plafonné à un Sigma d'environ 35. Ceci confirme notre hypothèse selon laquelle au-delà d'un seuil de lissage, les stimuli deviennent abstraits et suggèrent donc un grand nombre de perceptions. L'entropie des silhouettes augmente graduellement jusqu'à ce que celles-ci se prêtent à toutes les interprétations. Elles deviennent semblables à du bruit ou aux taches d'un test de Rorschach (1921). Il est alors possible que les réponses des sujets ne reviennent alors plus assez fréquemment pour être considérées comme un mode perceptuel.

Le même phénomène est constaté lorsque nous nous penchons sur le niveau de complexité périmétrique d'une silhouette plutôt que son niveau de lissage. Nous remarquons alors que les images les moins complexes ont un niveau d'ambiguïté moins élevé que les plus complexes, appuyant encore une fois l'hypothèse selon laquelle les

images très simples inspirent trop d'interprétations distinctes et sont même moins ambiguës que des silhouettes très complexes.

Une complexité périmétrique d'environ 6 semble être le niveau optimal d'ambiguïté et présente en moyenne 2.2 modes.

1 Imprévus rencontrés lors de cette recherche.

Deux éléments majeurs ont ralenti le fonctionnement de cette recherche et ont contribué à en limiter les résultats.

1.1 Premièrement, une coquille dans le programme servant à produire les stimuli a retardé l'expérience de plusieurs mois. En effet, à chaque silhouette l'horloge de l'ordinateur fournissait l'origine d'un générateur de nombre pseudo-aléatoire, ce qui compromet la randomisation assignée au point de vue et au niveau de lissage des images. Nous avons constaté ce problème seulement après avoir testé 16 sujets pilotes (pour un total de 391 essais) et 83 participants (3150 essais). Leurs résultats étant inutilisables à cause de cette erreur, nous avons dû régénérer des stimuli en corrigeant le code, recruter et tester de nouveaux participants. Notre procédure d'origine enregistrait des réponses plus riches. Nous demandions aux participants de dessiner les différentes parties de l'image qu'ils percevaient sur la silhouette présentée. Par exemple, s'ils avaient perçu un oiseau, ils devaient indiquer où ils voyaient les ailes, la tête et les pattes de l'animal en utilisant différentes couleurs et une légende. Cette étape rendait la procédure beaucoup plus longue. Nous avons décidé, faute de temps, de la rejeter lors des expériences subséquentes.

1.2 Le second imprévu rencontré est le fait qu'un volet de l'expérience, l'impact du point de vue sur l'ambiguïté, dut être abandonné lors des analyses statistiques. En effet, cette recherche avait été conçue pour tenir compte à la fois de l'effet du point de vue et de celui du lissage sur l'ambiguïté d'une silhouette. Des analyses exploratoires ont été effectuées afin de déterminer si, tous niveaux de lissage confondus, certains points de vue entraînaient une perception plus ambiguë. Cependant, le nombre d'essais limité que nous avons eu l'occasion d'effectuer a rendu impossible l'exploration de la variable *point de vue*. Nous considérons l'image du modèle de face, sans aucune élévation, comme étant le point de vue d'origine. De celui-ci, un grand nombre d'angles de présentation ont été produits aléatoirement. L'angle d'élévation et de longitude étaient variables, le rayon constant, créant ainsi une sphère de points de vue entourant le modèle. Cette sphère a ensuite été divisée en quatre groupes de points de vue adjacents, d'après le point de vue d'origine, couvrant chacun le quart d'une sphère (les énantiomorphes, ou images miroir, d'une silhouette donnant la même image, les opposées étaient considérées identiques). Les divisions entre chacun des groupes étaient arbitraires, mais aucune solution ne nous est venue afin de contourner cette notion de hasard. Nous n'avons pas pu déceler de différence significative selon les différents angles de présentations.

Ce fait est surprenant compte tenu de la grande quantité d'exemples recensés dans la littérature où le point de vue influence la facilité ou rapidité de reconnaissance d'un objet comme le *viewing from above bias* (Mamassian & Landy, 1998) et le point de vue canonique (Palmer, Rosch, & Chase, 1980; Blanz, Tarr, & Bülhoff, 1999; Verfaillie & Boutsen, 1995).

2 Limites de l'étude et améliorations possibles.

Tel que mentionné plus haut, la recherche présentée dans l'article de ce mémoire compte certains handicaps. Les lacunes majeures sont ici relevées, accompagnées de suggestions afin de pallier à celles-ci lors d'une potentielle réplique de cette étude.

2.1 Présence d'un thème redondant dans le choix des stimuli.

Un premier élément ayant pu influencer la validité des données est le fait que tous les modèles originaux utilisés pour produire les stimuli étaient des animaux. Pour certains stimuli peu dégradés, le modèle de base était aisément identifiable. Bien qu'il ait été spécifié aux sujets qu'ils ne devaient s'imposer aucune restriction quant à ce qu'ils pouvaient percevoir dans les silhouettes, il est possible qu'ils aient constaté par eux-mêmes que des représentations d'animaux revenaient fréquemment. Si tel est le cas, il est probable qu'ils aient ensuite restreint leur interprétation à cette catégorie faunique. Cette supposition semble appuyée par le fait que 49.07% des réponses données par les participants désignaient des animaux.

2.1.1 Améliorations possibles dans le choix des stimuli.

Il serait intéressant lors d'une étude subséquente de ne pas limiter le choix des modèles à un thème particulier, mais plutôt d'essayer de couvrir un spectre d'objets aussi large que possible, permettant ainsi d'éviter ce biais thématique.

Afin de s'assurer malgré tout que les sujets n'ont pas perçu de constante dans les modèles présentés, il serait également possible de demander directement aux sujets, suite au *testing*, s'ils ont reconnu un motif quelconque dans le choix des stimuli et s'ils s'y limitaient pour le choix de leurs réponses.

2.2 Manque de redondance dans les stimuli présentés

Une autre limite pouvant avoir influé sur les données est la très grande quantité de stimuli produits ainsi que le choix que nous avons fait de les sélectionner aléatoirement lors de la présentation au sujet. En fait, 500 stimuli avec différents points de vue et niveaux de lissage ont été produits pour chacun des 57 modèles tridimensionnels d'animaux. De cette banque totale de 28 500 stimuli produits, seulement 6 525 ont été présentés. Le rationnel derrière le choix de créer un très grand nombre de stimuli était de couvrir un spectre de niveaux de lissage et de points de vue aussi large que possible. Cependant, cela implique également qu'un même stimulus a très rarement été présenté à plus d'un sujet. Ce manque de redondance dans les stimuli présentés nous a empêché de vérifier la constance des réponses inter-sujets pour une même silhouette. Nous avons donc dû regrouper des silhouettes similaires (même modèle 3d, points de vue confondus, niveau de lissage semblable), mais non pas identiques afin de déterminer leur ambiguïté.

2.2.1 Effectuer un plus grand nombre d'essais

Une possibilité pour contrer cette limite, serait de recruter de nouveaux participants et d'augmenter le nombre d'essais total. En effet, la grande quantité de stimuli produits n'est problématique que dans la mesure où nous retrouvons une grande disparité entre la quantité de stimuli produits et les stimuli présentés. À mesure que la quantité d'essais augmente, la répétition de stimuli identiques présentés à plusieurs sujets sera plus fréquente.

2.2.2 Réutilisation de certains stimuli.

Une autre solution possible serait de limiter la quantité de stimuli en réutilisant certains d'entre eux. Il faudrait sélectionner, parmi nos stimuli, les silhouettes qui semblent particulièrement propices à créer de l'ambiguïté et tester de nouveaux participants avec cette banque de stimuli plus restreinte.

Dans le même ordre d'idées, il serait intéressant de vérifier la stabilité des réponses chez un même sujet. Dans l'expérience centrale de ce mémoire, la programmation proscrivait la réutilisation d'un même stimulus pour un même sujet. C'est-à-dire qu'un sujet ne s'est jamais fait présenter deux fois la même silhouette. Il serait intéressant, lors d'une réplication de l'expérience de permettre la présentation à plus d'une reprise d'un stimulus à un même sujet, permettant ainsi de vérifier si l'objet qu'il y perçoit est toujours le même. Lors de ces présentations répétées, une rotation pourrait être imposée aux silhouettes afin que la reconnaissance ne soit pas trop aisée. Il est possible que cette rotation encourage de nouvelles perceptions pour un même stimulus.

2.2.3 Amélioration possible lors de la production des stimuli

Une seconde option permettant la présentation d'un même stimulus à plusieurs participants serait de recréer les stimuli selon certains points de vue et niveau de lissage prédéfinis. Par exemple, chaque modèle tridimensionnel pourrait être présenté sous six points de vue, comportant chacun six niveaux de lissage, ce qui donnerait une banque de 2 052 stimuli, nettement plus restreinte que celle de 28 500 utilisés dans la présente étude. Le fait de réduire la quantité de points de vue et de niveaux de lissage présentés limiterait probablement la validité écologique des résultats, mais aurait l'avantage de prévenir un groupement non justifié. Par exemple, les groupements selon le point de vue que nous avons effectués lors des analyses étaient arbitraires. Ainsi, deux points de vue

pourtant très proches pouvaient être traités comme appartenant à deux groupes différents.

3 Surestimation possible du niveau d'ambiguïté

3.1 Réponses distinctes, perception identique.

Un élément pouvant potentiellement avoir mené à une surestimation de l'ambiguïté est que certaines réponses considérées comme distinctes représentent en fait la même perception. Par exemple, face à la même image d'un oiseau vu du dessous, les ailes déployées, il est possible qu'un sujet réponde oiseau et un autre avion, alors que leur interprétation de la silhouette était en fait la même, tous deux ayant vu l'avant (la tête), les ailes et la queue au même endroit. Ce problème peut potentiellement être rencontré pour différents animaux à l'allure similaire, notamment des quadrupèdes. La même silhouette interprétée de la même façon pourrait contenir des réponses distinctes telles que chien et léopard. Si tel est le cas, il est possible que certaines réponses considérées comme ambiguës lors de l'analyse ne l'étaient pas. Un moyen de contourner ce problème serait, lors de chaque essai, de demander aux participants de nommer les différentes parties de l'objet perçu et de leur demander d'assigner à chacune une couleur correspondante. Cette légende pourrait s'ajouter au lexique déjà utilisé afin de déterminer quelles réponses sont équivalentes. Les cas où la réponse est distincte, mais les coloriages identiques, seraient considérés comme des réponses similaires. Tel qu'expliqué dans la section 1.1, nous avons originellement programmé cette étape dans notre procédure, mais avons dû l'abandonner afin d'effectuer un maximum d'essais dans un temps restreint.

3.2 Surestimation de l'ambiguïté due à la diversité des sujets

Nous avons envisagé la possibilité que certains sujets répondent constamment la même catégorie de base à chaque fois qu'un modèle particulier leur est présenté. Par exemple, nous pourrions retrouver deux ou trois catégories de base pour un même modèle, mais chacune d'elles ne serait due qu'aux réponses d'un sujet en particulier. Ces images ne seraient alors pas ambiguës au sens où on l'entend généralement, c'est-à-dire qu'une même personne peut y percevoir deux choses. Il s'agit en fait d'images dans lesquelles différentes personnes perçoivent différentes choses. Il est donc important que ces images ne soient pas traitées comme ambiguës puisqu'elles surestimeraient l'ambiguïté réelle des silhouettes. Afin de vérifier cette hypothèse, nous avons calculé le nombre de catégories de base répondues en moyenne par sujet, par modèle, selon le niveau de complexité périmétrique. Nous constatons qu'il n'y a pas de sujets ayant donnés une très grande quantité de réponses à certains niveaux de complexité, ce qui indique que l'effet est bien intra-sujet.

3.3 Surestimation de l'ambiguïté due à la diversité des points de vue

Dans le même ordre d'idées, il est possible que la diversité des points de vue présentés pour un même modèle ait surestimé l'ambiguïté des certains d'entre eux. Comme nous l'avons expliqué, un grand nombre de points de vue étaient présentés pour chacun des modèles. Il est envisageable, par exemple, que pour un même modèle au même niveau de lissage, un point de vue soit interprété par tous les sujets comme ayant l'air d'un chat alors qu'un autre est exclusivement interprété comme un requin. En soit, aucune de ces images n'est ambiguë; elles n'ont chacune qu'un mode, mais lorsqu'elles

sont mises en commun dans le même groupement, elles donneront l'impression que ce modèle, à ce niveau de lissage est ambigu, qu'il contient plusieurs modes.

Un moyen d'éviter ce biais serait de contrôler le point de vue lorsque nous moyennons les modes pour chaque modèle et niveau de complexité. Par exemple, chaque niveau de complexité pourrait être subdivisé en quatre points de vue. La moyenne du nombre de modes pour ce niveau de complexité serait ensuite faite d'après la moyenne du nombre de modes de chacun des groupes de points de vue. Nous n'avons cependant pas assez d'essais actuellement afin d'effectuer une telle analyse. De plus, les groupements selon le point de vue seraient encore une fois faits de façon arbitraire. Des points de vue relativement similaires pourraient être assignés à des groupes différents.

Une seconde option serait, comme nous l'avons expliqué au point 2.2.3, de recréer des stimuli en se limitant à six points de vue. Tel qu'expliqué alors, la validité écologique des résultats serait limitée, mais chacun des groupes de point de vue serait plus représentatif.

Références

- Attneave, F. (1957). Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology* **53**(4): 221-227.
- Attneave, F. and M. D. Arnoult (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin* **53**(6): 452-471.
- Bagby, J. W. (1957). A cross-cultural study of perceptual dominance in binocular rivalry. *Journal of Abnormal and Social Psychology* **54**: 331-334.
- Biederman, I. (1987). Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review* **94**(2): 115-147.
- Bisiach, E., R. Ricci, et al. (1999). Unilateral neglect and disambiguation of the Necker cube. *Brain* **122**: 131-140.
- Blanz, V., Tarr, M. J., & Bühlhoff, H. H. (1999). What object attributes determine canonical views? *Perception*, **28**(5), 575-599.
- Boring, E. G. (1930). A New Ambiguous Figure. *The American Journal of Psychology* **42**: 444.
- Gregory, R. L. (2002) *Perceptions as hypotheses* in A. Noe, & E. Thompson, (Ed). *Vision and mind: Selected readings in the philosophy of perception*. (pp. 111-133). Cambridge, MA, US: MIT Press; US.
- Hegde, J. and D. Kersten (2010). A Link between Visual Disambiguation and Visual Memory. *Journal of Neuroscience* **30**(45): 15124-15133.
- Helmholtz, H. V. (1988). *Concerning the perceptions in general* in Benjamin, L. T. Jr. (Ed). *A history of psychology: Original sources and contemporary research*.

(pp. 97-100). *New York, NY, England: McGraw-Hill Book Company.*

Hunt, J. M. and J. P. Guilford (1933). Fluctuation of an Ambiguous Figure in Dementia Præcox and in Manic-depressive Patients. *Journal of Abnormal and Social Psychology* **27**.

Jastrow, J. (1899). The Mind's Eye. *Popular science monthly* **54**: 299-312.

Kersten, D., P. Mamassian, et al. (2004). Object Perception as Bayesian Inference. *Annual Review of Psychology* **55**(1): 271-304.

Kihlstrom, J. F. (2006, March). A new reversible figure and an old one. Paper presented at the annual meeting of the Society of Experimental Psychologists, La Jolla, Ca., March 25, 2006.

Kramer, P. and S. Yantis (1997). Perceptual grouping in space and time: Evidence from the Ternus display. *Perception & Psychophysics* **59**: 87-99.

Langer, M. S. and H. H. Bülthoff (2001). A prior for global convexity in local shape-from-shading. *Perception* **30**: 403-410.

Mamassian, P. and R. Goutcher (2001). Prior knowledge on the illumination position. *Cognition* **81**: B1-B9.

Mamassian, P. and R. Goutcher (2005). Temporal dynamics in bistable perception. *Journal of Vision* **5**(4)(7): 361-375.

Mamassian, P. and M. S. Landy (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research* **38**: 2817-2832.

- Marr, D. (1982). *Vision, A computational investigation into the human representation and processing of visual information. WH San Francisco: Freeman and Company.*
- Munger, D. (2008). Casual Fridays: TK-421, why can't you spin that woman in reverse?.
- Palmer, S., Rosch, E., & Chase, P. (1980). Canonical perspective and the perception of objects. *Attention and Performance IX* (pp. 135-151). Hillsdale, NJ: Lawrence Erlbaum: J. Long, A. Baddeley.
- Pearson, J. and J. Brascamp (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Sciences* **12**(9): 334-341.
- Piccolino, M. and N. J. Wade (2006). Guest editorial essay: Flagging early examples of ambiguity I. *Perception & Psychophysics* **35**(7): 861-864.
- Rock, I. (1956). The orientation of forms on the retina and in the environment. *American Journal of Psychology* **69**: 513-528.
- Rock, I. (1974). The perception of disoriented figures. *Scientific American* **230**(1): 78-85.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382-352.
- Rubin, E. (1921). *Visuell wahrgenommene Figuren*. Copenhagen: Gyldendalske.
- Shilane, P., Min, P., Kazhdan, M., & Funkhouser, T. (2004). [The Princeton Shape Benchmark](#) Shape Modeling International, Genova, Italy, June 2004
- Sundareswara, R. and P. R. Schrater (2008). Perceptual multistability predicted by search model for Bayesian decisions. *Journal of Vision* **8**(5): 12-12.

Troje, N. F. and M. McAdam (2010). The viewing-from-above bias and the silhouette illusion. *i-Perception* **1**(3): 143-148.

Verfaillie, K., & Boutsen, L. (1995). A corpus of 714 full-color images of depth-rotated objects. *Perception & Psychophysics*, *57*(7), 925-961.

Watson, A. B. Perimetric Complexity of Binary Digital Images, *The Mathematica Journal*, 2012.