

Université de Montréal

**L'utilisation de l'information visuelle en
reconnaissance d'expressions faciales d'émotion**

par

Caroline Blais

Département de Psychologie
Faculté des Arts et des Sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Ph.D
en Psychologie
option Sciences cognitives et neuropsychologie

Septembre 2012

© Caroline Blais, 2012

Résumé

L'aptitude à reconnaître les expressions faciales des autres est cruciale au succès des interactions sociales. L'information visuelle nécessaire à la catégorisation des expressions faciales d'émotions de base présentées de manière statique est relativement bien connue. Toutefois, l'information utilisée pour discriminer toutes les expressions faciales de base entre elle demeure encore peu connue, et ce autant pour les expressions statiques que dynamiques. Plusieurs chercheurs assument que la région des yeux est particulièrement importante pour arriver à « lire » les émotions des autres. Le premier article de cette thèse vise à caractériser l'information utilisée par le système visuel pour discriminer toutes les expressions faciales de base entre elles, et à vérifier l'hypothèse selon laquelle la région des yeux est cruciale pour cette tâche. La méthode des Bulles (Gosselin & Schyns, 2001) est utilisée avec des expressions faciales statiques (Exp. 1) et dynamiques (Exp. 2) afin de trouver quelles régions faciales sont utilisées (Exps. 1 et 2), ainsi que l'ordre temporel dans lequel elles sont utilisées (Exp. 2). Les résultats indiquent que, contrairement à la croyance susmentionnée, la région de la bouche est significativement plus utile que la région des yeux pour discriminer les expressions faciales de base. Malgré ce rôle prépondérant de la bouche, c'est toute de même la région des yeux qui est sous-utilisée chez plusieurs populations cliniques souffrant de difficultés à reconnaître les expressions faciales. Cette observation pourrait suggérer que l'utilisation de la région des yeux varie en fonction de l'habileté

pour cette tâche. Le deuxième article de cette thèse vise donc à vérifier comment les différences individuelles en reconnaissance d'expressions faciales sont reliées aux stratégies d'extraction de l'information visuelle pour cette tâche. Les résultats révèlent une corrélation positive entre l'utilisation de la région de la bouche et l'habileté, suggérant la présence de différences qualitatives entre la stratégie des patients et celle des normaux. De plus, une corrélation positive est retrouvée entre l'utilisation de l'œil gauche et l'habileté des participants, mais aucune corrélation n'est retrouvée entre l'utilisation de l'œil droit et l'habileté. Ces résultats indiquent que la stratégie des meilleurs participants ne se distingue pas de celle des moins bons participants simplement par une meilleure utilisation de l'information disponible dans le stimulus : des différences qualitatives semblent exister même au sein des stratégies des participants normaux.

Mots-clés : Perception visuelle, Stratégies d'extraction de l'information, Expressions faciales, Émotions de base, Différences individuelles, Bulles, Mouvements oculaires.

Abstract

The ability to recognize facial expressions is crucial for the success of social communication. The information used by the visual system to categorize static basic facial expressions is now relatively well known. However, the visual information used to discriminate the basic facial expressions from one another is still unknown, and this is true for both static and dynamic facial expressions. Many believe that the eye region of a facial expression is particularly important when it comes to reading others' emotions. The aim of the first article of this thesis is to determine which information is used by the visual system in order to discriminate between the basic facial expressions and to verify the validity of the hypothesis that the eye region is crucial for this task. The Bubbles method (Gosselin & Schyns, 2001) is used with static (Exp. 1) and dynamic (Exp. 2) facial expressions in order to determine which facial areas are used for the task (Exp. 1) and in which temporal order these facial areas are used (Exp. 2). The results show that, in contrast with the aforementioned belief, the mouth region is significantly more useful than the eye region when discriminating between the basic facial expressions. Despite this preponderant role of the mouth, it is the eye area—not the mouth area—that is underutilized by many clinical populations suffering from difficulties in recognizing facial expressions. This observation could suggest that the utilization of the eye area varies as a function of the ability to recognize facial expressions. The aim of the second article in

this thesis is thus to verify how individual differences in the ability to recognize facial expressions relate to the visual information extraction strategies used for this task. The results show a positive correlation between the ability of the participants and the utilization of the mouth region, suggesting the existence of qualitative differences between the strategy of clinical patients and of normal participants. A positive correlation is also found between the ability of the participants and the utilization of the left eye area, but no correlation is found between the ability and the utilization of the right eye area. These results suggest that the difference between the strategies of the best and the worst participants is not only that the best ones use the information available in the stimulus more efficiently: rather, qualitative differences in the visual information extraction strategies may exist even within the normal population.

Keywords : Visual perception, Information extraction strategies, Facial expressions, Basic emotions, Individual differences, Bubbles, Eye movements.

Table des matières

Résumé	ii
Abstract	iv
Table des matières	vi
Liste des figures	ix
Liste des abbréviations	xi
Remerciements	xii
Chapitre 1: Introduction générale	0
1.1 Préambule	1
1.2 L'expression faciale : sa transmission.....	2
<i>1.2.1 Les expressions faciales représentent-elles bien l'émotion vécue?</i>	<i>3</i>
<i>1.2.2 L'universalité des expressions faciales d'émotions.....</i>	<i>6</i>
<i>1.2.3 Les mouvements faciaux impliqués dans les expressions faciales ...</i>	<i>10</i>
1.3 L'expression faciale : son décodage.....	12
<i>1.3.1 Information utile pour reconnaître les expressions faciales : le haut vs. le bas de visage</i>	<i>12</i>
<i>1.3.2 Information utile pour reconnaître les expressions faciales : les fréquences spatiales.....</i>	<i>16</i>
<i>1.3.3 Information utile pour reconnaître les expressions faciales : la méthode des Bulles</i>	<i>19</i>
<i>1.3.4 La région des yeux est-elle spéciale?</i>	<i>25</i>

1.3.5 <i>Les expressions faciales dynamiques</i>	30
1.4 Résumé et objectifs de la présente thèse	33
Chapitre 2: Article 1	37
Chapitre 3: Article 2	85
Chapitre 4: Discussion générale	112
4.1 Retour sur les objectifs	113
4.2 Objectif 1 : Clarifier la stratégie visuelle utilisée pour discriminer les expressions faciales de base entre elles	114
4.2.1. <i>Stratégie des humains : Montre-moi ta bouche et je te dirai ce que tu ressens!</i>	114
4.2.2. <i>Information contenue dans les différentes régions du visage : Observateur idéal et analyse de mouvement.</i>	116
4.2.3. <i>Information utilisée vs. Information disponible</i>	117
4.2.4. <i>Mouvements oculaires et utilisation de l'information visuelle</i>	120
4.2.5 <i>Contribution de la région des yeux au percept de l'expression faciale d'émotion.</i>	121
4.3 Objectif 2 : Vérifier l'impact de différences individuelles au niveau de l'habileté à reconnaître les expressions faciales sur les stratégies visuelles pour extraire l'information	126
4.3.1 <i>Modulation quantitative vs. qualitative des stratégies visuelles</i>	127
4.3.3 <i>Différences hémisphériques en reconnaissance d'expressions faciales</i>	130

4.4 Limites et perspectives futures	136
4.4.1 <i>Les émotions complexes</i>	136
4.4.2 <i>L'impact de divers facteurs cognitifs, affectifs et sociaux sur les stratégies visuelles</i>	137
4.4.3 <i>Expressions faciales posées vs. spontanées</i>	138
4.4.4 <i>Mouvements oculaires et expressions faciales</i>	139
4.4.5 <i>Conclusion</i>	147
Références	148

Liste des figures

Chapitre 1- Introduction générale

Figure 1. Exemple d'une partie hypothétique pour le système visuel.....	14
Figure 2. Exemple de filtrage fréquentiel.....	17
Figure 3. Exemple de création d'un stimulus avec la méthode des Bulles.	22

Chapitre 2- Article 1

Figure 1. Illustration of the creation of a bubbled stimulus.....	76
Figure 2. Classification plane and classification volume.....	78
Figure 3. ROI analysis without the time dimension.....	80
Figure 4. ROI analysis with the time dimension.....	81
Figure 5. Classification plane obtained with the data of Smith et al. (2005).....	82
Figure 6. Classification plane obtained by the ideal observer.....	83
Figure 7. Relative variance of the amplitude of motion in different areas of a face.....	84

Chapitre 3- Article 2

Figure 1. Illustration of the creation of a bubbled stimulus.....	107
-------------------------------------------------------------------	-----

Figure 2. Thresholded average classification images obtained in Exps. 1 and 2.....	109
Figure 3. Thresholded individual differences classification images obtained in Exps. 1 and 2.....	110
Figure 4. ROI analysis.....	111

Chapitre 4- Discussion générale

Figure 1. Analyse ROI à travers le temps.....	134
Figure 2. Cartes de la proportion moyenne des fixations oculaires.....	142
Figure 3. Cartes des corrélations entre les fixations oculaires et l'habileté des participants.....	145

Liste des abréviations

- CI : Classification image
- CV : Classification volume
- FACS : Facial Action Coding System
- FFA : Fusiform Face Area
- FWHM : Full width half maximum
- ROI : Region-of-interest
- SD : Standard deviation
- STS : Sulcus temporal supérieur
- V1 : Cortex visuel primaire

Remerciements

Merci à mes directeurs, Frédéric Gosselin et Martin Arguin, pour m'avoir si bien guidée tout au long de ce périple. Merci M. Arguin, pour m'avoir ouvert la porte de votre laboratoire dès mes premiers balbutiements en recherche; pour le support que vous m'avez fourni à chaque fois que j'en avais besoin; et pour vos interventions rigoureuses et formatrices. Chacune de nos rencontres s'est révélée stimulante et motivante. Merci Frédéric, pour ton incroyable disponibilité; pour la formation exceptionnelle que tu m'as donnée; pour m'avoir appris à toujours douter et à remettre tout en question; pour ta passion contagieuse; pour la confiance que tu m'as accordée en m'impliquant dans tes nombreux projets; et pour les superbes opportunités scientifiques que tu m'as offertes en m'envoyant dans les laboratoires de tes collaborateurs.

Merci à tous mes compagnons de laboratoire. Nicolas, l'absurdité rafraichissante de nos discussions fera partie intégrante de mes souvenirs de thésarde. Catherine, jamais je n'ai autant apprécié faire des « bootstraps » qu'en ta compagnie! Karine et Catherine, merci pour les nombreux services que vous m'avez rendus tout au long de cette thèse. Verena, merci pour les stimulantes discussions que nous avons eues. Et enfin, merci Cynthia, pour les nombreuses heures que nous avons partagées à fouiller des données; pour la justesse de tes interventions; et pour ta capacité à me ramener à la source du problème lorsque

je m'égarais dans les boucles et les conditions booléennes. Merci, surtout, pour l'amitié que tu m'as offerte.

Merci à mes nouvelles collègues de l'UQO, en particulier à Geneviève et Hélène, pour vos conseils précieux durant les derniers miles de cette thèse. Débuter ma carrière à vos côtés est un réel plaisir et c'est avec enthousiasme que j'entrevois nos projets scientifiques futurs ensemble.

Merci à ma famille et à mes amis. Merci à mes parents, pour votre support inconditionnel et ce, dans toutes les sphères de ma vie. Merci à mes grands-parents, pour avoir nourri mon appétit intellectuel dès mon plus jeune âge en me fournissant roman, atlas, dictionnaires, etc. Merci à mes beaux-parents, pour votre présence et votre aide constante. Merci Geneviève, pour l'incroyable amie que tu es; pour ta facilité à reprendre la discussion où nous l'avions laissée, même lorsque plusieurs mois se sont écoulés; et pour être toujours présente dans ma vie, peu importe la distance qui nous sépare.

Finalement, merci à mes deux amours. Dan, pour nos innombrables discussions passionnantes; pour ton savoureux sens de la contradiction; pour ton enthousiasme et ton incroyable capacité à « pondre » des idées, toutes plus passionnantes les une que les autres; et surtout, pour me faire rire au quotidien, pour m'endurer dans mes pires moments, et pour l'estime (d'ailleurs

légèrement exagéré!) que tu as pour moi. Ma douce Anaïs, merci pour tes sourires coquins, pour ta joie de vivre et pour me rappeler ce qui est important dans la vie.

Chapitre 1

Introduction générale

1.1 Préambule

Imaginez que vous discutez tranquillement avec un ami sur la terrasse d'un restaurant et que soudainement, son visage se déforme en une expression de peur intense. Une chaîne de réactions se déclenche automatiquement : votre rythme cardiaque accélère, et avant même de vérifier l'origine de la réaction de votre ami, vous vous déplacez rapidement en direction inverse de l'endroit fixé par celui-ci. Vous vous retournez finalement, toujours en état d'alerte, pour constater qu'un vélo vient d'être renversé par une voiture à quelques mètres de vous.

Dans cette situation hypothétique, la reconnaissance de l'expression de peur dans le visage de votre ami vous aura permis d'enclencher les réactions physiologiques nécessaires pour vous protéger d'une menace inconnue. La reconnaissance des expressions faciales d'émotions est cruciale non seulement pour la survie de l'organisme, mais également pour l'adoption de comportements adéquats dans nos relations sociales. Par exemple, vous ne débuterez probablement pas une discussion avec votre patron de la même façon si celui-ci vous accueille dans son bureau avec une expression de colère ou s'il affiche au contraire un sourire bienveillant. L'expression faciale joue donc un rôle de communication entre deux ou plusieurs organismes : elle permet à un individu d'exprimer l'état émotionnel dans lequel il se trouve, et elle permet à

ceux qui sont en contact avec cet individu d'inférer son état émotionnel et d'ajuster leur comportement en conséquence. De ce fait, l'expression faciale peut être examinée sous deux angles : (1) la transmission de l'information émotionnelle (i.e., le signal contenu dans un visage exprimant une émotion), et (2) le décodage de l'information émotionnelle (i.e., les stratégies visuelles utilisées par un observateur pour reconnaître l'émotion exprimée par un autre individu).

1.2 L'expression faciale : sa transmission.

Afin que l'expression faciale puisse assumer son rôle communicatif de manière efficace, une certaine systématisme doit nécessairement exister dans l'apparence visuelle de l'expression affichée pour une émotion donnée. Les recherches menées au cours du vingtième siècle sur la reconnaissance des expressions faciales suggèrent que les expressions faciales constituent non seulement un signal suffisamment systématique pour être identifiées hors contexte par des observateurs externes (Coleman, 1949; Ekman & Bressler, 1964; Ekman & Rose, 1965; Howell & Jorgenson, 1970; Lanzetta & Kleck, 1970), mais que le signal transmis est en partie partagé par les individus provenant de différentes cultures (Ekman, 1972; Ekman & Friesen, 1971; Elfenbein & Ambady, 2002; voir cependant Naab & Russell, 2007; Russell, 1994).

1.2.1 Les expressions faciales représentent-elles bien l'émotion vécue?

Une question légitime à se poser lorsque l'on étudie la reconnaissance des expressions faciales relève de la concordance entre l'apparence d'une expression faciale et l'émotion interne vécue par l'émetteur. En effet, si l'expression faciale ne constitue pas un reflet juste de l'émotion vécue, l'intérêt scientifique qu'elle suscite serait plus ou moins justifié. Dans un monde idéal, le lien entre l'expression faciale et l'émotion vécue pourrait être étudié en ayant accès aux expressions faciales effectuées de manière spontanée par des individus dans des situations écologiques et en sachant l'émotion vécue par l'émetteur au moment précis où l'expression faciale fut capturée (e.g., photographiée ou filmée). Ces expressions faciales spontanées pourraient ensuite être présentées à des observateurs naïfs afin de vérifier s'ils identifient bien l'émotion ressentie par l'émetteur. Il est malheureusement impossible d'atteindre cet idéal puisque l'accès à l'émotion vécue est au mieux indirect (e.g., par un rapport verbal de l'émetteur, ou par inférence à partir de la situation dans laquelle l'émotion a été générée). De plus, il est très difficile de capturer une expression faciale spontanée. Par exemple, pour des raisons éthiques et techniques, le participant est souvent conscient d'être filmé, ce qui peut influencer sa façon de se comporter. Même s'il n'est pas conscient d'être filmé, le simple fait de savoir qu'il participe à une étude ou qu'il est observé pourrait avoir un impact sur l'apparence visuelle des expressions faciales du participant. En effet, les normes sociales peuvent nous amener à tenter de

masquer certaines expressions faciales (e.g., colère) et à en exagérer d'autres (e.g., joie). Une telle hypothèse a d'ailleurs été proposée par Ekman (1972), dans sa théorie neuro-culturelle des expressions faciales.

Coleman (1949) est l'un de ceux qui ont tenté de capturer des expressions faciales spontanées. Dans son étude, Coleman (1949) a filmé des individus alors qu'il les exposait à différentes situations ayant pour but de générer des émotions spécifiques. Par exemple, l'une des situations consistait à bander les yeux du participant, à placer un serpent sur la table devant lui sans l'avertir, et à retirer le bandeau rapidement. Il a ensuite présenté un sous-ensemble de ces vidéos de même qu'une liste des situations émotionnelles utilisées à un autre groupe de participants. La tâche de ce nouveau groupe de participant consistait à associer chaque expression faciale à l'une des situations émotionnelles sur la liste. Les participants avaient une performance significativement supérieure au hasard dans cette tâche, appuyant l'idée que les expressions faciales capturées par les films reflétaient bien l'émotion vécue. Malheureusement, Coleman (1949) a utilisé les vidéos de deux individus seulement sur les neuf individus filmés, ce qui remet en doute la généralisation possible de ses résultats. De plus, les participants étaient conscients d'être filmés, ce qui pourrait, tel que discuté ci-haut, avoir influencé leur réaction émotionnelle. D'autres études ont obtenu des résultats indiquant un lien entre les expressions faciales spontanées et l'émotion vécue mais la plupart

comparaient seulement deux situations (e.g., situations stressantes vs. non stressantes; Ekman & Bressel, 1964; Ekman & Rose, 1965; Howell & Jorgenson, 1970; Lanzetta & Kleck, 1970). Par ailleurs, certaines études ont, au contraire, observé une performance au niveau du hasard dans la catégorisation d'expressions spontanées (Hess & Blair, 2001; Motley & Camden, 1988; Naab & Russell, 2007; Wagner, 1990; Wagner, Lewis, Ramsay, & Krediet, 1992; Wagner, MacDonald, & Manstead, 1986; Yik, Meng, & Russell, 1998). Les expressions faciales spontanées utilisées dans ces études étaient toutefois sujettes aux problèmes mentionnés ci-haut. De plus, certains sont parvenus à créer des modèles computationnels capable de catégoriser des expressions faciales spontanées (Douglas-Cowie, Cowie, & Schröder, 2000; Grimm, Dastidar, & Kroschel, 2006 ; Sebe et al., 2004), ce qui suggère la présence d'un signal suffisamment systématique dans ces expressions pour qu'elles puissent être classifiées.

Comme nous ne vivons pas dans un monde idéal, la plupart des études de jugement d'expressions faciales utilisent comme stimuli des expressions faciales simulées plutôt que spontanées. Dans cette situation, la tâche des participants consiste à indiquer, généralement parmi six ou sept choix d'émotions, lequel correspond à l'expression faciale simulée par l'acteur. Les performances dans ce type de tâche sont bien supérieures au hasard, indiquant que le signal contenu dans l'expression faciale est suffisant pour mener à une

reconnaissance adéquate de l'émotion simulée par l'acteur (e.g., Adolphs et al., 2005; Simon, Craig, Gosselin, Belin, & Rainville, 2008; Smith, Cottrell, Gosselin, & Schyns, 2005). Par contre, certains ont soulevé la possibilité que les expressions faciales simulées constituent une forme spéciale d'expression qui, de manière similaire au langage, respecte un ensemble de conventions compris par tous, mais qui ne représente en rien l'apparence des expressions faciales spontanées (Hunt, 1941; Landis, 1924). L'un des principaux contre-arguments à cette hypothèse provient des études appuyant l'universalité des expressions faciales simulées.

1.2.2 L'universalité des expressions faciales d'émotions

Charles Darwin fut un pionnier dans l'exploration de l'apparence des différentes expressions faciales d'émotions, et son ouvrage « *The Expression of the Emotions in Man and Animals* » créa une immense controverse lors de sa parution en 1872. En effet, alors que la majorité de la population—incluant les scientifiques—était adepte de la doctrine créationniste et concevait les émotions comme un phénomène spécifique à l'espèce humaine, Darwin proposa que certaines expressions faciales étaient partagées par les humains et les animaux. Il proposa que les mouvements faciaux impliqués dans certaines expressions faciales aient dans un passé lointain rempli une fonction adaptative, aient été sélectionnés à travers l'évolution, et soient devenus innés chez plusieurs espèces. Cette proposition, en plus d'appuyer la théorie évolutionniste et de

remettre en question les croyances religieuses de nombreux scientifiques, ouvrit la porte à tout un champ de recherche visant à vérifier l'universalité des expressions faciales d'émotions.

L'hypothèse que les mouvements faciaux associés à une émotion spécifique sont innés prédit que l'on devrait trouver une forte similarité dans l'apparence de l'expression faciale pour une émotion donnée à travers les différentes cultures. Une telle similarité devrait également faire en sorte que les individus d'une culture arrivent à reconnaître les expressions faciales d'individus ayant évolués au sein d'une autre culture que la leur. Plusieurs études se sont intéressées à l'effet de la culture sur la reconnaissance des expressions faciales, mais l'étude d'Ekman et Friesen (1971) a eu un impact particulier dans le domaine étant donné la population étudiée. Ces chercheurs se sont rendus en Nouvelle-Guinée pour tester des individus ayant évolués au sein d'une tribu isolée et primitive (e.g., sans système d'écriture). Les individus testés n'avaient jamais vu de films ou de magazines, ne parlaient pas l'anglais et n'avaient jamais vécu à proximité de colonies caucasiennes. Les chercheurs ont présenté des photographies d'acteurs exprimant l'une des six émotions suivantes : la joie, le dégoût, la colère, la tristesse, la peur et la surprise. Ces six émotions ont été proposées comme étant les émotions de base par Ekman (1972), en partie parce qu'elles sont retrouvées de manière systématique à travers les études de jugements d'expressions faciales, et en partie parce qu'elles peuvent être

identifiées par des individus provenant de cultures différentes. Les participants de la tribu sont parvenus à identifier les expressions faciales présentées avec une performance bien au-delà du hasard, quoiqu'ils aient eu tendance à confondre la peur et la surprise.

Ekman et Friesen (Ekman, 1972) ont ensuite vérifié si les individus provenant d'une culture caucasienne arrivaient à reconnaître les émotions exprimées par des individus de cette tribu isolée de Nouvelle-Guinée. Ils ont demandé à des individus de la tribu (i.e., individus qui n'avaient pas participé à la première étude) d'effectuer l'expression faciale qu'ils feraient dans différentes situations hypothétiques, et ont filmé leurs performances. Ils ont ensuite demandé à des participants caucasiens qui n'avaient jamais vu d'individus provenant de cette tribu d'identifier l'émotion exprimée dans les vidéos. Les participants caucasiens ont identifié les expressions de colère, de dégoût, de tristesse et de joie avec des performances bien au-delà du hasard, mais ont confondu les expressions de peur et de surprise (i.e., celles que les Nouveau-Guinéen avaient eux-mêmes de la difficulté à discriminer). Les chercheurs ont conclu que certaines expressions faciales sont universelles : elles ont une apparence similaire et sont reconnues par toutes les cultures, même celles qui n'ont pas pu apprendre de « système de conventions », tel que proposé plus haut, suite au contact de la culture occidentale.

Plus récemment, Elfenbein & Ambady (2002) ont effectué une méta-analyse des études qui s'étaient intéressées à l'effet de la culture sur la reconnaissance des expressions faciales. Leur analyse a permis de montrer que 96.9% de ces études ont trouvé que les individus d'une culture donnée arrivent à reconnaître l'émotion exprimée par les individus d'une autre culture avec des performances au-delà du hasard. Ces résultats appuient l'hypothèse qu'il existe une certaine systémativité dans l'apparence visuelle des expressions faciales d'émotions, systémativité qui permet aux individus de différentes cultures de communiquer entre eux leurs états d'âme.

Par ailleurs, malgré que la reconnaissance interculturelle des expressions faciales soit possible avec un niveau de performance au-delà du hasard, un avantage pour reconnaître les expressions faciales des individus provenant de la même culture que nous a été observé (voir Elfenbein & Ambady, 2002 pour une méta-analyse). Ce résultat suggère que malgré que l'apparence visuelle des expressions faciales soit suffisamment similaire pour permettre une reconnaissance interculturelle, des nuances existent dans la façon exacte d'exprimer une émotion, et ces nuances varient d'une culture à l'autre. Ces variations pourraient d'ailleurs être la source des modulations observées dans les stratégies développées par le système visuel pour reconnaître les expressions faciales (Jack, et al., 2009; Jack, Caldara, & Schyns, 2012; Jack et al., 2012). Néanmoins, le fait que des individus ayant évolué au sein d'une culture

complètement isolée arrivent à décoder les expressions faciales simulées par des occidentaux (et vice-versa) appuie l'idée que les expressions faciales simulées sont représentatives de l'émotion interne vécue. Cette observation appuie également l'hypothèse qu'il existe une certaine systématisme dans l'apparence de l'expression faciale d'une émotion donnée. Maintenant, on peut donc se poser la question : quel est le signal contenu dans les expressions faciales (i.e., qu'est-ce qui caractérise l'apparence visuelle de chaque expression faciale d'émotion)?

1.2.3 Les mouvements faciaux impliqués dans les expressions faciales

Paul Ekman et coll. (1978; 1982) se sont attaqués à la tâche extrêmement laborieuse de développer une méthode pour caractériser l'apparence des expressions faciales, et ont ainsi donné naissance au système de codage *Facial Action Coding System* (FACS). Ce système est basé sur l'analyse de tous les changements visibles possibles dans l'apparence du visage suivant l'activation d'un ou de plusieurs muscles du visage. Chacun de ces changements est appelé « unité d'action ». Ekman et ses collaborateurs ont ainsi défini 46 unités d'action. Par exemple, la 38^e unité d'action représente la dilatation des narines, et est le résultat de la contraction du muscle nasalis. En plus de ces unités d'action isolées, Ekman et coll. ont défini des combinaisons d'unités d'action (i.e., constituées de deux à sept unités d'action) et les modifications d'apparence du visage qui y sont associées. Bien qu'un

entraînement soit nécessaire afin de parvenir à utiliser le FACS, ce dernier constitue un outil extrêmement intéressant puisqu'il permet de définir les expressions faciales d'émotions de manière objective. Avec ce système, il devient par exemple possible de comparer les unités d'action des expressions faciales d'émotions d'individus de différentes cultures, ou bien avec différentes psychopathologies.

Ainsi, le FACS fournit de l'information quant à la portion « transmission » de l'expression faciale. Par contre, il ne permet pas de savoir comment un observateur externe décode l'expression faciale d'un autre individu. Par exemple, ce n'est pas parce que l'expression de peur implique les unités d'action représentant le soulèvement complet des sourcils (portions internes et externes), le soulèvement de la paupière supérieure et le relâchement de la mâchoire qu'un observateur externe utilise nécessairement tous ces indices pour inférer l'émotion exprimée. En effet, il est nécessaire d'étudier les stratégies sous-jacentes à la reconnaissance des expressions faciales pour comprendre comment les humains arrivent à décoder ce signal social. L'objectif de cette thèse est d'étudier la façon dont les humains (ici, uniquement des caucasiens) extraient l'information visuelle pour discriminer toutes les expressions faciales d'émotions de base entre elles.

1.3 L'expression faciale : son décodage.

La perception et l'association du percept avec les connaissances constituent deux des principaux processus impliqués dans la reconnaissance des expressions faciales (Adolphs, 2002). Par exemple, pour reconnaître une expression faciale de peur, l'information visuelle pour se créer un percept adéquat du visage apeuré doit être extraite, et ce percept doit être associé aux connaissances que nous avons acquises par rapport à un tel stimulus : dans quel genre de contexte il s'est produit (e.g., situations dangereuses), quels ont été les comportements des gens qui affichaient cette expression (e.g., crier, fuir, figer), et quelles conséquences ont suivi (e.g., blessure, évitement du danger). La présente thèse s'intéresse à la portion perceptive de la reconnaissance des expressions faciales, plus spécifiquement aux stratégies visuelles impliquées dans cette tâche. Nous nous attarderons donc principalement sur ce processus dans la présente section.

1.3.1 Information utile pour reconnaître les expressions faciales : le haut vs. le bas de visage

L'une des questions centrales à la perception des expressions faciales porte sur l'étude de l'information visuelle utilisée pour reconnaître ou catégoriser un tel stimulus. Les premières études à s'être intéressées à cette question l'ont abordé du point de vue des régions faciales utiles pour reconnaître une expression faciale d'émotion (Bassili, 1979; Calder, Young,

Keane, & Dean, 2000; Coleman, 1949; Dunlap, 1927; Frois-Wittmann, 1930; Hanawalt, 1944; Nummenmaa, 1964; Plutchik, 1962; Ruckmick, 1921). La procédure utilisée dans ces études consistait à diviser les stimuli faciaux en régions faciales (i.e., généralement le haut du visage vs. le bas du visage), à présenter les régions faciales isolées ou non isolées aux participants et à mesurer leur performance dans ces différentes conditions. Les résultats obtenus dans ces différentes études ne permettent pas de conclusion claire quant à l'importance relative des différentes régions de visage. En effet, certaines suggèrent que la portion inférieure du visage soit plus importante pour la reconnaissance des expressions faciales (Dunlap, 1927; Ruckmick, 1921); certaines suggèrent que le haut et le bas du visages ont une importance égale (Coleman, 1949; Frois-Wittman, 1930); et certaines suggèrent que l'importance des différentes régions du visage dépend de l'expression faciale à reconnaître (Bassili, 1979; Calder et al., 2000; Hanawalt, 1944; Nummenmaa, 1964; Plutchik, 1962).

La procédure utilisée dans ces études présente toutefois une faiblesse majeure. En effet, la division du visage en régions définies à priori, par exemple le haut vs. le bas du visage, obligeait les chercheurs à se positionner sur ce que constitue une partie, ou une caractéristique, du visage expressif. Or, cette question était la problématique même de leur recherche (i.e., quelle région de visage est importante pour la reconnaissance des expressions faciales), donc la

définition a priori des régions introduisait une certaine circularité à ces études. En effet, ce n'est pas parce que les humains ont donné un nom à différentes régions anatomiques du visage (ex. les yeux, le nez, la bouche), que ces différentes régions constituent une « partie » pour le système visuel. Ainsi, la division du visage expressif en « parties » définies arbitrairement rend difficile l'interprétation des résultats obtenus. Par exemple, si une « partie » pour le système visuel inclut une portion du visage expressif contenue à la fois dans le haut et dans le bas du visage (e.g., dans l'expression de dégoût, les plis qui se forment sur l'arête du nez *et* les plis qui se forment autour du nez; voir Figure 1), alors on peut s'attendre à ce que la présentation isolée du haut ou du bas du visage altère également la performance.

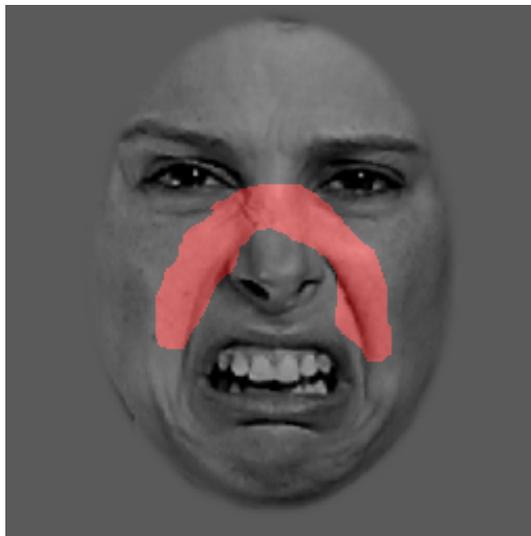


Figure 1. Exemple d'une partie hypothétique pour le système visuel. Cette « partie » n'est pas reliée à une étiquette verbale telle que le nez ou la bouche. De plus, elle inclut à la fois le haut et le bas du visage.

Plus récemment, l'information utilisée pour reconnaître les expressions faciales d'émotions a été investiguée (Smith et al., 2005) à l'aide de la méthode des Bulles (Gosselin & Schyns, 2001), qui présente l'avantage de ne pas exiger de définition a priori des « parties » constituantes des expressions faciales. De plus, la méthode utilisée par Smith et al. (2005) offrait la possibilité d'investiguer non seulement les « parties » de visage utilisées pour reconnaître une expression faciale, mais également dans quelles fréquences spatiales ces parties sont utilisées. L'étude de l'utilisation des fréquences spatiales pour le traitement de divers types de stimuli (e.g., mots, visages, objets) a suscité beaucoup d'intérêt dans les deux dernières décennies (e.g., Costen, Parker, & Craw, 1996; Gaspar, Sekuler, & Bennett, 2008; Gold, Bennett, & Sekuler, 1999; Kumar & Srinivasan, 2011; Näsänen, 1999; Smith & Schyns, 2009; Willenbockel et al., 2010). L'utilisation des fréquences spatiales dans le traitement des expressions faciales a toutefois très peu été investiguée, bien que récemment quelques études s'y soient intéressées (Kumar & Srinivasan, 2011; Smith et al., 2005; Smith & Schyns, 2009). La théorie sous-jacente à ces études sera présentée ici puisqu'elle est centrale dans le domaine de la cognition visuelle et fait partie inhérente de la méthode utilisée dans la présente thèse.

1.3.2 Information utile pour reconnaître les expressions faciales : les fréquences spatiales

L'une des théories en vogue dans le domaine de la vision propose que le système visuel décompose toute stimulation visuelle en ses fréquences spatiales constituantes, de manière similaire à une analyse de Fourier (De Valois, De Valois, & Yund, 1979; Maffei & Fiorentini, 1973). Selon cette théorie, le traitement visuel précoce reposerait sur un ensemble de canaux, chacun étant particulièrement sensible à une bande de fréquences spatiales spécifique. La Figure 2 donne un aperçu de l'information spatiale représentée par les basses vs. les hautes fréquences spatiales : les basses fréquences spatiales codent essentiellement la structure globale d'un stimulus (Fig. 2a), alors que les hautes fréquences spatiales codent plutôt les détails fins (Fig. 2b). Les cellules simples et les cellules complexes du système visuel primaire (i.e., V1) ont été proposées comme substrat physiologique de ce mécanisme de filtrage (De Valois, Albrecht, & Thorell, 1982).

La théorie des fréquences spatiales a nourri tout un pan de la recherche sur les composantes visuelles utilisées pour reconnaître différents types de stimuli. En effet, si le système visuel décompose un stimulus en ses fréquences spatiales constituantes, il est probable que certaines des fréquences spatiales de ces stimuli soient plus utiles que d'autres pour leur reconnaissance, soit parce que le système visuel est plus sensible à ces dernières ou parce qu'elles codent certaines portions plus discriminante du stimulus. Par exemple, le spectre

fréquentiel (i.e., l'énergie contenue dans chacune des fréquences spatiales) d'un visage suit une fonction du type $1/\text{fréquence spatiale}^2$ (i.e., $1/f^2$). Les basses fréquences spatiales sont donc celles qui contiennent le plus d'énergie (i.e., qui sont les plus dominantes dans un stimulus de visage). Pourtant, il semble que les fréquences spatiales moyennes (i.e., situées entre 8 et 20 cycles par visage) soient optimales pour la reconnaissance des visages (Costen, Parker, & Craw, 1996; Gold, Bennett, & Sekuler, 1999; Näsänen, 1999; Willenbockel et al., 2010).

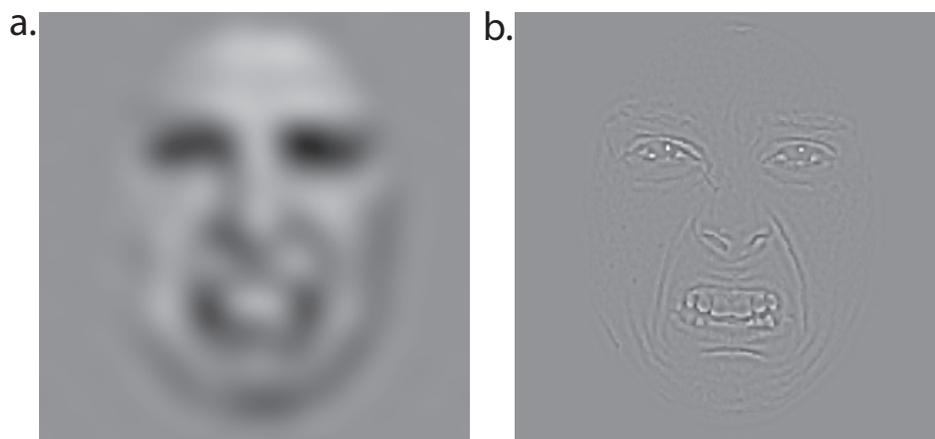


Figure 2. Exemple de filtrage fréquentiel. Visage exprimant le dégoût et filtré de sorte que seules les basses fréquences spatiales (a) ou seules les hautes fréquences spatiales (b) soient présentes.

Peu d'études se sont intéressées aux fréquences spatiales utilisées efficacement pour la reconnaissance des expressions faciales. Les efforts ont plutôt été dirigés vers la mise en évidence d'une voie sous-corticale (i.e., colliculus supérieur – pulvinar – amygdale) très rapide qui serait utilisée pour

reconnaitre la peur, et qui repose principalement sur le traitement des basses fréquences spatiales (Johnson, 2005; Vuilleumier, Armony, Driver, & Dolan, 2003). L'importance de cette voie pour le traitement des expressions faciales est toutefois remise en question (de Gelder, van Honk & Tamietto, 2011; Pessoa & Adolphs, 2010; Tamietto & de Gelder, 2010). Par ailleurs, quelques rares études se sont intéressées spécifiquement aux fréquences spatiales utilisées pour reconnaître les expressions faciales. Kumar & Srinivasan (2011) ont montré que la reconnaissance de l'expression faciale de joie repose sur le traitement des basses fréquences spatiales, alors que la reconnaissance de l'expression faciale de tristesse repose sur le traitement des hautes fréquences spatiales. Smith & Schyns (2009) ont quant à eux montré que les basses fréquences spatiales sont plus utiles pour la reconnaissance des expressions de joie, de surprise, de dégoût et de colère, alors que les hautes fréquences spatiales sont plus utiles pour la reconnaissance des expressions de peur, de tristesse et de neutralité. L'étude de Smith et al. (2005) a examiné quelles régions faciales étaient utilisées et dans quelles bandes de fréquences spatiales ces dernières étaient utilisées. Ainsi, cette étude fait le pont entre deux dimensions de l'information visuelle utilisée pour reconnaître les expressions faciales : les « parties » de visage, et les fréquences spatiales. Afin d'arriver à étudier conjointement ces deux dimensions, Smith et al. (2005) ont utilisé, tel que mentionné ci-haut, la méthode des Bulles.

1.3.3 Information utile pour reconnaître les expressions faciales : la méthode des Bulles

La méthode des Bulles permet de révéler l'information visuelle utilisée de manière efficace par des participants pour faire une tâche sur un stimulus visuel, par exemple catégoriser une expression faciale d'émotion. Cette méthode consiste à échantillonner aléatoirement une ou plusieurs des dimensions (e.g., dimension de l'étendue spatiale—i.e., largeur et hauteur, dimension des fréquences spatiales, dimension temporelle) d'un stimulus visuel, et à vérifier l'impact d'un tel échantillonnage sur certains aspects du comportement ou de l'état du participant, par exemple sa performance ou son électroencéphalogramme. La version de la méthode des Bulles utilisée par Smith et al. (2005) pour examiner l'information utile pour la reconnaissance des expressions faciales d'émotions échantillonnait les dimensions de l'étendue spatiale et des fréquences spatiales. Puisque l'une des expériences centrales à la présente thèse utilisait la même version de la méthode des Bulles que celle utilisée par Smith et al. (2005), celle-ci sera ici décrite en détails.

Dans l'étude de Smith et al. (2005), la tâche des participants consistait à catégoriser des expressions faciales présentées sur un écran d'ordinateur. À chaque essai, le visage présenté était échantillonné aléatoirement sur les dimensions des fréquences spatiales et de l'étendue spatiale selon la procédure qui suit. Le visage était d'abord décomposé en cinq bandes de fréquences

spatiales. Cette décomposition avait pour résultat de générer cinq images du visage expressif, représenté dans des fréquences spatiales distinctes (voir Fig. 3, rangée supérieure). Pour chacune de ces cinq images, un masque de bulles était généré aléatoirement (voir Fig. 3, rangée du milieu). Une bulle consiste en une ouverture gaussienne à travers laquelle le stimulus est visible. À chaque essai, les bulles étaient positionnées à des endroits différents, sélectionnés aléatoirement, sur les cinq images de visage. La position des bulles variait donc d'un essai à l'autre, mais également d'une bande de fréquence à l'autre. La taille d'une bulle variait en fonction de la bande de fréquence à laquelle elle était destinée, de sorte que le nombre de cycles par image révélé par une bulle était constant à travers les bandes de fréquences spatiales. Le nombre de bulles variait également en fonction de la bande de fréquence de sorte que la quantité de pixels révélés soit constante d'une bande de fréquences spatiales à l'autre. Chacun des cinq masques de bulles était ensuite appliqué à l'image de visage filtré qui lui correspondait, de sorte que seules les portions du visage filtré situées aux endroits où des bulles étaient placés étaient visibles (voir Fig. 3, rangée inférieure). Finalement, les cinq visages filtrés et bullés étaient regroupés en un seul stimulus, dans lequel seules certaines régions du visage, dans certaines bandes de fréquence, étaient visibles. Par exemple, dans le stimulus final présenté sur la dernière rangée de la Figure 3, la bouche est visible et est représentée par les quatre bandes de fréquences spatiales les plus élevées. Un stimulus tel que celui présenté à la Figure 3 était produit à chaque

essai, et la tâche du participant consistait à le catégoriser comme étant une expression de colère, de dégoût, de peur, de joie, de tristesse, de surprise ou de neutralité.

L'hypothèse sous-jacente à la méthode des Bulles est que la probabilité que le participant catégorise correctement l'expression faciale augmente lorsque l'information nécessaire pour y parvenir est visible dans le stimulus, alors qu'elle diminue lorsque l'information nécessaire pour y parvenir est masquée. Ainsi, suite à de nombreux essais, il est possible d'inférer quelle information était nécessaire pour réussir la tâche en corrélant les masques de bulles et la performance correspondante à chaque essai. Plus spécifiquement, dans le cas de l'étude de Smith et al. (2005), la méthode des Bulles permettait de révéler quelles régions du visage étaient utilisées de manière efficace¹ pour catégoriser chacune des expressions faciales d'émotions, et dans quelle(s) bande(s) de fréquences spatiales ces régions étaient utiles.

¹ La notion d'efficacité réfère ici au fait que l'information était corrélée avec la performance. En d'autres mots, il est possible qu'une autre portion de l'information contenue dans le visage soit échantillonnée et même traitée par le système visuel, mais cette information n'augmente pas la probabilité de catégoriser correctement l'expression du visage.

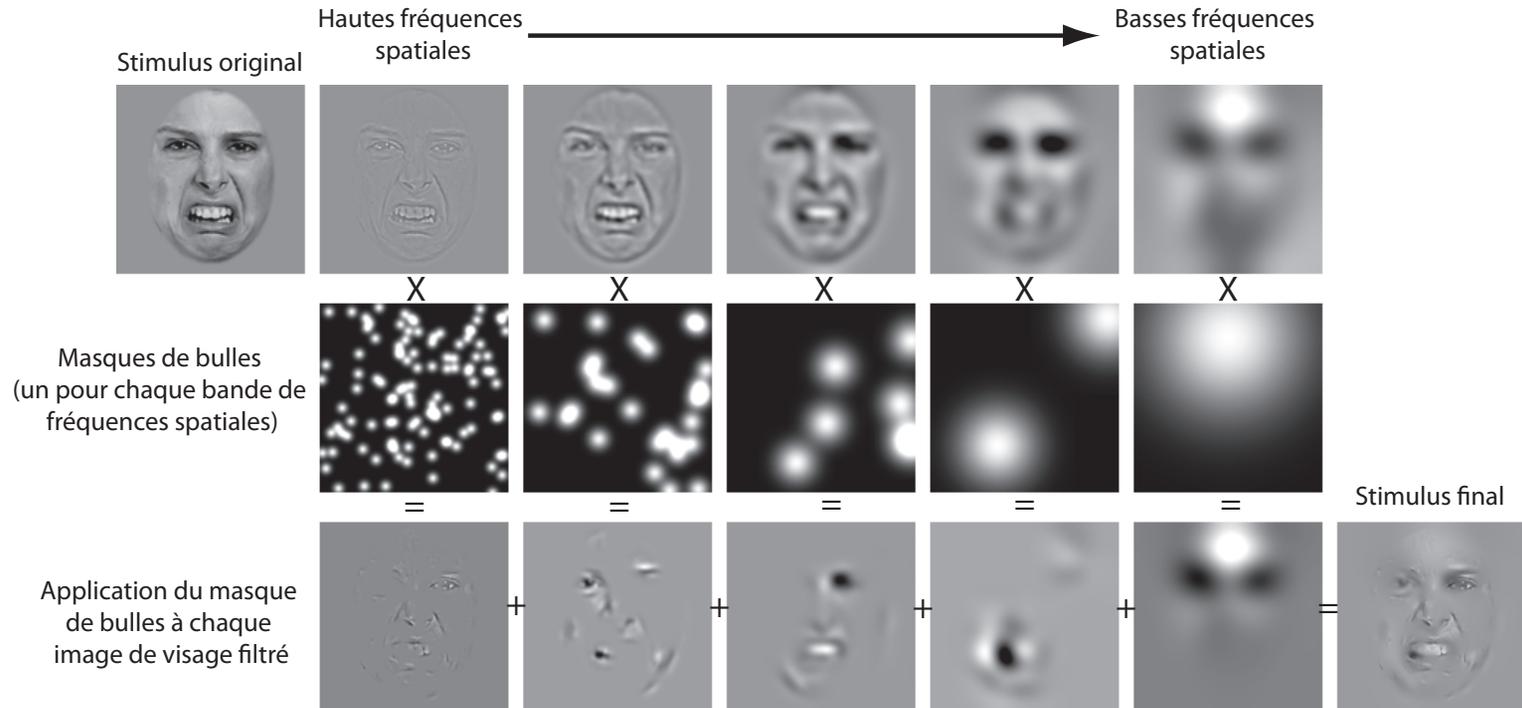


Figure 3. Exemple de création d'un stimulus avec la méthode des Bulles. Cette procédure est répétée à chaque essai de l'expérience afin de générer aléatoirement des stimuli qui échantillonnent différentes portions de l'information à chaque essai.

Les résultats de Smith et al. (2005) montrent que selon la catégorie à laquelle appartient l'expression faciale présentée, l'information utilisée efficacement diffère grandement. Par exemple, la région des yeux est utilisée efficacement pour reconnaître l'expression faciale de peur, alors que les plis nasogéniens constituent la région la plus utilisée pour reconnaître l'expression faciale de dégoût. Les auteurs ont également soumis à leur tâche un observateur idéal (i.e., un modèle informatique qui utilise tous les pixels des images afin de catégoriser l'expression faciale présentée). L'analyse de l'observateur idéal leur a permis de vérifier l'information disponible—par opposition à l'information utilisée par les observateurs humains—dans chaque expression faciale. Ils ont montré qu'en tant que « transmetteur » d'émotion, l'expression faciale a évolué de sorte que l'apparence de chaque expression faciale diffère le plus possible de celle des autres. Ils ont également montré que l'humain, en tant que « décodeur » d'expression faciale, a développé des stratégies visuelles qui ont pour but de relever les ambiguïtés résiduelles. Par exemple, les expressions de peur et de surprise sont celles qui partagent le plus de signal. En effet, la région des yeux et la région de la bouche ont toutes les deux une apparence similaire dans ces deux expressions : les yeux sont grand ouverts, et la bouche forme un « O ». Pourtant, les stratégies visuelles utilisées par les humains pour reconnaître ces deux expressions sont très différentes : ce sont surtout les yeux qui sont utilisés efficacement pour reconnaître la peur, alors que c'est principalement la bouche qui est utilisée pour reconnaître la surprise.

Ainsi, l'étude de Smith et al. (2005) a permis une avancée importante dans nos connaissances sur l'information visuelle utilisée pour reconnaître chacune des expressions faciales de base. Plusieurs questions demeurent toutefois en suspens. Par exemple, l'examen de l'information utile pour chaque expression prise isolément est intéressant, mais on peut se demander si certaines portions de l'information visuelle revêtent une importance supérieure à travers *toutes* les expressions faciales de base. En effet, le système visuel ne sait pas quelle expression faciale il traite avant qu'une certaine quantité d'information soit extraite. Puisque ce dernier présente des capacités limitées (Levin & Simons, 1997; Simons & Rensink, 2005), il est probable qu'il alloue un maximum de ressources à l'information visuelle la plus discriminante à travers toutes les expressions faciales. Une telle stratégie lui permettrait de réduire le plus possible l'ambiguïté quant à l'expression faciale traitée avant d'allouer des ressources vers une région contenant du signal plus spécifique à une expression donnée. L'un des buts de la première étude présentée dans cette thèse est de vérifier quelles « parties » de visage sont utilisées efficacement à travers toutes les expressions faciales.

Malgré que peu d'études aient directement examiné l'information visuelle utilisée à travers toutes les expressions faciales—et que les résultats de celles qui l'ont fait ne font pas consensus, voir section 1.3.1—une croyance prédominante dans le domaine accorde une importance particulière à la région

des yeux pour la reconnaissance des expressions faciales d'émotions (e.g., Itier & Batty, 2009; Lischke et al., 2012; Eisenbarth & Alpers, 2011). D'ailleurs, une quantité impressionnante d'expressions dans notre langage—e.g., avoir des fusils dans les yeux, avoir des étoiles dans les yeux, avoir le regard assassin, etc.—suggère que cette croyance est partagée par la population non-scientifique. Les études qui pourraient avoir alimenté cette croyance seront présentées dans la prochaine section.

1.3.4 La région des yeux est-elle spéciale?

La région des yeux n'a jamais été démontrée comme cruciale pour discriminer les expressions faciales de base les unes des autres. La croyance qui attribue une importance spéciale à la région des yeux est probablement basée sur des résultats provenant d'études menées en enregistrement des mouvements oculaires, sur le développement du traitement visuel des visages chez les nourrissons et sur le traitement visuel des expressions faciales chez des populations cliniques.

Typiquement, les fixations oculaires durant le traitement des visages se situent sur les caractéristiques internes du visage (e.g., yeux, nez, bouche; Yarbus, 1967). Plusieurs études ont montré que parmi les régions internes du visage, la région des yeux est la plus fixée durant l'identification de visages (Blais et al., 2008; Henderson, Williams, & Falk, 2005; Janik, Wellens,

Goldberg, & Dell’Osso, 1978). Une étude d’Eisenberg & Alpers (2011) avait pour but d’examiner les patrons de mouvements oculaires durant le traitement des expressions faciales. Les auteurs ont conclu que le biais pour fixer la région des yeux trouvé en identification de visages est également présent durant le traitement des expressions faciales. Ces résultats pourraient avoir contribué à alimenter la croyance que la région des yeux est « spéciale » pour la reconnaissance des expressions faciales. Toutefois, puisqu’une dissociation peut être observée entre les fixations attentionnelles et les fixation oculaires (Arizpe, Kravitz, Yovel & Baker, 2012; Jonides, 1981; Posner, 1980; voir cependant Rayner, 1998; Deubel & Schneider, 2003; Godijn & Theeuwes, 2003), la prudence s’impose lorsque l’on veut inférer l’information utilisée à partir de l’examen des fixations oculaires. De plus, une lecture plus approfondie de l’étude d’Eisenberg & Alpers (2011) suggère qu’en fait les auteurs ne trouvent jamais de différence statistiquement significative entre la densité de fixations oculaires—ni sur le nombre de fixations ni sur la durée de celles-ci—sur la région des yeux vs. sur la région de la bouche. Autrement dit, les yeux ne sont pas significativement plus fixés que la bouche. Les résultats observés en mouvement oculaires durant le traitement des expressions faciales ne permettent donc aucunement de conclure à une nature spéciale de la région des yeux.

Les études menées chez les nourrissons montrent que ces derniers préfèrent regarder la région des yeux plutôt que d'autres régions d'un visage (Hainline, 1978; Haith, Bergman, & Moore, 1977; Maurer, 1985). Ils préfèrent également regarder un visage avec les yeux ouverts qu'un visage avec les yeux fermés (Batki et al., 2000) et préfèrent regarder un visage dont le regard est dirigé vers eux qu'un visage dont le regard est dirigé vers un autre endroit (Farroni, Csibra, Simion, & Johnson, 2002). L'enregistrement de l'électroencéphalogramme des nourrissons révèle également une augmentation de l'amplitude de la composante N170—une composante associée au traitement des visages chez les adultes—pour des visages dont le regard est dirigé vers eux par rapport à des visages dont le regard est dirigé ailleurs (Farroni, et al., 2002). Ces résultats ont été interprétés par plusieurs comme une preuve que le cerveau humain est prêt dès la naissance à détecter l'information sociale. Certains ont même suggéré l'existence d'un module cérébral inné dont les fonctions seraient de détecter la présence d'yeux dans l'environnement et de traiter leur direction (Baron-Cohen, 1995; Batki et al., 2000; voir cependant Johnson, 2005 et Johnson & Morton, 1991). Encore une fois, ces résultats ont probablement eux aussi contribué à la croyance que la région des yeux a une importance cruciale pour la reconnaissance des expressions faciales.

La troisième série de résultats qui pourrait avoir alimenté la croyance que la région des yeux est spéciale dans le traitement des expressions faciales

provient des études sur les stratégies visuelles de populations cliniques. Une sous-utilisation des yeux ou une diminution des fixations oculaires sur la région des yeux est observée chez plusieurs populations cliniques présentant une altération de la capacité à reconnaître les expressions faciales. Par exemple, suite à une lésion bilatérale des amygdales, la patiente SM s'est retrouvée avec un déficit sévère à reconnaître l'expression faciale de peur (Adolphs, Tranel, Damasio & Damasio, 1994). Une étude utilisant la méthode des bulles et l'enregistrement des mouvements oculaires a permis de montrer que son déficit est lié à une sous-utilisation de la région des yeux et à une diminution des fixations oculaires dirigées vers cette même région (Adolphs et al., 2005). De manière surprenante, le simple fait de donner l'instruction à SM de regarder la région des yeux pendant qu'elle effectue une tâche de catégorisation d'expressions faciales ramène sa performance à un niveau normal pour reconnaître l'expression faciale de peur. Une diminution de l'utilisation de la région des yeux (Baron-Cohen, Wheelwright, & Jolliffe, 1997; Gross, 2008; Spezio, Adolphs, Hurley, & Piven, 2007a; 2007b) et une diminution des fixations oculaires sur la région des yeux ont également été observée chez les individus autistes (Corden, Chilvers, & Skuse, 2008; Pelphrey et al. 2002; Spezio, Adolphs, Hurley, & Piven, 2007a; 2007b), lesquels présentent aussi un déficit à reconnaître les expressions faciales (Humphreys, Minshew, Leonard, & Behrmann, 2007; voir Harms, Martin, & Wallace, 2010 pour une revue de la littérature). Une sous-utilisation de la région des yeux a été observée chez des patients prosopagnosiques (Bukach, Bub, Gauthier, & Tarr, 2006; Bukach,

LeGrand, Kaiser, Bub, & Tanaka, 2008; Caldara et al., 2005; Rossion, Kaiser, Bub, & Tanaka, 2009), i.e., souffrant d'un trouble spécifique à la reconnaissance des visages familiers suite à une lésion cérébrale. Or, les patients prosopagnosiques présentent eux aussi un déficit dans la reconnaissance des expressions faciales (Caldara et al., 2007; Humphreys, Avidan, & Behrmann, 2007). D'ailleurs, une sous-utilisation de la région des yeux durant le traitement de l'expression faciale de peur a été observée chez la patiente prosopagnosique PS (Caldara et al., 2007). Enfin, une diminution des fixations oculaires sur la région des yeux pendant le traitement des expressions faciales est retrouvée chez les personnes âgées (Sullivan, Ruffman, & Hutton, 2007), qui présentent elles aussi un déficit dans la reconnaissance de certaines expressions faciales de base, notamment la colère, la peur et la tristesse (Sullivan & Ruffman, 2004).

Bien entendu, l'observation d'un lien entre un déficit de la reconnaissance des expressions faciales et une diminution de l'utilisation de la région des yeux n'implique pas nécessairement que cette région est la plus importante pour cette tâche. En effet, il est possible qu'un déficit commun à toutes ces populations, par exemple un déficit à traiter les hautes fréquences spatiales, soit sous-jacent au trouble de la reconnaissance des expressions faciales *et* au traitement de la région des yeux dans un visage exprimant une émotion, ce qui expliquerait leur cooccurrence. Une autre possibilité, que nous soulevons dans cette thèse, serait que la région des yeux contienne en fait moins

d'information que la région de la bouche et soit donc touchée plus rapidement par la diminution des ressources visuo-cognitives. Bref, avant d'affirmer que la région des yeux est cruciale pour la reconnaissance des expressions faciales, il est nécessaire de vérifier directement son utilisation, relativement aux autres régions du visages, durant cette tâche. L'un des articles de cette thèse porte sur cette problématique.

1.3.5 Les expressions faciales dynamiques

Les études qui se sont intéressées à l'information utile pour reconnaître les expressions faciales ont, pour la plupart, utilisé des stimuli statiques. Or, les expressions faciales ont de manière inhérente une nature dynamique. Un modèle de la perception des visages, développé par Haxby et collaborateurs, propose d'ailleurs qu'étant donné la nature dynamique des expressions faciales, des systèmes neuraux différents sont impliqués dans le traitement de ces stimuli vs. dans le traitement de l'identité d'un visage (Haxby, Hoffman, & Gobbini, 2000; 2002). Le modèle de Haxby est en partie basé sur le modèle de Bruce et Young (1986), qui postulait l'existence de modules fonctionnels distincts pour le traitement de l'identité et de l'expression faciale. Le modèle de Haxby pousse toutefois l'idée plus loin en proposant l'existence de deux types de représentations : la représentation des aspects invariants des visages et la représentation des aspects changeants des visages. La représentation des aspects invariants serait davantage recrutée pour l'identification des visages,

puisque'une identité doit pouvoir être reconnue malgré les changements transitoires pouvant survenir dans l'apparence d'un visage, par exemple lorsque la personne parle ou exprime une émotion. La représentation des aspects changeants serait quant à elle davantage recrutée pour le traitement des expressions faciales. Haxby et collaborateurs ont proposé que le gyrus fusiforme (communément appelé le *Fusiform Face Area* ou *FFA* dans le domaine de la perception des visages) serait impliqué dans le traitement des aspects invariants, alors que le sulcus temporal supérieur (i.e., STS) serait impliqué dans le traitement des aspects changeants.

Le STS a été associé au traitement du mouvement biologique (Grossman & Blake, 2002; Puce et al., 1998). Il est donc plausible que son implication dans le traitement des expressions faciales soit reliée au mouvement inhérent à ces dernières. Par ailleurs, cette région est aussi activée durant le traitement des expressions faciales statiques (Furl, van Rijsbergen, Treves, Friston, & Dolan, 2007; Haxby, Hoffman, & Gobbini, 2000; Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001; Tsuchiya, Kawasaki, Oya, Howard, & Adolphs, 2008). Haxby et collaborateurs expliquent cette observation par la proposition que la perception adéquate d'une expression faciale statique implique la connaissance des mouvements faciaux sous-jacents à l'apparence finale (et statique) de cette expression. Ainsi, les expressions faciales statiques et dynamiques partagent certains systèmes neuronaux. Toutefois, des études qui ont directement comparé les réseaux neuronaux recrutés par les deux types de

stimuli indiquent une plus grande activation des régions cérébrales impliquées dans le traitement des expressions faciales (e.g., FFA, STS, cortex orbitofrontal, régions parahippocampiques) avec les stimuli dynamiques qu'avec les stimuli statiques (Kilts et al., 2003; LaBar et al., 2003; Sato et al., 2004; Trautmann, Fehr, & Herrmann, 2009). Ces différences sur le plan de l'activité cérébrale se reflètent d'ailleurs au niveau comportemental. En effet, plusieurs études suggèrent que le mouvement facilite la reconnaissance des expressions faciales (Bassili, 1978; Chiller-Glaus et al., 2011; Matsuzaki & Sato, 2008), quoique cet effet semble plus robuste lorsque les expressions faciales sont subtiles (Ambadar, Schooler, & Cohn, 2005; Bould, Morris, & Wink, 2008; Kätsyri & Sams, 2008). Par ailleurs, les stimuli dynamiques contiennent évidemment plus d'information que les stimuli statiques : les stimuli statiques présentent généralement l'apex de l'expression, alors que les stimuli dynamiques présentent l'apex en plus de tous les mouvements faciaux qui le précèdent. Cette différence au niveau de la quantité d'information contenue dans les deux types de stimuli pourrait donc expliquer la facilitation observée sur le plan comportemental de même que l'augmentation de l'activité cérébrale dans les régions reliées au traitement des expressions faciales avec les stimuli dynamiques. D'ailleurs, lorsque la quantité d'information contenue dans le stimulus est contrôlée, l'efficacité du système visuel (i.e., la quantité de signal nécessaire pour un humain vs. pour un observateur idéal pour catégoriser correctement l'expression présentée) est égale (Barker et al., 2012), voire moindre (Hammal, Gosselin, & Fortin, 2008), avec les stimuli dynamiques

qu'avec les stimuli statique. Toutefois, ces dernières études ne prenaient pas en compte l'information de *mouvement*, mais bien l'information au niveau de la dimension de l'étendue spatiale à différents moments dans le temps. Il est probable que l'information de mouvement soit aussi utilisée par le système visuel durant le traitement des expressions faciales dynamiques.

La facilitation observée avec les expressions faciales dynamiques est généralement interprétée comme un appui à l'idée que le décours temporel des expressions faciales contient de l'information critique pour leur reconnaissance (Ekman, Friesen, & Hagar, 2002). Puisqu'il est très rare que nous soyons exposés à des expressions faciales statiques, il est probable que le système visuel ait développé des stratégies d'extraction de l'information qui soient adaptées au traitement des expressions faciales dynamiques et qui tirent avantage de la signature temporelle des expressions faciales. L'un des objectifs de cette thèse est d'investiguer les stratégies utilisées pour extraire l'information visuelle utile à la reconnaissance des expressions faciales dynamiques, ce qui n'a jamais été fait auparavant.

1.4 Résumé et objectifs de la présente thèse

L'idée que l'information visuelle utile à la reconnaissance des expressions faciales varie en fonction de l'expression présentée est maintenant généralement bien acceptée (Bassili, 1979; Cunningham, Kleiner, Bülhoff, &

Wallraven, 2004; Ekman, 1982; Hanawalt, 1944; Nummenmaa, 1964; Nusseck, Cunningham, Wallraven, Bülthoff, 2008; Plutchik, 1962; Smith et al., 2005). Toutefois, étant donnée les capacités limitées du système visuo-cognitif (Levin & Simons, 1997; Simons & Rensink, 2005), et le fait qu'il ignore quelle expression faciale il traite avant de l'avoir au moins partiellement traitée, il est probable qu'il ait développé des stratégies d'extraction de l'information visuelle qui lui permettent de réduire autant que possible l'ambiguïté quant à l'expression faciale traitée. Si c'est le cas, cette stratégie consisterait à extraire l'information la plus discriminante à travers toutes les expressions faciales.

Le premier article de cette thèse, intitulé « The eyes are not the window to basic emotions », vise à (1) investiguer la stratégie d'extraction de l'information utilisée pour discriminer toutes les expressions faciales de base entre elles et ce, pour des expressions faciales présentées de manière statique ou dynamique (i.e., la portion *décodage* de l'expression faciale); (2) vérifier, à l'aide de modèles computationnels, où se situe l'information la plus discriminante à travers toutes les expressions faciales de base et ce, pour des expressions faciales présentées de manière statique ou dynamique (i.e., la portion *transmission* de l'expression faciale); et (3) mettre en relation les

stratégies utilisées par les humains (ici, de culture caucasienne)² et l'information contenue dans les expressions faciales de base.

Le deuxième article de cette thèse, intitulé « Individual differences in the visual extraction strategies of facial expressions », vise à vérifier si la stratégie d'extraction de l'information visuelle révélée dans la première étude de cette thèse est reliée à l'habileté de chaque individu à reconnaître les expressions faciales. En effet, une modification de la stratégie d'extraction de l'information visuelle a été observée chez plusieurs populations cliniques dont la reconnaissance des expressions faciales est altérée (i.e., voir section 1.3.4; Adolphs et al., 2005; Caldara et al., 2007; Langner, Becker, & Rinck, 2009; Lee et al., 2011; Spezio et al., 2007a; 2007b). Dans la population normale, des variations individuelles sont observées au niveau de l'habileté à reconnaître les expressions faciales (Besel & Yuille, 2010; Moeller, Ewing Lee, & Robinson, 2011; Suzuki, Hoshino, & Shigemasu, 2006; Young & Brunet, 2011), mais les stratégies d'extraction de l'information visuelle n'ont jamais été mises en relation avec ces variations. Cette problématique sera donc examinée dans le deuxième article de cette thèse.

² Des différences ont été observées au niveau des stratégies d'extraction de l'information visuelle en fonction de l'ethnie de l'observateur (Blais et al., 2008 ; Jack et al., 2009 ; 2011 ; 2012). Dans cette thèse, seuls des individus Caucasiens ont pris part aux expériences dans le but d'homogénéiser les groupes et les stratégies révélées. Ainsi, les références aux « observateurs humains » tout au long de la thèse signifient en fait « observateurs humains de culture caucasienne ». D'autres études devront être menées afin de mieux comprendre les stratégies utilisées par les autres cultures.

Les deux articles présentés dans cette thèse sont basés sur deux expériences utilisant la méthode des Bulles (voir section 1.3.3). La première expérience utilisait des expressions faciales statiques et les dimensions de l'étendue spatiale et des fréquences spatiales étaient échantillonnées. La deuxième expérience utilisait des expressions faciales dynamiques et les dimensions de l'étendue spatiale, des fréquences spatiales, et du temps étaient échantillonnées.

Chapitre 2

Article 1

THE EYES ARE NOT THE WINDOW TO BASIC EMOTIONS

Caroline Blais¹, Cynthia Roy¹, Daniel Fiset², Martin Arguin¹ & Frédéric
Gosselin¹

1. Département de psychologie, Université de Montréal.
2. Département de psychoéducation et de psychologie, Université du Québec en Outaouais

Address correspondence to

Frédéric Gosselin, Ph.D.

Article published in *Neuropsychologia* (2012), 50, 2830-2838.

2.1 Abstract

Facial expressions are one of the most important ways to communicate our emotional state. In popular culture and in the scientific literature on face processing, the eye area is often conceived as a very important—if not the most important—cue for the recognition of facial expressions. In support of this, an underutilization of the eye area is often observed in clinical populations with a deficit in the recognition of facial expressions of emotions. Here, we used the Bubbles technique to verify which facial cue is the most important when it comes to discriminating among eight static and dynamic facial expressions (i.e., six basic emotions, pain and a neutral expression). We found that the mouth area is the most important cue for both static and dynamic facial expressions. We conducted an ideal observer analysis on the static expressions and determined that the mouth area is the most informative. However, we found an underutilization of the eye area by human participants in comparison to the ideal observer. We then demonstrated that the mouth area contains the most discriminative motions across expressions. We propose that the greater utilization of the mouth area by the human participants might come from remnants of the strategy the brain has developed with dynamic stimuli, and/or from a strategy whereby the most informative area is prioritized due to the limited capacity of the visuo-cognitive system.

2.2 Introduction

Interactions with other individuals are an important part of our everyday life, and having the skills to adequately transmit our state of mind and decode that of others is crucial for the success of social communication. The facial expression of emotions is one of the most important ways of communicating those states (Mehrabian, 1968), and the study of the perceptual information available to decode facial expressions, as well as the actual information used by human observers to achieve this task, has preoccupied researchers at least since the publication of Darwin's seminal book *The Expression of Emotions in Man and Animals* (1872). It is now well accepted that the various facial expressions differ from one another in terms of where the information is available across the different facial areas (e.g., Bassili, 1979; Cunningham, Kleiner, Bülthoff, & Wallraven, 2004; Ekman, 1982; Hanawalt, 1944; Nummenmaa, 1964; Nusseck, Cunningham, Wallraven, Bülthoff, 2008; Plutchik, 1962; Smith et al., 2005). However, some areas of the face may convey more information than others when it comes to discriminating *all* emotions or a significant subset of them.

In popular culture, the eyes are often portrayed as the most important emotional cue. This is, for instance, reflected in our metaphorical language: "The eyes are the window to the soul", "I could see the fear in his eyes", "His eyes were filled with anger", "There was passion in her eyes", "Love showed in

his eyes”, “Her eyes welled with tears”, and so on. In fact, a considerable part of the literature that emerged from cognitive science research suggests that the eyes are particularly important for face recognition (e.g., Sekuler et al., 2004; Butler et al., 2010; Caldara et al., 2005; Schyns, Bonnar, & Gosselin, 2002). Research on infant development also suggests that the eye region is “special” from very early on in life. For example, infants show a preference for looking at the eye region in comparison to other facial areas (Hainline, 1978; Haith, Bergman, & Moore, 1977; Maurer, 1985) and prefer to look at faces with direct eye contact (Farroni, Csibra, Simion, & Johnson, 2002). The existence of an innate gaze module, dedicated to the task of detecting the presence of eyes, has even been proposed (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000).

Moreover, research on clinical populations has shown that the eye region is processed less efficiently or is processed in an abnormal way in many neurological pathologies leading to social impairments (e.g., Adolphs et al., 2005; Lee, Gosselin, Wynn, & Green, 2010; Spezio, Adolphs, Hurley, & Piven, 2007). One example of this is the finding that SM, a patient with a bilateral amygdala lesion suffering from a major deficit at categorizing the expression of fear, processes eye information less effectively than do control subjects (Adolphs et al., 2005). Furthermore, this patient's performance with the expression of fear returns to normal when she is instructed to look at the eyes. Similarly, schizophrenia patients rely less on the high spatial frequencies in the

eye region than control participants when categorizing fear (Lee, Gosselin, Wynn, & Green, 2010). Adults with autism, who show a deficit in the categorization of facial expressions (Humphreys, Minshew, Leonard, & Behrmann, 2007; Harms, Martin, & Wallace, 2010), have also been shown to process the eye region less efficiently than normal participants (Baron-Cohen et al., 1997; Spezio, Adolphs, Hurley, & Piven, 2007a; 2007b). Acquired prosopagnosic patients, who have been shown to process the eye region less efficiently than controls (Bukach, Bub, Gauthier, & Tarr, 2006; Bukach, LeGrand, Kaiser, Bub, & Tanaka, 2008; Caldara et al., 2005; Rossion, Kaiser, Bub, & Tanaka, 2009), also suffer from a deficit in discriminating facial expressions of emotions (Humphreys, Avidan, & Behrmann, 2007). Thus, many neuropsychological phenomena associated with a failure in facial expression categorization involve an underutilization of the eyes.

From the observations listed above, it is tempting to conclude that the eyes are more important than any other facial area for the discrimination of facial emotions. However, the studies that have directly addressed the question of which facial features are useful for the discrimination of basic facial expressions of emotions have led to contradictory results. Some have found that the lower part of the face was more important than the upper part of the face (Dunlap, 1927; Ruckmick, 1921) and some have found no greater importance of one part of the face over another (Baron-Cohen, Wheelwright, & Jolliffe, 1997; Coleman, 1949; Frois-Wittman, 1930). In a related vein, eye-tracking

studies that have examined how ocular fixations are distributed on faces during the recognition of facial expressions have found a roughly equal sampling of the mouth and eye areas (Eisenberg & Alpers, 2011; Jack et al., 2009). However, eye fixation patterns are partly dissociable from information use (Arizpe, Kravitz, Yovel, & Baker, 2012; Jonides, 1981; Posner, 1980; see however Rayner, 1998; Deubel & Schneider, 2003; Godijn & Theeuwes, 2003), and the question of interest here is what visual information is actually used to discriminate facial expressions from one another.

Our primary aim here is to discover which facial information is the most important when it comes to discriminating a significant subset of facial emotions. The data that will be analyzed in this paper is part of a larger project in which we examined many dimensions of the visual information extraction strategies employed for the discrimination of static and dynamic facial expression of emotions using the Bubbles technique (Gosselin & Schyns, 2001). Here, we will focus on the spatial and the temporal dimensions of the data. Our secondary aim is to examine if the use of information varies between static and dynamic stimuli. We will also present a novel analysis of the data from Smith et al. (2005), who have used static Bubbles to verify which facial information is the most important to discriminate *each* expression of basic emotion from one another.

2.3 Methods

2.3.1. Participants

Forty-one Caucasian participants (14 males; 24.2 years old on average) with normal or corrected-to-normal visual acuity took part in the experiment with static stimuli, and 59 different participants (30 males; 23.9 years old on average) took part in the experiment with dynamic stimuli. All procedures were carried out with the ethics approval of the Université de Montréal.

2.3.2. Materials and stimuli.

Stimuli were displayed on a calibrated high-resolution CRT monitor with a refresh rate of 60 Hz. The experimental program was written in Matlab, using functions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The stimulus width subtended 5.72 degrees of visual angle.

The stimuli were created using a validated database composed of highly recognizable pictures and videos of 10 Caucasian actors' faces spontaneously expressing one of eight states (i.e., anger, disgust, fear, happy, pain, sad, surprise, and a neutral expression; Roy et al., 2007). The database as well as normative data are available at the following address: <http://mapageweb.umontreal.ca/gosselif/STOIC.rar>. The videos consisted of 15 frames, displayed at a rate of 30 Hz, starting with a neutral expression that

naturally deployed into one of the eight expressions mentioned above and ending with the apex of the expression. They lasted a total duration of 500 ms. The static stimuli consisted of the apex of the facial expressions mentioned above, displayed for a duration of 500 ms. All the stimuli were gray-scaled and their luminance was normalized. To minimize the head movements that occurred while the actors made the facial expressions, the stimuli were also spatially aligned frame by frame to ensure that the eyes and nose were located at about the same spatial coordinates across frames and stimuli.

To reveal the visual information useful for the discrimination of facial expressions, we used the Bubbles technique. The Bubbles technique consists of randomly sampling the visual information contained in a stimulus, such that, on each trial, a different subset of this information is rendered available to the participant. The performance of the participant with these subsets of information indicates which parts of the stimulus are most useful in performing the task. Here, we sampled the static expressions on the space (i.e., x , y coordinates of the face) and on the spatial frequency dimensions; and we sampled the dynamic expressions on the space, spatial frequency and time dimensions.

For each trial, the creation of a bubbled stimulus went as follows: first, the image of a facial expression was decomposed into five spatial frequency bands (128-64, 64-32, 32-16, 16-8, 8-4 cycles/image, or 86-43, 43-21.5, 21.5-

10.8, 10.8-5.4, 5.4-2.7 cycles/face; the remaining low frequency bandwidth served as a constant background; see Fig. 1a, top row) using the Laplacian pyramid (Burt & Adelson, 1983). With dynamic stimuli, the spatial frequency decomposition was performed on each frame of the videos (see Fig. 1b, top row, for an example with the third spatial frequency band). Then, independently for each spatial frequency band, the bubbles' locations (i.e., a bubble is a Gaussian aperture through which the information is visible) were randomly selected (see Fig. 1a and 1b, middle row). On the space dimension, the size of the bubbles (FWHM : 14.1, 28.3, 56.5, 113.0, and 226.1 pixels) was adjusted as a function of the frequency band so that each bubble revealed 1.5 cycles of spatial information. Because the size of the bubbles increased as the spatial scale became coarser, the number of bubbles differed across scales to keep the size of the sampled area constant across frequency bands. The size of the bubbles also varied as a function of spatial frequency on the time dimension (i.e., with dynamic bubbles), such that their duration increased as the spatial frequency band increased (FWHM: 14.1, 28.3, 56.5, 113.0, and 226.1 pixels on the space dimension and 7.3, 6.1, 5.1, 4.2 and 3.5 frames on the time dimension). This was done to take into account the faster processing of lower spatial frequencies (Hughes, Fendrich, Reuter-Lorenz, 1990; Parker, Lishman, & Hughues, 1992). A pointwise multiplication was then performed between the bubbles' masks and the filtered images (see Fig. 1a and 1b, bottom row). Finally, the information revealed by the bubbles was fused across the five frequency bands to produce an experimental stimulus (Fig. 1a).

Insert Fig. 1 near here

2.3.3. Procedure

Each participant completed 4000 trials divided into experimental sessions comprising 160 trials each. On each trial, the sequence of events went as follows: a fixation point was first displayed in the center of the screen for 200 ms and was immediately replaced by the stimulus (i.e., a bubbled image or video of a facial expression). The stimulus was displayed for 500 ms and was then replaced by a homogenous grey screen that remained visible until the participant responded. The participant was instructed to press on the keyboard key that corresponded to the facial expression he had perceived. Responses were not restricted by time pressure. No accuracy feedback was provided. The accuracy was maintained at 56% correct on average (i.e., halfway between chance and perfect performance) across all expressions by adjusting the total number of bubbles on the stimulus on a trial-by-trial basis using QUEST (Watson & Pelli, 1983). We used a constant number of bubbles across expressions because we did not want this parameter to become a cue for the recognition. A threshold of 56% correct (midway between chance and perfect performance) was chosen to make sure that the performance would reach neither ceiling for facial expressions that are easier to recognize, such as happy, nor floor for facial expressions that are more difficult to recognize, such as fear.

2.4. Results

2.4.1. *Classification plane and classification volume*

A mean of 144.3 bubbles (SD: 119.6) and of 241.3 bubbles (SD: 253.6) were necessary to maintain the average performance at 56% for the static and the dynamic version of the experiment respectively.

The visual information useful in categorizing facial expressions of emotions was determined using an analysis procedure that amounts to a multiple linear regression on the bubbles masks (explanatory variables) and on the participant's response accuracy (predictor variable). In other words, for each participant, each facial expression, and each spatial frequency band, a weighted sum of all the bubbles centers was calculated, using the accuracies transformed into z-score values as weights. This resulted in 3D volumes (or 4D volumes in the dynamic version of the experiment) of regression coefficients that will be referred to as classification volumes. These classification volumes were then summed across participants, leading to one classification volume per expression and per spatial frequency. These classification volumes were transformed into z-scores. We used the voxels (i.e., volume elements) outside the face area to calculate the mean and standard deviation of the distribution of the null hypothesis. Since we were particularly interested in verifying the location of the most useful visual information across all facial expressions in a categorization

task, we then summed the classification volumes across the eight expressions tested and across the five spatial frequency bands, and we normalized the resulting classification volumes by dividing them by the square root of 40 (i.e., 8 expressions x 5 spatial frequency bands). We then smoothed the 2D classification volume (or classification planes) using a Gaussian window with a FWHM of 28.3 pixels (equivalent to the spatial extent of the bubbles that revealed the information at second finest scale) and the 3D classification volume using a Gaussian window with a FWHM of 28.3 pixels on the spatial dimension and of 6.1 frames on the temporal dimension (equivalent to the temporal extent of the bubbles that revealed the information at second finest scale); and we transformed one last time the pixels (or voxels) of the classification plane (or volume) into z-scores, using the pixels (or voxels) outside the face area to calculate the mean and standard deviation of the distribution of the null hypothesis. In order to determine if the facial information significantly correlated with accuracy, we applied the *Pixel* test ($p < 0.05$, $Z_{crit} = 3.76$ and 4.31 for the classification plane and classification volume, respectively) to the classification plane and volume. The statistical threshold provided by this test corrects for multiple comparisons while taking the spatial correlation inherent to structured images into account (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005). The classification plane and classification volume are displayed in Fig. 2a and 2b respectively. The non-significant pixels are depicted in gray, and the different colors indicate the z-score values of the significant pixels (and voxels).

Insert Fig. 2 near here

A quick visual inspection of the classification plane and volume reveals that the eye and mouth regions are the most important facial areas. To further characterize the information available in the classification plane and classification volume in terms of facial features, we conducted a region-of-interest (i.e., ROI) analysis on six facial areas (i.e., the eyes, the eyebrows, the frown lines, the nose, the nasolabial folds, and the mouth; see inset in Fig. 4). In the first part of this analysis, we were interested in verifying the relative importance of the different face areas and in comparing the result obtained for dynamic and static stimuli. We therefore collapsed the classification volume on the temporal dimension.

2.4.2. ROI analysis without the time dimension.

We kept only the portion of the static and dynamic “classification planes” that corresponded to the highest 5% regression coefficients. This ensured that the same number of pixels was considered for the ROI analysis on the data from the static and the dynamic version of the experiment. We then calculated, separately for the static and the dynamic classification planes, the proportion of the total number of these pixels that fall on each facial feature, and divided this proportion by the total number of pixels in that feature, thus

normalizing the proportion for feature size (e.g., Gibson et al., 2007). The results of this analysis are summarized in Fig. 3. For both static and dynamic facial expressions of emotions, the mouth area is more important than the eye area (i.e., 8.56 and 15.51 times more important on average for the static and the dynamic stimuli, respectively). To make sure that the average classification planes reflected the strategy of most participants rather than only a few participants, we created 1,000 classification planes using random subsamples of 20 participants (results are robust to changes in the size of this subsample of participants). We calculated the average of the z-score values for the mouth area and for the eye area, and we calculated the ratio of these two values (mouth / eyes). The ratios were higher than one (i.e., mouth > eyes) on 99.6% of the dynamic classification planes and on 100% of the static classification planes. We also verified that this result was not an artifact of the smoothing applied on our classification planes by performing the same analysis on unsmoothed classification planes. The mouth remained more important than the eye area on 96.7% of the dynamic classification planes and on 100% of the static classification planes.

Insert Fig. 3 near here

2.4.3. ROI analysis with the time dimension.

We then examined in what order the different facial features became useful for the categorization of dynamic facial expressions. We calculated for each frame

of the 3D classification volume, the proportion of the total number of significant pixels (i.e., as determined with the *Pixel* test) that fell on each facial feature across the frames and divided this proportion by the total number of pixels in that feature. The relative importance of each facial feature across time is displayed in Fig. 4. The mouth is the first area to become useful for the discrimination of facial expressions, around 100 ms after stimulus onset. The left eye and the eyebrows are the second areas to become useful (around 234 ms after stimulus onset), followed by the frown lines and the nasolabial folds (around 368 ms after stimulus onset). Except for a short period of time between 268 and 333 ms, the mouth remains the most useful area throughout the stimulus presentation. We also looked at the relative importance of the mouth and the “metropolitan” eye area by grouping the eyes, the frown lines, and the eyebrows (see the dotted black curve in Fig. 4). Again, except for a short period of time around 300 ms after stimulus onset, the mouth remains the most useful area throughout stimulus duration.

Insert Figure 4 near here

Thus, in both ROI analyses, the mouth is the most useful area when it comes to discriminating facial expressions of emotions. It is conceivable that the mouth was particularly informative in the set of stimuli we used, or that the parameters used in our experiments affected the participants’ strategy, and that our results overestimate the importance of the mouth area. To test this, we reanalyzed the

data of Smith et al. (2005) with the same procedure as described above. These authors also applied the Bubbles method (i.e., the same version we used in our study with the static stimuli) in a discrimination task of facial expressions of emotions. However, they used a completely different set of stimuli (the California Facial Expressions—i.e., CAFE—database; Dailey, Cottrell, & Reilly, 2001) with a slightly different subset of facial expressions (i.e., they did not use the pain expression). Moreover, they adjusted the performance differently from what we did. They manipulated the number of bubbles such that accuracy was approximately equal (i.e., accuracy threshold of 75%) across the seven facial expressions. Here, we decided to allow accuracy to vary across the expressions to prevent the number of bubbles from becoming a cue for discriminating between facial expressions of emotions (i.e., the average accuracy was not controlled separately for each expression, but was instead controlled across expressions; the accuracy threshold was of 56%). Another difference between the experiment of Smith et al. (2005) and ours was the stimulus duration. They displayed the stimuli until the participant's response, whereas we displayed the stimuli for 500 ms. Because the aim of Smith et al. (2005) was to verify which information was useful in categorizing each facial expression, they analyzed each expression and spatial scale independently. Here, we re-analyzed their data to reveal the information that was useful across all expressions irrespective of spatial scale.

2.4.4 Reanalysis of Smith et al. (2005)

Similarly to our study, their experiment provided one classification plane per expression and per frequency band. We therefore added the five frequency bands and the seven expressions of their classification planes, we smoothed the resulting classification plane using a Gaussian window with a FWHM of 28.3 pixels, and we z-scored it using the pixels outside the face area to calculate the mean and standard deviation of the distribution of the null hypothesis. The resulting classification plane is presented in Fig. 5. The pixels with the highest regression coefficients (i.e., we kept the same number of pixels as displayed in Fig. 2a) are displayed in colors. These results, which are quite consistent with our own, clearly show that the mouth area is favored over the eye area during the categorization of facial expressions.

Insert Fig. 5 near here

2.5. Discussion

We examined what visual information is most useful in discriminating the basic facial emotions and pain from one another using the *Bubbles* technique. We observed a clear preference for the mouth for both static and dynamic stimuli. The results obtained for the discrimination of dynamic facial expressions also show that the mouth remains the most useful area throughout

the stimulus duration. We also replicated our finding in a re-analysis of the data from Smith and colleagues (2005) despite very different experimental parameters: the stimulus set used, the variance of performance across expressions (i.e., constant in the experiment from Smith and colleagues, varying in our experiments), and the stimulus duration.

Our findings may appear inconsistent with the eye movement literature, which shows a roughly equal sampling of the eye and mouth areas (e.g., Eisenberg & Alpers, 2011; Jack et al., 2009), but they need not be. There is a partial dissociation between where the eye fixations land and what visual information is actually processed (Arizpe et al., 2012; Jonides, 1981; Posner, 1980; see however Rayner, 1998; Deubel & Schneider, 2003; Godijn & Theeuwes, 2003). The eyes are smaller than the mouth, and are therefore represented by higher spatial frequencies. They may thus need to be processed within the fovea, since the high density of cones found in this area of the retina makes it more suitable for processing high spatial frequencies. The mouth, which is represented by lower spatial frequencies, may be adequately processed in parafoveal regions and may thus also be processed while the eye fixations land close to the eye area. In other words, when the eye fixations fall on the eye area, both the eye and the mouth area may be processed, whereas when the eye fixations fall on the mouth area, only the mouth area may be processed. This could explain why the proportion of fixations falling on the eye and on the

mouth area is similar even if the mouth area is more useful than the eye area for the recognition of facial expressions.

Why is the mouth area more important than any other facial area for the accurate categorization of facial expressions of basic emotions? One conceivable explanation for this finding is that the mouth is the most informative area of the face (i.e., it contains more signal). To verify this possibility, we submitted an *ideal observer*—a model observer that uses all the available information optimally—to the same static facial discrimination task as our human observers (e.g., see Smith et al., 2005). The classification plane of the model observer reveals which areas of the face are informative in discriminating the expressions from one another.

2.5.1 Ideal observer analysis on static facial expressions.

On each trial, the model observer was presented with a stimulus of the experiment that we conducted with the human observers. The same mask of bubbles was applied to the face and Gaussian white noise was added to the stimulus in order to keep the average accuracy at the same level as the one used with the humans (i.e., 56%). The amount of noise was adjusted on a trial-by-trial basis using QUEST (Watson & Pelli, 1983). The same mask of bubbles was also applied to all the other faces of the stimulus set, and the ideal observer calculated the correlation between the target stimulus presented and every other face. The facial expression of the face that had the highest correlation with the target stimulus was the model's response. We then computed the classification

plane of the ideal observer using the same procedure as explained in the Results section (see Fig. 6).

Insert Fig. 6 near here

The model observer mostly used the mouth and eye areas, confirming that these face areas are the most informative when it comes to discriminating the expressions included in our study from one another. Most importantly, the ideal observer shows that the mouth area contains more information than the eye area (see also Fig. 3). This may explain at least in part why this area was used most by human participants. However, a more rigorous analysis of the similarities and differences in the visual extraction strategies of human observers and of the model observer reveals that the relative utilization of the mouth and the eyes is different for the ideal observer and for the human participants. Indeed, the ratio of the proportion of diagnostic pixels (i.e., top 5% pixels) that fell on the mouth vs. on the eye area was much greater for human participants (i.e., 8.56 and 15.51 on average for the static and dynamic stimuli respectively) than for the ideal observer (i.e., 3.08). To test if the ratio of the mouth vs. eyes utilization was statistically significantly higher for the humans than for the ideal observer, we created 1,000 classification planes using random subsamples of 20 participants (results are robust to changes in the size of this subsample of participants), and calculated the ratios of the average of the z-scores in the mouth area and in the eye area. We compared the ratio found in

each of these classification planes to the one found for the ideal observer, and performed a sign test on the result of this comparison. This analysis indicated a significant difference between the ratio of the mouth and eyes utilization for the humans and for the ideal observer ($p < 0.001$ for both static and dynamic stimuli). Again, we made sure that this result was not an artifact of the smoothing of our classification planes by repeating the analysis on unsmoothed classification planes: A significant difference was again found between the ratio of the mouth and eyes utilization for the humans and for the ideal observer ($p < 0.001$). Thus, pixel-wise, informativity does not account entirely for the human preference for the mouth. What else then could explain this preference?

One possibility comes from the inherently dynamic nature of facial expressions. It may be the case that the movements of the mouth contain much more information for the discrimination of natural facial expressions than the movements of any other facial area. Moreover, the human brain may have learned to use these motion cues to discriminate facial expressions and remnants of this strategy may influence how humans recognize static facial expressions. To test the first part of this hypothesis, we measured how the amplitude of the movements in different areas of the expressive face varies across our dynamic stimuli.

2.5.2. Motion analysis on the dynamic facial expressions.

We first calculated, using a three-step search method (Koga, Iinuma, Hirano, Iijima, & Ishiguro, 1981) with a spatial granularity of 10 x 10 pixels, the surface-based motion occurring between each step of two frames in our dynamic faces (the results of this analysis are robust to parameter changes). This resulted in one motion vector for each 10 x 10 area of our stimuli. We then calculated the amplitude of those vectors and, finally, we calculated the variance of these amplitudes for each area across all the stimuli and averaged these values across all frames (see Fig. 7). The variance values were transformed into z-scores using the average and the standard deviations of the variances across all the facial areas. The more variance there is in an area across all stimuli, the more this area gives information about the expression portrayed. It is clear, from this analysis, that the mouth area is by far the most informative for the categorization of all the expressions.

Insert Fig. 7 near here

Therefore, the greater utilization of the mouth area in comparison to the eye area by the human observers could be explained in part by the mouth area conveying most of the movement information in dynamic stimuli. As proposed above, it is possible that the brain has elaborated a strategy over the course of its development that is suitable for dynamic expressions, since they appear in an ecological environment. Whenever the brain is exposed to a facial expression

categorization task, a similar strategy is used no matter the state—static or dynamic—of the expressions. We are not suggesting that facial expression recognition relies only on motion cues, but that the importance of movement for the memory representations of facial expressions of emotions should not be underestimated. If this hypothesis is true, the facial expressions would benefit from being represented by neurons that integrate both the movement and the shape of an object. Interestingly, the superior temporal sulcus (STS), a region of the cortex that has been proposed as a site of integration for these two kinds of information (Vaina, Solomon, Chowdhury, Sinha & Belliveau, 2001), is involved in the processing of both static (Furl, van Rijsbergen, Treves, Friston, & Dolan, 2007; Haxby, Hoffman, & Gobbini, 2000; Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001; Pessoa & Padmala, 2007; Tsuchiya, Kawasaki, Oya, Howard, & Adolphs, 2008) and dynamic facial expressions (Ishai, 2008; Said, Moore, Engell, Todorov, & Haxby, 2010).

2.5.3 A strategy for a system with a limited capacity

Another potential explanation for the greater utilization of the mouth area by the human observers in comparison to the ideal observer may be the limited capacity of the human visual system (Levin & Simons, 1997; Simons & Rensink, 2005). Indeed, with limited capacity, a strategy whereby the most informative area is favored at the expense of other areas may be selected. The more resources are available, the more the other areas—for instance the second

most informative area, the eyes—receive processing. This could also explain, at least in part, why an underutilization of the eye area has often been reported in clinical populations that show a deficit in facial expression recognition. Patients suffering from a brain lesion or from brain dysfunction related to facial expression recognition most likely have less visual resources to devote to facial expression discrimination than does the healthy population. They could have just enough resources to process the mouth but not the mouth and the eyes. Since the mouth is very informative, the patients are capable of performing the task. However, since the eyes also convey crucial information, they are impaired compared to healthy individuals. This last proposition is congruent with the relative difficulty of revealing a deficit in basic emotion recognition in the autistic population (Adolphs, Sears, & Piven, 2001; Baron-Cohen et al., 1997; Grossman, Klin, Carter, & Volkmar, 2000; Ogaï et al., 2003; Ozonoff, Pennington, & Rogers, 1990; Prior, Dahlstrom, & Squires, 1990; Spezio, Adolphs, Hurley, & Piven, 2007; Teunisse and de Gelder, 1994; Volkmar, Sparrow, Rende, & Cohen, 1989) and, therefore, the need to use very sensitive tasks—for instance the facial expression megamix (Humphreys, Minschew, Leonard, & Behrmann, 2007). Interestingly, this population has been shown to underutilize the eye area compared to healthy individuals and to rely more on the mouth area during the processing of the facial expression of basic emotions (Spezio, Adolphs, Hurley, & Piven, 2007a; 2007b).

Of course, the two explanations proposed above, in sections 2.5.2 and 2.5.3, are speculative and more research will be needed to understand why the mouth area is so important for the recognition of the facial expression of basic emotions. It will also be important to use techniques other than Bubbles because each technique employed to probe the use of visual information may interact with this use of information.

2.5.5 Conclusion

Even if our results show that humans use the mouth area more than the eye area to discriminate the basic facial expressions from one another, the importance of the eye area should not be underestimated. Our participants effectively used the eye area, though less so than the mouth area. In fact, the eye area is the most important visual cue for the recognition of fear (Adolph et al., 2005; Smith et al., 2005; Gosselin, Spezio, Tranel, & Adolphs, 2011). Tasks using composite facial expressions (e.g., smiling mouth with angry eyes) show that the top and the bottom parts of facial expressions interact to create the final percept of the facial emotion. This is consistent with our finding that both the eye and the mouth areas are useful for the recognition of facial expressions. Moreover, research suggests that the eye area becomes more important when recognizing complex mental states (Baron-Cohen et al., 1997; Baron-Cohen, Wheelwright, & Jolliffe, 2010). Thus, our aim here is not to negate the importance of the eye area in the field of Social Neuroscience but, rather, to

rehabilitate the importance of the mouth area for the recognition of facial expressions.

2.6 Acknowledgments

We would like to thank Marie Smith for kindly giving us access to her data. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (CRSNG) to Frédéric Gosselin and by a scholarship from the Fonds Québécois de Recherche sur la Nature et les Technologies (FQRNT) to Caroline Blais.

2.7. References

- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P. G. & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68-72
- Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience*, *13*, 232–240.
- Arizpe, J., Kravitz, D. J., Yovel, G., & Baker, C. I. (2012). Start position strongly influences fixation patterns during face processing: Difficulties with eye movements as a measure of information use. *PLoS ONE*, *7*, e31106. doi:10.1371/journal.pone.0031106.
- Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes" ? Evidence from normal adults and adults with autism or Asperger syndrome. *Visual Cognition*, *4*, 311-331
- Bassili, J. N. (1979). Emotion Recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, *37*, 2049-2058.
- Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, *23*, 223-229.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433-436.

- Bukach, C. M., Bub, D. N., Gauthier, I., Tarr, M. J. (2006). Perceptual expertise effects are not all or none: Spatially limited perceptual expertise for faces in a case of prosopagnosia *Journal of Cognitive Neuroscience*, 18, 48–63.
- Bukach, C. M., LeGrand, R., Kaiser, M., Bub, D. N., & Tanaka, J. W. (2008). Preservation of mouth region processing in two cases of prosopagnosia. *Journal of Neuropsychology*, 2, 227-244.
- Burt, P., & Adelson, E. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31, 532–540.
- Caldara, R., Rossion, B., Mayer, E., Smith, M., Gosselin, F. & Schyns, P. G. (2005). Does prosopagnosia take the eyes out from faces? Evidence for a defect in the use of diagnostic facial information in a single-case. *Journal of Cognitive Neuroscience*, 17, 1-15.
- Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, 5, 659-667
- Coleman, J. (1949). Facial expressions of emotion. *Psychological Monographs*, 63, 1–296.
- Cunningham, D. W., Kleiner, M., Bühlhoff, H. H., & Wallraven, C. (2004). The components of conversational facial expressions. Proceedings of the 1st Symposium on Applied perception in graphics and visualization, doi:10.1145/1012551.1012578.
- Dailey M., Cottrell G.W., Reilly J. (2001). California facial expressions, CAFE.

Unpublished digital images, University of California, San Diego, Computer Science and Engineering Department.

Darwin, C. (1965). *The expression of the emotions in man and animals*.

Chicago: University of Chicago Press. (Original work published 1872).

Deubel, H. & Schneider, W. X. (2003). Delayed saccades, but not delayed manual aiming movements, require visual attention shifts. *Annals of the New York Academy of Sciences*, 1004, 289-296.

Dunlap, K. (1927). The role of the eye muscles and mouthmuscles in the expression of emotions. *Genetic Psychology Monographs*, 2, 199–233

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes : Scanning emotional facial expressions. *Emotion*, 11, 860-865.

Ekman, P. (1982). *Emotion in the human face*, second edition. Cambridge, UK: Cambridge University Press.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, 99, 9602-9605.

Frois-Wittman, J. (1930). The judgement of facial expression. *Journal of Experimental Psychology*, 13, 113–151

Furl, N., van Rijsbergen, N. J., Treves, A., Friston, K. L. & Dolan, R. J. (2007). Experience-dependant coding of facial expression in superior temporal sulcus. *Proceedings of the National Academy of Science*, 104, 13485-13489.

- Gibson, B., Lazareva, O. F., Gosselin, F., Schyns, P. G. & Wasserman, E. A. (2007). Non-accidental properties underlie shape recognition in mammalian and non-mammalian vision. *Current Biology*, *17*, 336-340
- Godjin, R. & Theeuwes, J. (2003). Parallel allocation of attention prior to the execution of saccade sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 882-896.
- Gosselin, F. & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition. *Vision Research*, *41*, 2261-2271.
- Gosselin, F., Spezio, M.L., Tranel, D. & Adolphs, R. (2011). Asymmetrical use of eye information from faces following unilateral amygdala damage. *Social Cognitive and Affective Neuroscience*, *6*, 330-337.
- Grossman, J. B., Klin, A., Carter, A. S., & Volkmar, F. R. (2000). Verbal bias in recognition of facial emotions in children with Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *41*, 213–225.
- Hainline, L. (1978). Developmental changes in visual scanning of face and non-face patterns by infants. *Journal of Experimental Child Psychology*, *25*, 90-115.
- Haith, M., Bergman, T., & Moore, M. (1977). Eye contact and face scanning in early infancy. *Science*, *198*, 853-855.
- Hanawalt, N. (1944). The role of the upper and lower parts of the face as the basis for judging facial expressions: II. In posed expressions and “candid camera” pictures. *Journal of General Psychology*, *31*, 23–36.

- Harms, M., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychological Review, 20*, 290-322.
- Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed neural human system for face perception. *Trends in Cognitive Science, 4*, 223-233.
- Hughes, H. C., Fendrich, R., & Reuter-Lorenz, P. A. (1990). Global versus local processing in the absence of low spatial frequencies. *Journal of Cognitive Neuroscience, 2*, 272-282.
- Humphreys, K., Avidan, G. and Behrmann, M. (2007). A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. *Experimental Brain Research, 176*, 2, 356-373.
- Humphreys, K., Minschew, N., Leonard, G. L., & Behrmann, M. (2007). A fine-grained analysis of facial expression processing in high-functioning adults with autism. *Neuropsychologia, 45*, 685-695.
- Ishai, A. (2008). Let's face it : It's a cortical network. *NeuroImage, 40*, 415-419.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P., & Caldara, R. (2009). Cultural Confusions Show That Facial Expressions are Not Universal. *Current Biology, 19*, 1543-1548.
- Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. L. A. Baddeley (Ed.), *Attention and Performance IX*. Hillsdale, NJ: Erlbaum.

- Koga, T., Linuma, K., Hirano, A., Iijima, Y., & Ishiguro, T. (1981). Motion Compensated Interframe Coding for Video Conferencing. *Proc. Nat. Telcommun. Conf., New Orleans*, 5.3.1-5.3.5
- Lee, J., Gosselin, F., Wynn, J. K., & Green, M. F. (2011). How do schizophrenia patients use visual information for decoding facial emotion? *Schizophrenia Bulletin*, 37, 1001-1008
- Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, 4, 501-506
- Maurer, D. (1985). Infants' perception of facedness. In Field, T. & Fox, M. (Eds). *Social perception in infants*, Norwood, NJ: Ablex
- Mehrabian, A. (1968). Communication without Words. *Psychology Today*, 2, 53-56.
- Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Cognitive Brain Research*, 12, 225-231.
- Nummenmaa, T. (1964). *The language of the face* (Jyvaskyla studies in education, psychology, and social research). Jyvaskyla, Finland.
- Nusseck, M., Cunningham, D. W., Wallraven, C., & Bülthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision*, 8, 1-23
- Ogai, M., Matsumoto, H., Suzuki, K., Ozawa, F., Fukuda, R., Uchiyama, I., Suckling, J., Isoda, H., Mori, N., & Takei, N. (2003). fMRI study of recognition of facial expressions in high-functioning autistic patients.

Neuroreport, 14, 559–563.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1990). Are there emotion perception deficits in young autistic children? *Journal of Child Psychology and Psychiatry*, 31, 343–361.

Pessoa, L. & Padmala, S. (2007). Decoding near-threshold perception of fear from distributed single-trial brain activation. *Cerebral Cortex*, 17, 691-701.

Parker, D. M., Lishman, J. R., Hughues, J. (1992). Temporal integration of spatially filtered visual images. *Perception*, 21, 147-160.

Pelli, D. G. (1997). The Video Toolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10, 437-442.

Plutchik, R. (1962). *The emotions : facts, theories, and a new model*. New York: Random House.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.

Prior, M., Dahlstrom, B., & Squires, T. (1990). Autistic children's knowledge of thinking and feeling states in other people. *Journal of Child Psychology and Psychiatry*, 31, 587–601.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.

Rossion, B., Kaiser, M.D., Bub, D., & Tanaka, J.W. (2009). Is the loss of diagnosticity of the eye region of the face a common aspect of acquired prosopagnosia? *Journal of Neuropsychology*, 3, 69-78.

- Roy, S., C. Roy, Fortin, Ethier-Majcher, Belin & Gosselin (2007). A dynamic facial expression database. *Journal of Vision*, 7, 944.
- Ruckmick, C. A. (1921). A preliminary study of the emotions. *Psychological Monographs*, 30, 30-35.
- Said, C. P., Moore, C. D., Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision*, 10, 1-12.
- Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present and future. *Trends in Cognitive Sciences*, 9, 16-20
- Smith, M., Cottrell, G., Gosselin, F. & Schyns, P. G. (2005). Transmitting and decoding facial expressions of emotions. *Psychological Science*, 16, 184-189
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007a). Abnormal use of facial information in high-functioning autism. *Journal of Autism and Developmental Disorders*, 37, 929–939.
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007b). Analysis of face gaze in autism using “Bubbles. *Neuropsychologia*, 45, 144–151.
- Teunisse, J.-P., & de Gelder, B. (1994). Do autistic children have a generalized face processing deficit? *International Journal of Neuroscience*, 77, 1–10.
- Tsuchiya, N., Kawasaki, H., Oya, H., Howard, M. A., & Adolphs, R. (2008). Decoding face information in time, frequency and space from direct intracranial recordings of the human brain. *PLoS ONE*, 3(12): e3892. doi:10.1371/journal.pone.0003892

- Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Beliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Science*, *98*, 11656-11661.
- Volkmar, F. R., Sparrow, S. S., Rende, R. D., & Cohen, D. J. (1989). Facial perception in autism. *Journal of Child Psychology and Psychiatry*, *30*, 591-598.
- Watson, A. B. & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113-20.

2.8. Figure Legends

Figure 1. Illustration of the creation of a bubbled stimulus. (a) Procedure with static stimuli. A facial expression image was decomposed into five spatial frequency bands (top row). Bubbles were then placed at random locations separately for each spatial frequency band (middle row). The information revealed by the bubbles (bottom row) was then fused across the five frequency bands to produce an experimental stimulus. The column delimited by the red rectangle indicates the spatial frequency band used on Fig. 1b to illustrate the procedure with dynamic stimuli (b) With dynamic stimuli, the procedure to create a bubbled stimulus was very similar to the procedure for static stimuli, except that the decomposition into five spatial frequency bands was performed on each frame. Here, we illustrate the procedure for the third frequency band and the first 10 frames of the video. The same procedure was repeated on each frequency band. Each frame of the video was decomposed into five frequency bands (top row). Bubbles were then placed randomly at different locations and frames, separately for each spatial frequency band (middle row). The bottom row illustrates the information revealed by the bubbles on each frame.

Figure 2. Classification plane obtained with the static stimuli (a) and classification volume obtained with the dynamic stimuli (b). The areas depicted in color were significantly correlated with accuracy.

Figure 3. ROI analysis calculated on the classification plane (i.e., static stimuli) and on the classification volume (i.e., dynamic stimuli). The time frames of the classification volume were collapsed in order to compare the relative importance of each area for the static and dynamic stimuli. The figure indicates the probability that a significant pixel falls on each ROI.

Figure 4. ROI analysis across time on the classification volume. The figure indicates, for each time frame, the proportion of significant pixels on each ROI.

Figure 5. Classification plane obtained with the data of Smith et al. (2005). The pixels depicted in color were those with the highest z-score values (i.e., we display the same number of pixels that used for the ROI analysis on the static version of our study).

Figure 6. Classification plane obtained by the ideal observer with the static stimuli. The areas depicted in color were significantly correlated with accuracy.

Figure 7. Relative variance of the amplitude of motion in different areas of a face across all the dynamic stimuli tested (i.e., 80 stimuli: 8 expressions x 10 identities). The pixels depicted in color were the 5% pixels with the highest z-score values.

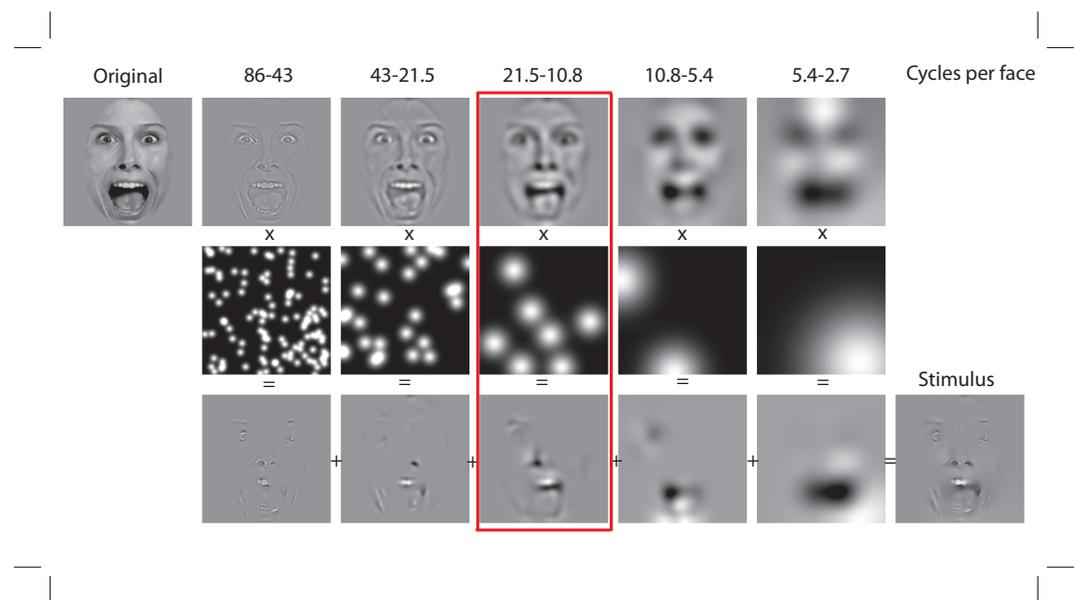


Figure 1a

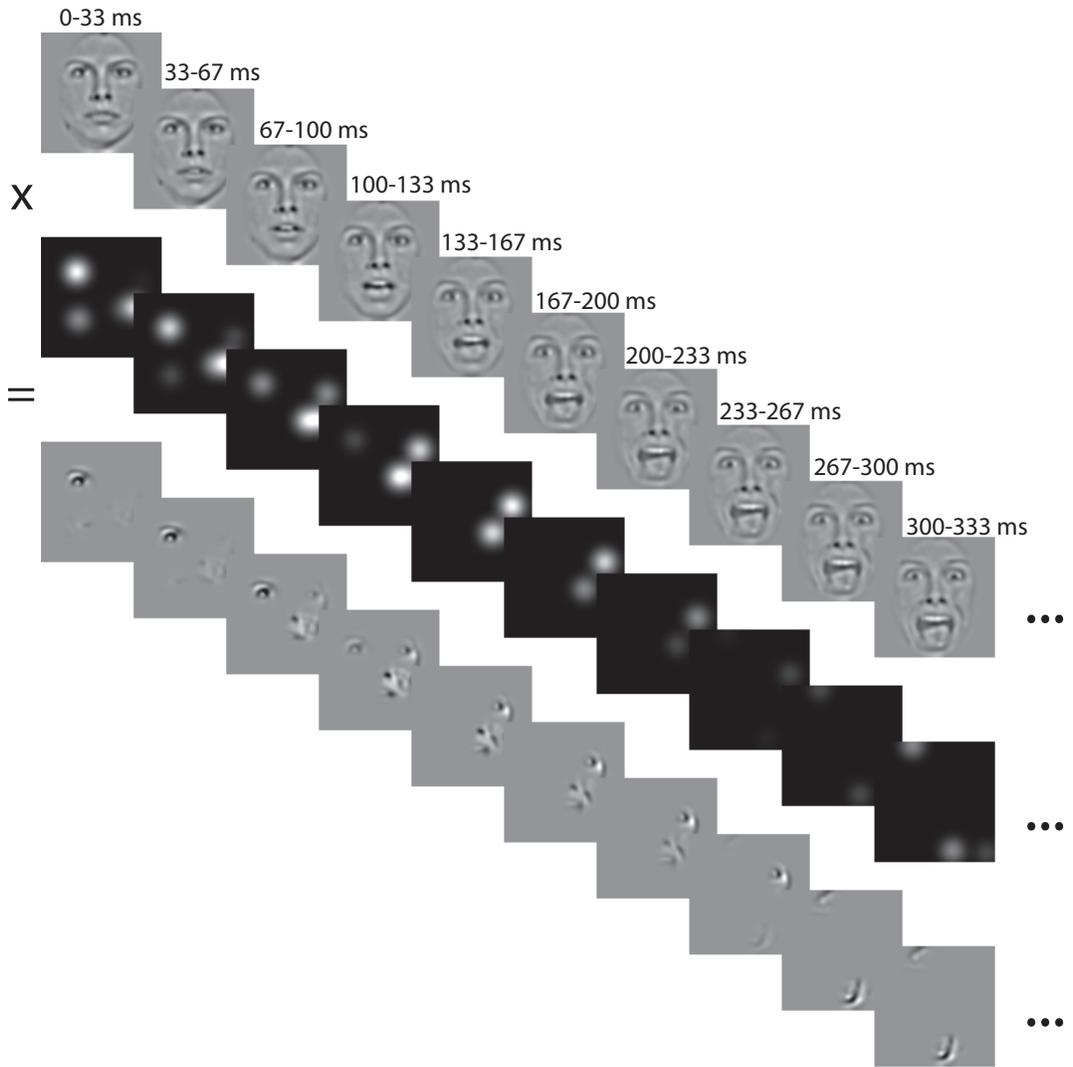


Figure 1b

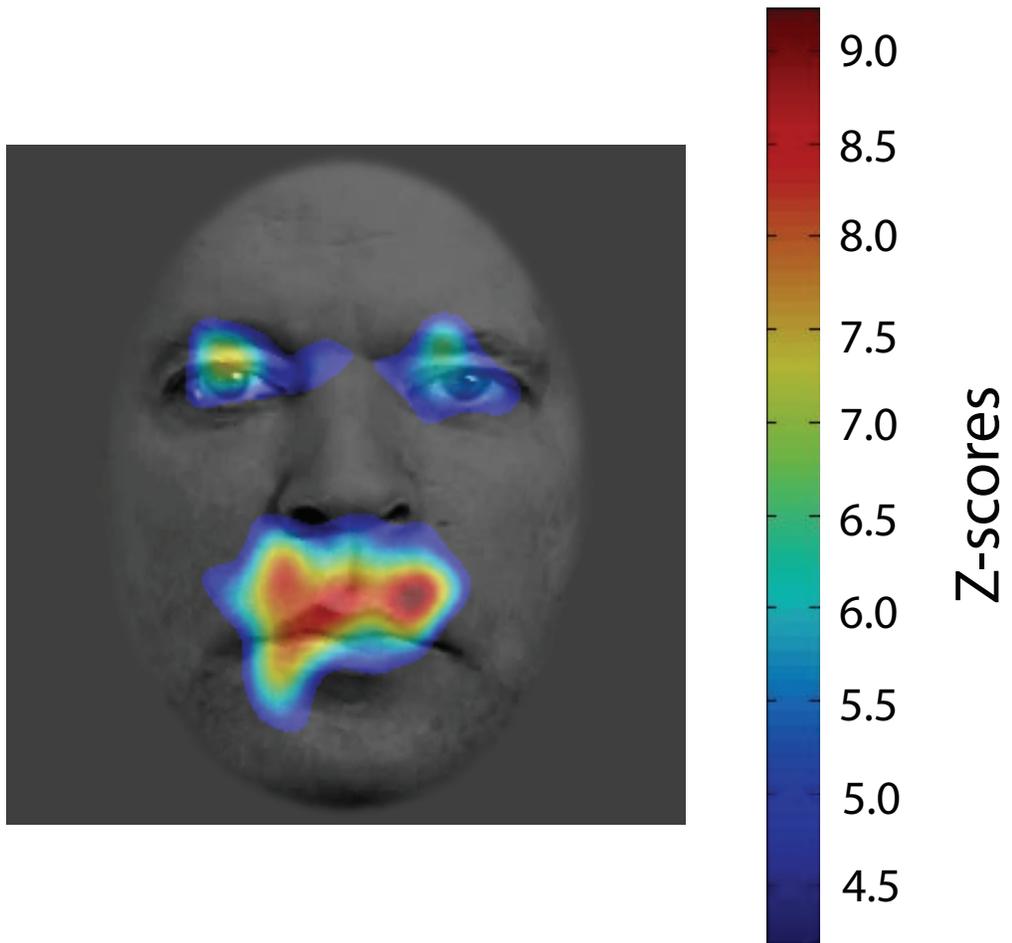


Figure 2a

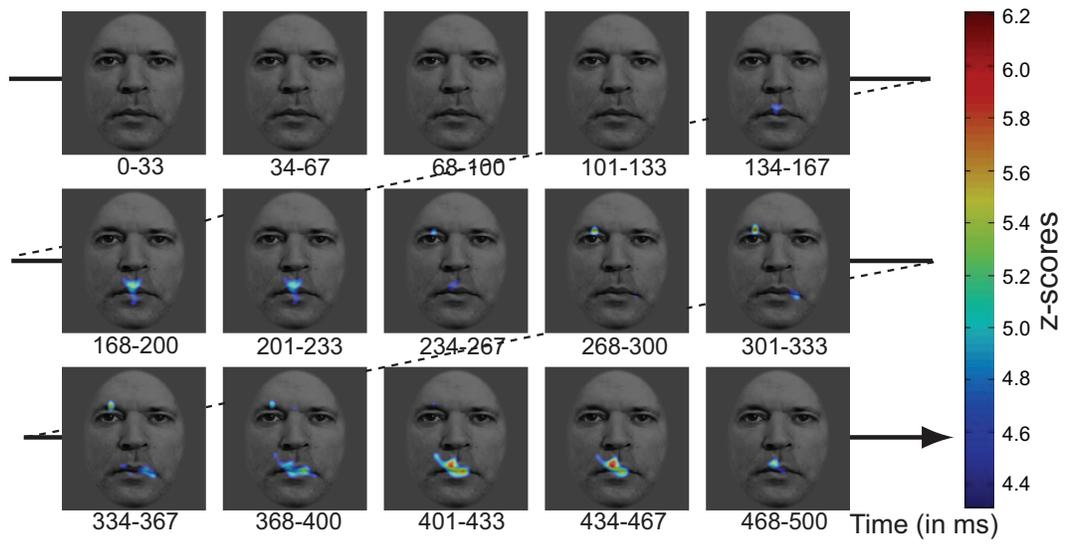


Figure 2b

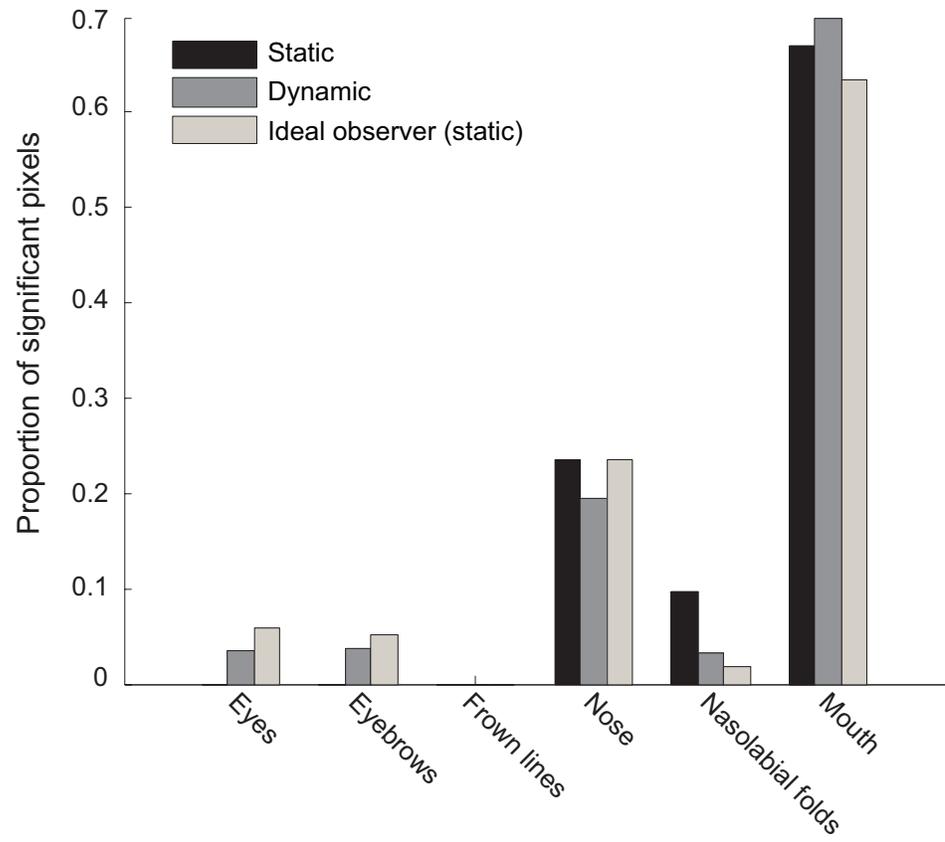


Figure 3

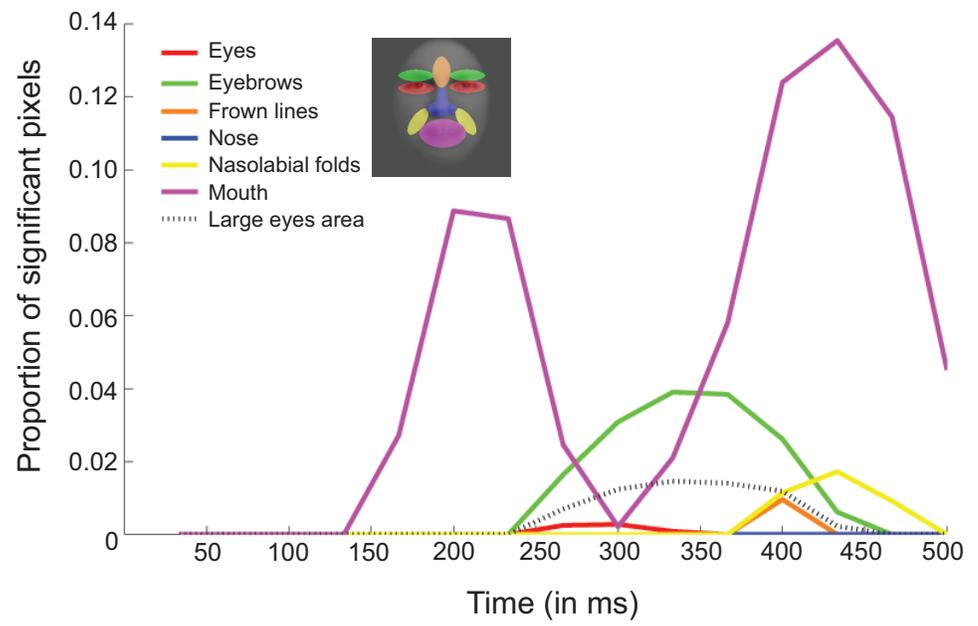


Figure 4

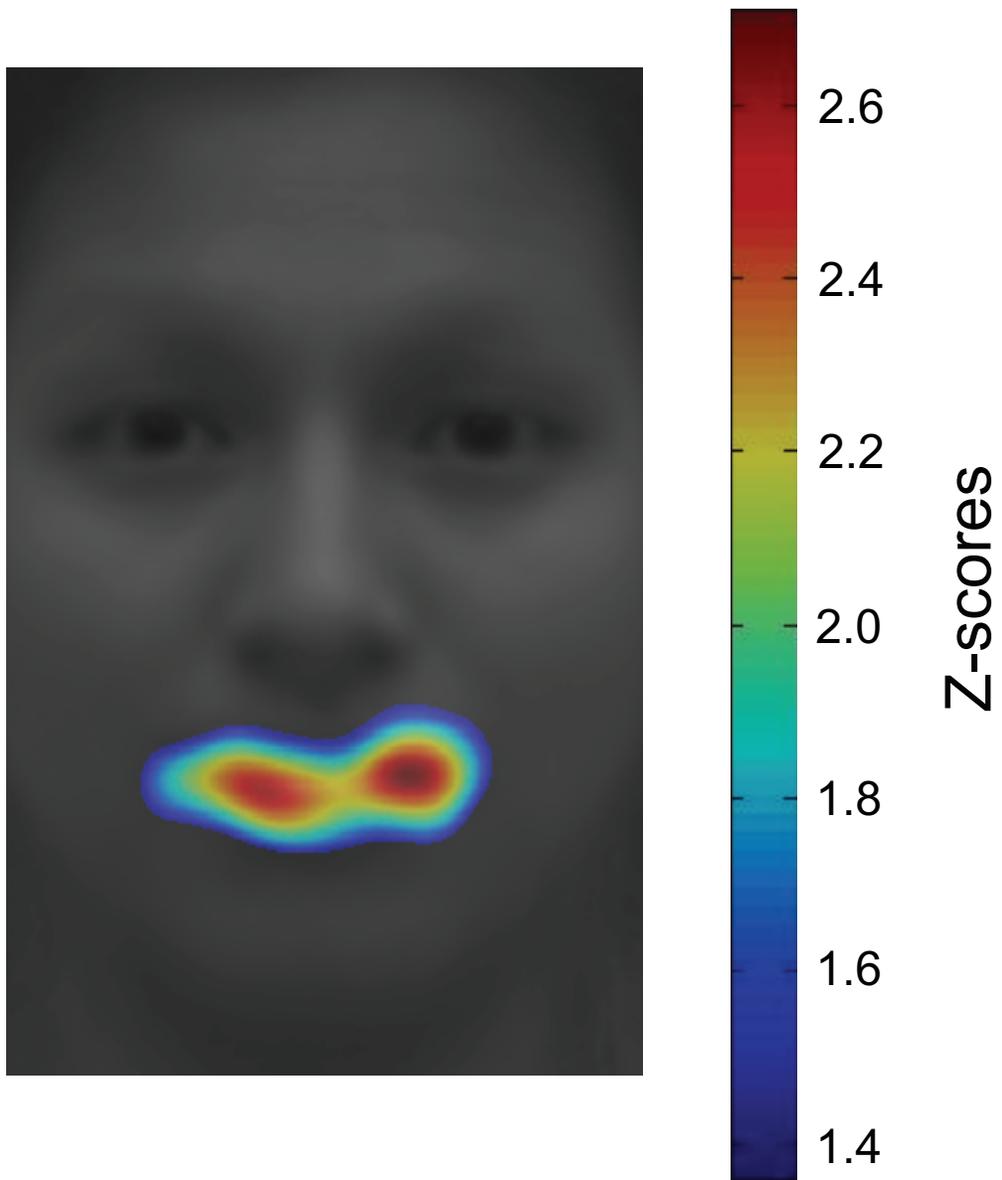


Figure 5

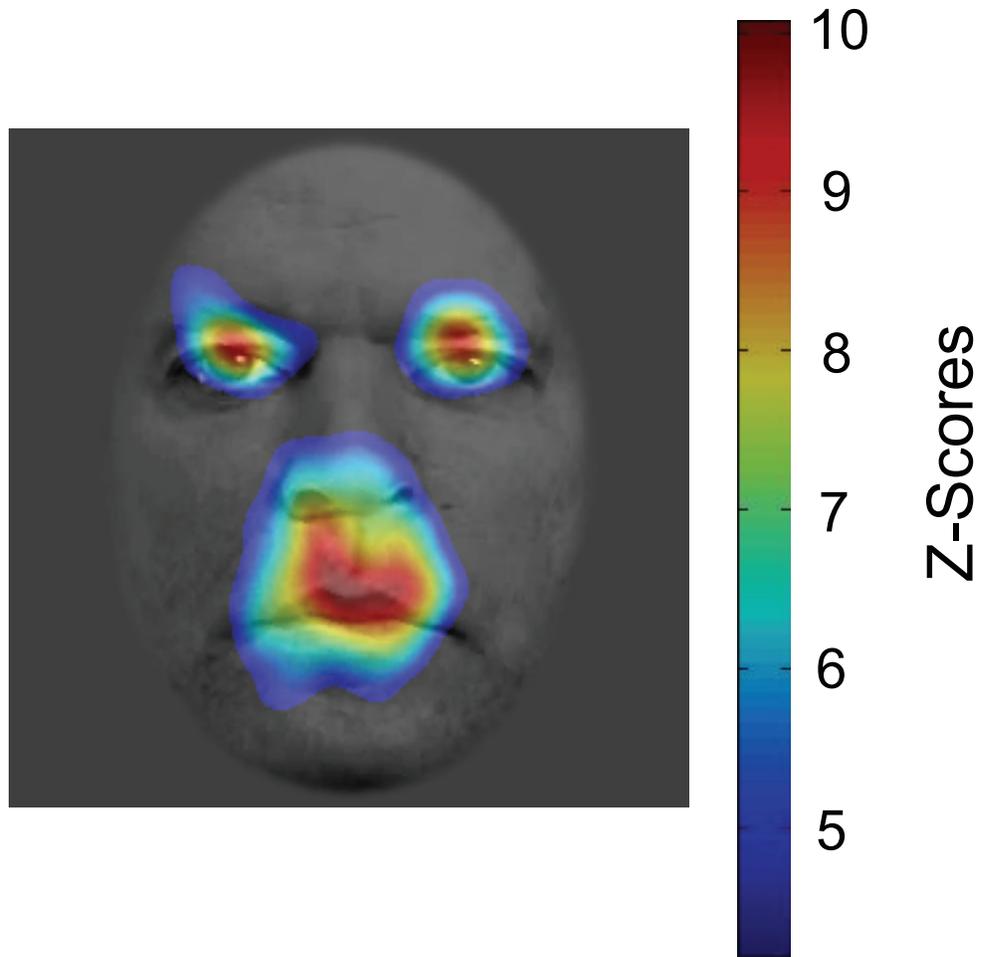


Figure 6

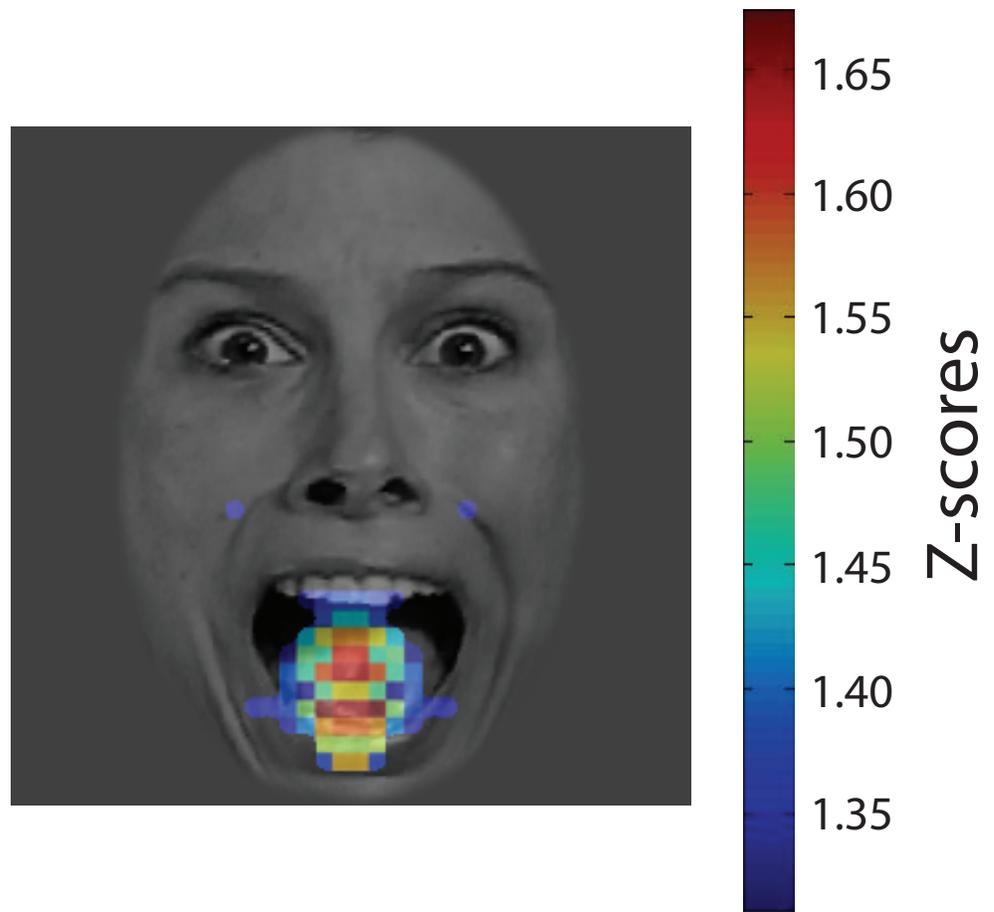


Figure 7

Chapitre 3

Article 2

**Individual differences in the visual extraction strategies for
recognizing facial expressions**

Caroline Blais¹, Daniel Fiset², Cynthia Roy¹, Martin Arguin¹, & Frédéric
Gosselin¹

¹ *Centre de Recherche en Neuropsychologie et Cognition, Département de
Psychologie, Université de Montréal, CP 6128, succ. Centre-ville, Montréal,
H3C 3J7, CANADA*

² *Département de Psychoéducation et de Psychologie, Université du Québec en
Outaouais, CP 1250, succ. Hull, Gatineau, J8X 3X7, CANADA*

Address correspondence to

Frédéric Gosselin, Ph.D.

3.1 Abstract

An underutilization of the eye region has been observed in many clinical populations with impaired facial expressions recognition. This raises the question of whether any decrease in the ability to recognize facial expressions is associated with decreased processing of the eyes. To answer this question, we used Bubbles to examine how the ability to recognize facial expressions relates to the visual processing strategies used. We found that processing of the left eye and the mouth is significantly correlated with the ability level for facial expression recognition, whereas processing of the right eye is not. The order in which the facial features are processed also vary as a function of the ability. Our results show qualitatively different processing strategies between the best and the worst facial expression recognizers. The present findings fail to support the idea that a decrease in the ability to recognize facial expressions is necessarily accompanied by a decrease in the processing of the complete eye area.

3.2 Introduction

The recognition of facial emotional expressions is a crucial skill for successful social interactions. As a matter of fact, a clear deficit in facial expression recognition has been described in many psychopathologies characterized by abnormal social interactions, such as schizophrenia (Mandal, Pandey, & Prasad, 2008) and autism spectrum disorders (Harms, Martin, & Wallace, 2010). Recently, these psychopathologies have been linked to atypical visual processing strategies. For instance, compared to normal individuals, schizophrenic patients rely less on the high spatial frequencies in the eye area to identify fearful faces (Lee, Gosselin, Wynn, & Green, 2010), and adults suffering from autism rely less on the eye area when categorizing facial expressions (Spezio, Adolphs, Hurley, & Piven, 2007a; 2007b). A link between the impaired recognition of emotional facial expressions and atypical eye processing has also been observed in patients with a brain lesion in the amygdala (Adolphs et al., 2005; Gosselin, Spezio, Tranel, & Adolphs, 2011), in a prosopagnosic patient (Caldara et al., 2007), and in normal aging (Sullivan, Ruffman, & Hutton, 2007). These observations raise the question of whether any decrease in the ability to recognize emotional facial expressions is associated with an atypical processing of the eye area.

To assess this possibility, we took advantage of the fact that the ability to recognize facial expressions varies among healthy individuals (Besel &

Yuille, 2010; Moeller, Ewing Lee, & Robinson, 2011; Young & Brunet, 2011). Thus, the aim of the present study is to determine whether variations in the ability to recognize emotional facial expressions in normal observers relate to the visual processing strategies used – in particular the processing of the eye area – to perform this task. To do so, we conducted new analyses on data from a larger project (see Blais et al., 2012) examining many dimensions of the visual information extraction strategies used for the discrimination of static and dynamic facial expression of emotions using the Bubbles technique (Gosselin & Schyns, 2001). The present study focused on the spatial and the temporal dimensions of the data. During the experiments in which this data set was collected, the participants' performance was controlled by adjusting the amount of information available (i.e., the number of bubbles on each trial) to maintain the criterion performance. The participants' ability was thus indexed using this measure (see Methods).

3.3 General Methods

We conducted two experiments examining the visual information strategies used to categorize static and dynamic facial expressions. Forty-one participants took part in Exp. 1 (static stimuli) and 59 different participants took part in Exp. 2 (dynamic stimuli). All had normal or corrected-to-normal visual acuity.

To reveal the visual information useful for the categorization of facial expressions, we sampled static and dynamic facial expressions with Bubbles, a classification image technique that reveals the facial areas and the spatial frequencies used by participants to successfully complete the task. During each of the 4000 trials, the participant categorized a static (i.e., Exp. 1) or a dynamic (i.e., Exp. 2) facial expression revealed through a mask of bubbles. The facial expression stimuli came from a validated database (Roy et al., 2007; see <http://mapageweb.umontreal.ca/gosselif/STOIC.rar>) and consisted of pictures or movies of 10 actors' faces expressing one of eight emotional states (i.e., anger, disgust, fear, happy, pain, sad, surprise, and a neutral expression). The movies (15 frames; total duration of 500 ms) started with a neutral expression that naturally deployed into one of the eight target expressions and ended with the apex of the expression. The static stimuli consisted in the apex of the facial expression and were displayed for 500 ms. The stimuli subtended 5.72 degrees of visual angle in width and 7.87 degrees in height. They were flipped over the vertical axis on half the trials to equalize the amount of information in both hemifields.

For each trial, the construction of a bubbled stimulus went as follows. First, the image of a facial expression was decomposed into five spatial frequency bands (86-43, 43-21.5, 21.5-10.8, 10.8-5.4, 5.4-2.7 cycles/face; the remaining low-frequency bandwidth served as a constant background; see Fig. 1a, top row) using the Laplacian pyramid (Burt & Adelson, 1983). With

dynamic stimuli, the spatial frequency decomposition was performed on each frame of the movies (see Fig. 1b, top row). Then, for each spatial frequency band independently, the bubbles' locations were randomly selected (see Fig. 1a and 1b, middle row). On the space dimension, the size of the bubbles increased as the spatial scale became coarser (FWHM : 14.1, 28.3, 56.5, 113.0, and 226.1 pixels), and the number of bubbles differed across scales to maintain the size of the sampled area constant across frequency bands. The size of the bubbles also varied as a function of spatial frequency on the time dimension (i.e., with dynamic bubbles), such that their duration increased as the spatial frequency band increased (FWHM: 7.3, 6.1, 5.1, 4.2 and 3.5 frames). This was done to take into account the faster processing of lower spatial frequencies (Hughes, Fendrich, Reuter-Lorenz, 1990; Parker, Lishman, & Hughes, 1992). A pointwise multiplication was then performed between the bubbles' masks and the filtered images (see Fig. 1a and 1b, bottom row). Finally, the information revealed by the bubbles was fused across the five frequency bands to produce an experimental stimulus (Fig. 1a, bottom row, rightmost image).

Accuracy was maintained at an average of 56.25% correct (i.e., halfway between chance and optimal performance) across all expressions by adjusting the total number of bubbles in the stimulus on a trial-by-trial basis using QUEST (Watson & Pelli, 1983).

Insert Fig. 1 near here

3.4 Results.

Means of 144.3 bubbles (SD: 119.6) and of 241.3 bubbles (SD: 253.6) were necessary to maintain the average performance at 56.25% correct with the static and the dynamic stimuli, respectively.

For each participant, each facial expression and each spatial frequency band, a classification image (i.e., CI; or a classification volume, CV, in the dynamic experiment) was constructed by calculating the weighted sum of all the bubbles centers of the bubbles masks presented to the participant. The weights consisted in the accuracy of the participant on each trial, transformed into z-scores. This resulted in 40 CIs or CVs per participant (i.e., eight expressions by five spatial frequency bands).

Then, two series of analyses were conducted. The first determined the average information use across participants and the second, examined how information use varied as a function of individual differences in the ability to recognize facial expressions. The Average CI and CV (i.e., CI_{average} and CV_{average}) were computed to help us with the interpretation of the Individual differences CI and CV (i.e., CI_{ID} and CV_{ID}), namely the meaning of a positive

vs. negative correlation. In particular, in the CI_{ID} , highly positive values may come from highly positive values in the individual CIs and highly positive ranks in the performance index. Conversely, they may also come from highly negative values in the individual CIs and highly negative ranks in the performance index. Thus, the absence of facial areas negatively correlated with a correct response in the Average CI and CV will indicate that only the first type of outcome occurred. The $CI_{average}$ (or $CV_{average}$) was computed by summing the individuals' CIs (or CV). The CI_{ID} (or CV_{ID}) was computed by performing a weighted sum of the individuals' CIs (or CVs) using an index of each participant's performance level as weight. The weights represented the rank of the participants on the number of bubbles they required to maintain an average accuracy of 56.25% correct. Thus, the participant that needed the fewest bubbles was ranked first and was given a positive weight (arbitrarily set to 1), the participant who required the most bubbles was ranked last and was given a negative weight (arbitrarily set to -1). The participants in between were given weights that varied linearly between -1 and 1 according to their rank¹. The resulting CIs and CVs were transformed into z-scores. The pixels outside the face area were used to calculate the mean and standard deviation of the distribution of the null hypothesis. To determine the spatial information used to categorize the eight expressions, the CIs and CVs were summed across the expressions tested and across the five spatial frequency bands. The CVs were also summed across the 15 temporal frames. The resulting CIs were smoothed using a Gaussian window with a FWHM of 28.3 pixels, which corresponds to

the size of the Gaussian window used with the second spatial frequency band during the experiment. To determine the facial information that correlated significantly with accuracy (i.e., in the CI_{average}) or with the individual differences (i.e., in the CI_{ID}), we applied the pixel test ($Z_{\text{crit}}=3.84$, $p < 0.01$) to the normalized CIs (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005). The statistical threshold provided by this test corrects for multiple comparisons while taking the spatial correlation inherent to structured images into account.

The thresholded CI_{average} and CI_{ID} for the static and for the dynamic stimuli are displayed in Figs. 2 and 3, respectively. The CI_{average} obtained with the static and the dynamic stimuli showed that on average, both eyes as well as the mouth areas are used by the participants to recognize facial expressions. Note also that there were no strong negative values in the CI_{average} . This indicates that in the CI_{ID} , highly positive values come from highly positive values in the individual CIs and highly positive ranks in the performance index. Conversely, strongly negative values come from highly positive values in the individual CIs and strongly negative ranks in the performance index. The CI_{ID} both showed that the utilization of the left eye and of the mouth areas was positively correlated with a high level of ability in recognizing facial expressions. No facial area was significantly negatively correlated with recognition ability. Taken together, these results indicate that the better (vs. the worse) a participant was at the task, the more (vs. the less) he used the left eye and the mouth area. However, the utilization of the right eye area did not vary

with the ability of the participants to recognize facial expressions (i.e., $p=0.57$ with static stimuli; $p=.65$ with dynamic stimuli).

Insert Figs. 2 and 3 near here

Since we used dynamic bubbles with the dynamic stimuli, it was possible to determine whether the temporal order in which the facial areas are processed varies according to the individual differences in the ability to process facial expressions. For this, one CV per participant was computed using the procedure described above. For each participant, a region-of-interest (ROI) analysis was conducted using the left eye, the right eye, and the mouth as ROIs. For each temporal frame, the average of the z-score values of the pixels located in each ROI was calculated. This resulted in three temporal vectors per participants (i.e., one for each ROI). A weighted sum of each vector was then computed across participants using as weights the same index of performance as in the previous analysis. Finally, a bootstrap procedure was used to transform these vectors into z-score values and to determine statistical significance.

The outcome of this analysis is presented in Fig. 4. The statistical thresholds (i.e., $p < 0.001$, two-tailed and Bonferroni corrected for 15 frames times three vectors) are represented by two black horizontal dotted lines. Early after stimulus onset, there was a positive correlation between the participants'

level of ability and the utilization of the left eye area, indicating that the better a participant was, the more he used that area. In contrast, there was a negative correlation between the participants' ability and the utilization of the mouth area and of the right eye area, indicating that the worse a participant was, the more he used these two areas early in stimulus processing. Around 133 ms following stimulus onset, the correlation between the participants' ability and the utilization of the mouth area and of the right eye area became positive. Finally, there was a negative correlation between the participants' ability and the utilization of the left eye area between 300 and 367 ms. These results indicate a clear variation in the temporal order in which the facial areas are processed as a function of an individual's ability to recognize facial expressions.

Insert Fig. 4 near here

3.5 Discussion

Visual processing strategies have been shown to be altered in many clinical populations suffering from impairment at recognizing emotional facial expressions. The present results show that even among a healthy population, qualitative differences can be observed in the visual processing strategies of individuals (see also Langner, Becker, & Rinck, 2009). Indeed, the utilization of the different facial areas varies as a function of the ability of an observer to

recognize basic emotional facial expressions: use of the left eye and the mouth areas is correlated with performance, while utilization of the right eye is not. Interestingly, a very similar pattern is observed for both static and dynamic facial expressions. The lack of correlation between right eye utilization and an individual's level of ability indicates that the difference between the high and low performance participants does not simply lie in a better processing of the informative facial features by the high performance participants. Had this been the case, the right eye would also have been correlated with performance since the utilization of this area is clearly part of the averaged strategy employed by our participants, and since the information in both hemifields was equalized by flipping the stimuli over the vertical axis on half of the trials. Instead, our results indicate qualitative differences among the visual processing strategies as a function of performance. That is, the high and low performance participants did not only differ in the amount of processing they did on the diagnostic areas, but they also differed in the actual source of the information subtending their performance.

The analysis examining the order in which the different facial areas are processed also confirms the presence of qualitative differences by showing a clear variation in the temporal processing of the different facial areas as a function of the ability to recognize emotional facial expressions. While the high performance participants make an early use of the left eye area, the low performance participants instead use the mouth and the right eye areas at the

beginning of the stimulus presentation. The early utilization of the left eye area by the high performance participants is intriguing, as we have shown in a previous study (Blais et al., 2012) that the averaged strategy of participants to recognize basic emotional facial expressions starts with the utilization of the mouth area, which is indeed the most informative area for this task. An early use of the left eye area has however been observed in a face identification task (Vinette, Gosselin, & Schyns, 2004). This raises the question of whether our high performance participants might actually be good in tasks related to faces in general and use a strategy that is efficient with most of these judgments, rather than one that is mostly efficient for the recognition of facial emotional expressions. A better characterization of what distinguishes high vs. low ability participants, and how being good at one kind of facial judgment generalizes to other facial judgments, appears as a relevant avenue for future research.

In the Introduction, we highlighted the frequent observation of an underutilization of the eye area in patients impaired at recognizing facial expressions. We asked the question of whether any impairment with emotional facial expressions leads to a gradual decrease in the utilization of eye area. Our results are inconsistent with this idea. Indeed, the absence of correlation between right-eye utilization and performance suggests that the underutilization of the eye area found in patients should be restricted to the left eye, which is not the case. Moreover, the correlation between utilization of the mouth and the ability to perform the task suggests that, if patients were an extreme on the

continuum existing in the healthy normal population, a differential use of the mouth should be found when patients are compared to healthy participants. Again, this is not verified in the results from patients with impaired recognition of facial expressions. Thus, the pattern observed in many clinical populations (i.e., strict underutilization of the eye area) is not what would be expected if these patients were part of a continuum whereby the worse someone is, the less he uses the mouth and the left eye areas. Rather, their visual processing strategy appears to have undergone qualitative changes relative to that of normal observers.

In summary, our results indicate that within the normal population, the visual processing strategies used are qualitatively influenced by the individuals' ability to recognize emotional facial expressions. They also strongly suggest that the systematic observation of an underutilization of the eye area by patients does not simply reflect an extreme on the continuum existing in the healthy normal population but rather that it signals a qualitative difference. Future research will help to better understand how different factors, for example traits of personality such as empathy or extraversion/introversion, or cognitive factors such as individual differences in the ability with different facial judgments, may influence the visual processing strategies, and how these factors and their impact on the visual processing may be related to the ability to recognize emotional facial expressions.

3.6 Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (CRSNG) to Frédéric Gosselin and by a scholarship from the Fonds Québécois de Recherche sur la Nature et les Technologies (FQRNT) to Caroline Blais.

3.7 References

- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P. G. & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68-72
- Besel, L.D.S., & Yuille, J. C. (2010). Individual differences in empathy: The role of facial expression recognition. *Personality and Individual Differences*, *49*, 107-112.
- Blais, C. Fiset, D., Roy, C., Arguin, M., & Gosselin, F. (2012). The eyes are not the windows to emotions. [Abstract]. International Neuropsychological Society, Montreal, Canada.
- Burt, P., & Adelson, E. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, *31*, 532–540.
- Caldara, R., Fiset, D., Blais, C., Schyns, P., Scheepers, C., & Meyer, E. (2007). Clarifying the nature of facial identity and facial expression representations with an acquired case of prosopagnosia. *Journal of Vision*, *7*(9): 628; doi: 10.1167/7.9.628
- Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*, 659-667
- Gosselin, F. & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition. *Vision Research*, *41*, 2261-2271.

- Gosselin, F., Spezio, M.L., Tranel, D. & Adolphs, R. (2011). Asymmetrical use of eye information from faces following unilateral amygdala damage. *Social Cognitive and Affective Neuroscience*, 6, 330-337
- Harms, M., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychological Review*, 20, 290-322.
- Hughes, H. C., Fendrich, R., & Reuter-Lorenz, P. A. (1990). Global versus local processing in the absence of low spatial frequencies. *Journal of Cognitive Neuroscience*, 2, 272-282.
- Langner, O., Becker, E. S., & Rinck, M. (2009). Social anxiety and anger identification: bubbles reveal differential use of facial information with low spatial frequencies. *Psychological Science*, 20, 666-670.
- Lee, J., Gosselin, F., Wynn, J. K., & Green, M. F. (2010). How do schizophrenia patients use visual information to decode facial emotion? *Schizophrenia Bulletin*, 37, 1001-1008.
- Mandal, M.K., Pandey, R.P., & Prasad, A.B. (1998). Facial expressions of emotions and schizophrenia: a review. *Schizophrenia Bulletin*, 24, 399–412.
- Moeller, S. K., Ewing Lee, E. A., & Robinson, M. D. (2011). You never think about my feelings: Interpersonal dominance as a predictor of emotion decoding accuracy. *Emotion*, 11, 816-824.
- Parker, D.M., Lishman, J.R., & Hughes, J. (1992). Temporal integration of spatially filtered visual images. *Perception*, 21, 147-160

- Roy, S., C. Roy, Fortin, Ethier-Majcher, Belin & Gosselin (2007). A dynamic facial expression database. *Journal of Vision*, 7, 944.
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007a). Abnormal use of facial information in high-functioning autism. *Journal of Autism and Developmental Disorders*, 37, 929–939.
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007b). Analysis of face gaze in autism using “Bubbles”. *Neuropsychologia*, 45, 144–151.
- Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *Journal of Gerontology: Psychological Sciences*, 62B, 53-60.
- Vinette, C., Gosselin, F. & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It’s in the eyes! *Cognitive Science*, 28, 289-301
- Watson, A. B. & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33, 113-20.
- Young, M. E., & Brunet, P. M. (2011). The influence of sociability—but not shyness—on the categorization of facial expressions of emotion. *Journal of Individual Differences*, 32, 65-73

3.8 Footnote

1. We also computed the CI_{ID} and CV_{ID} using the number of bubbles transformed into z-scores as weights rather than the ranked performance; the results were essentially the same.

3.9 Figure Legends

Figure 1a. Illustration of the construction of a bubbled stimulus. (a) Procedure with static stimuli. A facial expression image was decomposed into five spatial frequency bands (top row). Bubbles were then placed at random locations separately for each spatial frequency band (middle row). The information revealed by the bubbles (bottom row) was then fused across the five frequency bands to produce an experimental stimulus. The column delimited by the red rectangle indicates the spatial frequency band used on Fig. 1b to illustrate the procedure with dynamic stimuli (b) With dynamic stimuli, the procedure to produce a bubbled stimulus was very similar to the procedure for static stimuli, except that the decomposition into five spatial frequency bands was performed on each frame. Here, we illustrate the procedure for the third frequency band and the first 10 frames of the video. The same procedure was repeated on each frequency band. Each frame of the video was decomposed into five frequency bands (top row). Bubbles were then placed randomly at different locations and frames, separately for each spatial frequency band (middle row). The bottom row illustrates the information revealed by the bubbles on each frame.

Figure 2. Thresholded average classification images obtained in Exps. 1 and 2. The areas depicted in color were statistically significantly correlated with the participants' performance.

Figure 3. Thresholded individual differences classification images obtained in Exps. 1 and 2. The areas depicted in color were statistically significantly correlated with the participants' performance.

Figure 4. ROI analysis. The statistical thresholds are represented by the two black horizontal dotted lines. The temporal frames for which the z-score value is higher than the upper statistical threshold indicate a positive correlation between the ability to recognize emotional facial expressions and the utilization of that ROI (i.e., the better a participant is, the more he uses that area). In contrast, the temporal frames for which the z-score value is lower than the lower threshold indicate a negative correlation between the ability to recognize emotional facial expressions and the utilization of that ROI (i.e., the worse a participant is, the more he uses that area).

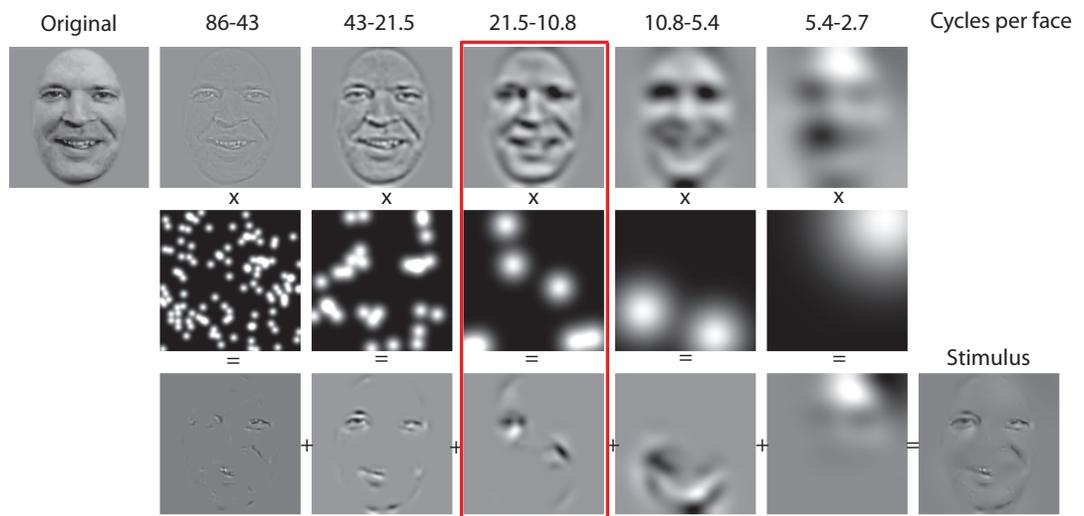


Figure 1a. Blais, Fiset, Roy, Arguin, & Gosselin

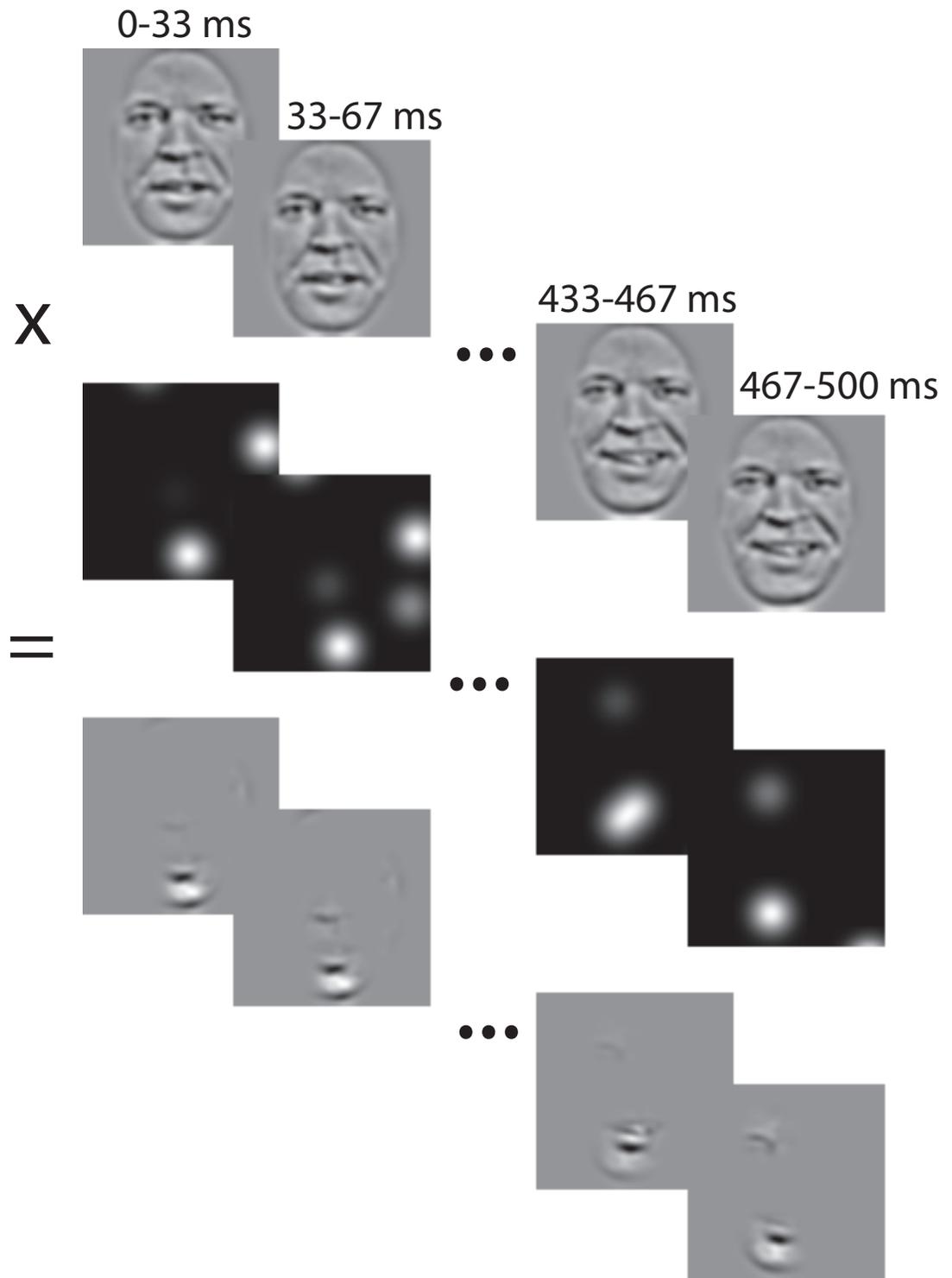


Figure 1b. Blais, Fiset, Roy, Arguin, & Gosselin

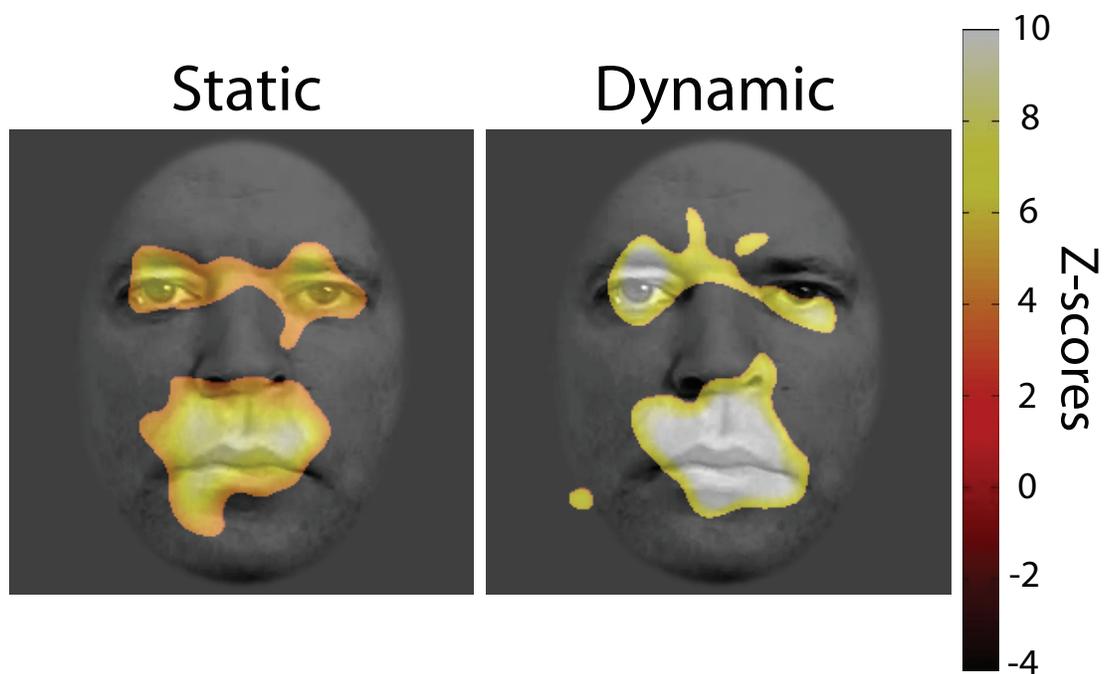


Figure 2. Blais, Fiset, Roy, Arguin, & Gosselin

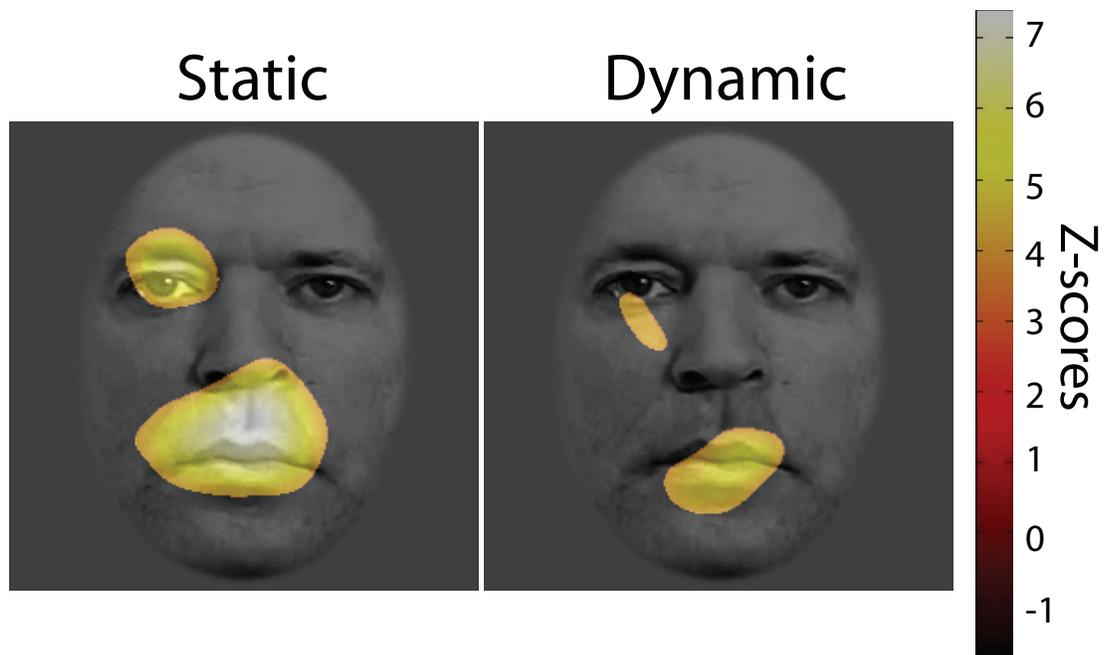


Figure 3. Blais, Fiset, Roy, Arguin, & Gosselin

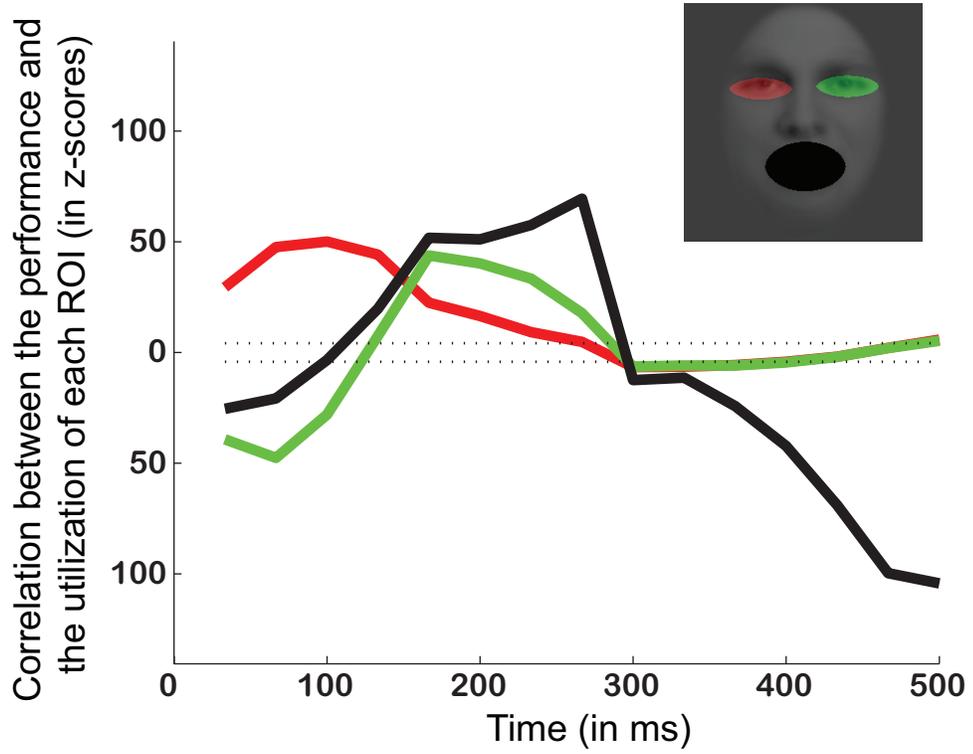


Figure 4. Blais, Fiset, Roy, Arguin, & Gosselin

Chapitre 4

Discussion générale

4.1 Retour sur les objectifs

L'objectif principal de cette thèse consistait à mieux comprendre les stratégies visuelles sous-jacentes à la reconnaissance des expressions faciales de base. L'information visuelle utile pour reconnaître certaines expressions faciales spécifiques commence à être mieux connue (e.g., Adolphs et al., 2005; Nusseck et al., 2008; Smith et al., 2005; Smith & Schyns, 2009). Par exemple, la région des yeux est particulièrement utile pour reconnaître l'expression faciale de peur (Adolphs et al., 2005; Smith et al., 2005). Toutefois, l'information visuelle utilisée pour discriminer plusieurs expressions faciales entre elles est beaucoup moins bien connue, et les études qui se sont intéressées à cette problématique ont utilisé des méthodes biaisées (voir section 1.3.1). Pourtant, puisque le système visuel ne peut savoir quelle expression faciale il traite avant d'avoir extrait une certaine quantité d'information, il est probable que certaines régions soient priorisées, du moins en début de traitement, parce qu'elles permettent de discriminer un maximum d'expressions faciales entre elles. La première étude de cette thèse avait donc pour principal objectif de clarifier la stratégie utilisée pour extraire l'information visuelle permettant de discriminer toutes les expressions faciales de base entre elles. La seconde étude visait à vérifier si la stratégie révélée dans la première étude est liée aux variations, au sein de la population normale, de l'habileté à reconnaître les expressions faciales.

4.2 Objectif 1 : Clarifier la stratégie visuelle utilisée pour discriminer les expressions faciales de base entre elles

Tel que mentionné dans la section 1.4, la première étude de cette thèse visait (1) à révéler les régions faciales utilisées par des observateur humains pour discriminer toutes les expressions de base (i.e., portion « décodage » de l'expression faciale); (2) à vérifier de manière objective quelles régions du visage contiennent l'information la plus discriminante à l'aide d'un observateur idéal; et (3) à comparer l'information utilisée par les observateurs humains à celle utilisée par l'observateur idéal. Tout ceci dans l'optique où une évaluation objective de l'hypothèse selon laquelle la région des yeux comporte une importance cruciale dans le processus de reconnaissance des expressions faciales d'émotions serait également rendue possible.

4.2.1. Stratégie des humains : Montre-moi ta bouche et je te dirai ce que tu ressens!

Les résultats indiquent que les observateurs humains utilisent la région des yeux de même que la région de la bouche pour discriminer les expressions faciales de base entre elles. Néanmoins, une comparaison de l'utilisation relative des deux régions montre que la région de la bouche est 8.5 fois plus utile que la région des yeux pour discriminer les expressions faciales de base statiques entre elles, et 15.5 fois plus utile pour discriminer les expressions faciales dynamiques entre elles. De plus, l'échantillonnage de la dimension temporelle avec les stimuli dynamiques a permis de vérifier dans quel ordre les

différentes régions du visage sont traitées. Les résultats montrent que la bouche est la première région à être traitée, et demeure la région la plus importante durant la majeure partie de la présentation de l'expression faciale.

Afin de s'assurer que nos résultats n'étaient pas simplement dus à l'ensemble de stimuli utilisé, une réanalyse de l'étude de Smith et al. (2005) a été effectuée. Dans cette étude, la méthode des Bulles avait été utilisée dans une tâche de reconnaissance d'expressions faciales de base, mais des visages différents avaient été utilisés et l'expression de douleur n'avait pas été présentée. Les analyses effectuées par les auteurs visaient principalement à vérifier quelle information visuelle, en terme de régions de visage et de fréquences spatiales, étaient utilisées par les participants pour identifier correctement chaque expression faciale. Une analyse supplémentaire, présentée dans la première étude de cette thèse, a permis de vérifier l'utilisation des différentes régions de visage pour discriminer toutes les expressions faciales entre elles. Les résultats de cette analyse répliquent les résultats obtenus dans notre étude (i.e., une plus grande utilisation de la région de la bouche). Il est également intéressant de noter que d'autres paramètres de l'étude de Smith et al. (2005) différaient de ceux utilisés pour la première étude de cette thèse. Par exemple, le nombre de bulles était ajusté différemment dans les deux études. Dans l'étude de Smith et al. (2005), le nombre de bulles était ajusté indépendamment pour chaque expression faciale, de sorte que la performance était égale à travers les sept expressions faciales utilisées. Avec une telle

procédure, il est possible que les participants soient arrivés à se servir du nombre de bulles comme indice pour identifier l'expression faciale présentée. Afin de contourner ce problème, notre étude contrôlait le nombre de bulles à travers toutes les expressions faciales, de sorte que la performance variait d'une expression faciale à l'autre. Le temps de présentation était lui aussi différent dans les deux études. Dans l'étude de Smith et al. (2005), les stimuli étaient présentés jusqu'à la réponse du participant. Dans notre étude, les stimuli étaient présentés pendant 500 ms. Ainsi, les résultats obtenus avec les données de Smith et al. (2005) sont très similaires à ceux obtenus dans la première étude de la présente thèse, et ce malgré que les deux études aient utilisé des stimuli différents, des temps de présentation différents, et aient ajusté la performance selon différentes procédures. La similarité des résultats obtenus malgré l'utilisation de paramètres expérimentaux différents milite en faveur de la généralisabilité de nos résultats.

4.2.2. Information contenue dans les différentes régions du visage : Observateur idéal et analyse de mouvement.

Afin de mieux comprendre pourquoi les participants donnent autant d'importance à la région de la bouche, un observateur idéal a été soumis à la tâche de reconnaissance d'expressions faciales avec des bulles. Cette procédure a permis de trouver quelles régions du visage contiennent l'information la plus discriminante pour les expressions faciales incluses dans l'ensemble de stimuli

utilisé pour la première expérience. L'observateur idéal utilisait lui aussi la région de la bouche dans une plus grande mesure que la région des yeux. Toutefois, le ratio de l'utilisation de la bouche vs. des yeux était supérieur pour les humains par rapport à pour l'observateur idéal, suggérant que les humains utilisent plus la bouche et moins les yeux par rapport à l'observateur idéal. De plus, une analyse du mouvement contenu dans les expressions faciales dynamiques utilisées dans la deuxième expérience montre que les mouvements survenant dans la région de la bouche contiennent davantage d'information quant à l'expression présentée.

4.2.3. Information utilisée vs. Information disponible

Les résultats de l'observateur idéal de même que ceux obtenus suite à l'analyse du mouvement contenu dans les expressions faciales dynamiques indiquent tous deux que la région de la bouche est hautement informative pour discriminer les expressions faciales de base entre elles. Ainsi, la plus grande utilisation de la région de la bouche que de la région des yeux par les participants est probablement due, du moins en partie, à la quantité d'information contenue dans cette région. Néanmoins, le fait que le ratio de l'utilisation de la bouche vs. des yeux soit supérieur chez les humains par rapport à chez l'observateur idéal indique que la quantité d'information ne permet pas à elle seule d'expliquer la stratégie des humains.

Deux possibilités ont été proposées dans l'article pour expliquer le biais des humains pour la bouche. Une première possibilité repose sur le fait que le système visuel est exposé de manière naturelle et répétée à des expressions faciales dynamiques, puisque c'est ainsi qu'elles surviennent en milieu écologique. Suivant cette exposition, il est probable que la stratégie qu'il ait développé pour décoder les expressions faciales soit en partie basée sur l'information contenue dans le mouvement. Dans ce cas, une stratégie dans laquelle beaucoup de ressources seraient dirigées vers la région de la bouche serait bien adaptée. Cette stratégie, une fois mise en place, serait utilisée même en l'absence de mouvement dans le stimulus (i.e., avec les expressions faciales présentées de manière statique). Cette hypothèse est consistante avec les données indiquant que la région cérébrale STS, connue pour son implication dans le traitement du mouvement biologique, soit impliquée à la fois dans le traitement des expressions faciales dynamiques et dans celui des expressions statiques (Furl, van Rijsbergen, Treves, Friston, & Dolan, 2007; Haxby, Hoffman, & Gobbini, 2000; Hoffman & Haxby, 1999; Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001; Pessoa & Padmala, 2007; Tsuchiya, Kawasaki, Oya, Howard, & Adolphs, 2008).

Une deuxième possibilité pour expliquer le biais des observateurs humains à utiliser la région de la bouche vient des limites inhérentes au système visuel. En effet, l'observateur idéal n'a aucune contrainte en terme de quantité d'information pouvant être traitée. Le système visuel humain est, quant à lui,

extrêmement limité. Les études sur le « change blindness » (Simon & Levin, 1997), durant lesquelles un changement grossier survenant dans la scène visuelle à traiter n'est pas remarqué par les participants, témoignent bien de ces limites. Ainsi, parce qu'il lui est difficile de traiter toute l'information utile, il est possible que le système visuel alloue davantage de ressources à la région du visage contenant le plus d'information (i.e., la bouche). Cette hypothèse pourrait d'ailleurs expliquer pourquoi une sous-utilisation de la région des yeux est observée chez plusieurs populations cliniques souffrant de difficultés à reconnaître les expressions faciales d'émotions (voir section 1.3.4). En effet, il est plausible qu'une lésion cérébrale ou un trouble développemental mène à une diminution des ressources, et donc à une diminution de l'attribution de ressources vers les régions moins informatives du visage.

Notons que les deux hypothèses susmentionnées ne sont pas mutuellement exclusives. Les ressources limitées du système visuel ont probablement un rôle à jouer dans le développement des stratégies d'extraction de l'information utile pour le traitement d'un stimulus visuel. Ainsi, des ressources limitées pourraient contraindre le système visuel à développer une stratégie dans laquelle beaucoup de ressources sont dirigées vers la bouche, puisque le mouvement qui survient dans cette région est très informatif pour discriminer les expressions faciales de base entre elles. Bien entendu, les hypothèses proposées ici sont spéculatives et devront être évaluées avec des études spécifiquement structurées pour répondre à cette question.

4.2.4. Mouvements oculaires et utilisation de l'information visuelle

Une stratégie dans laquelle une haute importance est accordée à la région de la bouche peut à première vue paraître inconsistante avec les patrons de mouvements oculaires observés durant la reconnaissance des expressions faciales. En effet, une répartition égale des fixations oculaires est typiquement retrouvée entre la région des yeux et la région de la bouche durant cette tâche (Eisenberg & Alpers, 2011; Jack et al., 2009). Il est toutefois important de noter que les fixations oculaires ne sont pas le reflet direct de l'information traitée. En effet, une dissociation existe entre l'attention et les fixations oculaires (Jonides, 1981; Posner, 1980; voir toutefois Rayner, 1998; Deubel & Schneider, 2003; Godijn & Theeuwes, 2003), et une étude récente suggère que les mouvements oculaires ne reflètent pas uniquement l'information utilisée, mais également des effets visuo-moteurs qui ne sont pas nécessairement modulés par la tâche cognitive en cours (Arizpe, Kravitz, Yovel, & Baker, 2012). Ainsi, l'observation d'une répartition égale des fixations oculaires durant la reconnaissance d'expressions faciales n'est pas inconsistante avec l'observation d'une *utilisation* supérieure de la région de la bouche. En effet, il est possible que la région de la bouche soit traitée pendant que la région des yeux est fixée—le traitement serait simplement effectué en dehors de la fovéa.

L'observation que la région de la bouche est particulièrement importante pour la reconnaissance des expressions faciales est d'ailleurs consistante avec des résultats obtenus dans une étude où les mouvements oculaires de

participants asiatiques et caucasiens étaient comparés alors qu'ils devaient catégoriser des expressions faciales d'émotions (Jack et al., 2009). Les asiatiques sont connus pour présenter des taux de reconnaissances plus faibles que les caucasiens avec certaines expressions négatives (Elfenbein & Ambady, 2002; Mesquita & Fridga, 1992; Russell, 1994). De manière intéressante, les représentants de ce groupe culturel présentaient des patrons de fixations oculaires dans lesquels la région de la bouche était sous-représentée. Bien sûr, une sous-représentation des fixations oculaires au niveau de la région de la bouche chez les asiatiques n'implique pas en soi un sous-traitement de la région de la bouche. Toutefois, à l'aide d'une analyse computationnelle, Jack et al. (2009) ont montré qu'une sous-utilisation de la région de la bouche pouvait expliquer les patrons d'erreurs observés chez les asiatiques. Ainsi, les résultats de Jack et al. (2009) appuient l'idée que la bouche est particulièrement importante pour la reconnaissance des expressions faciales en montrant qu'une sous-représentation des fixations oculaires sur cette dernière et que sa sous-utilisation sont liées à des erreurs de catégorisation.

4.2.5 Contribution de la région des yeux au percept de l'expression faciale d'émotion.

Le domaine de la reconnaissance des visages s'est vu inondé, au cours des dernières décennies, d'études visant à montrer que l'identification des visages est basée sur un mode de traitement configural (Diamond & Carey,

1986) ou holistique (Tanaka & Farah, 1993). Le modèle configural propose que le traitement de l'identité des visages est basé sur les relations spatiales entre les différentes parties d'un visage, par exemple la distance entre les deux yeux, ou la distance entre le nez et la bouche. Le modèle holistique propose quant à lui que l'identité est codée comme une représentation globale dans laquelle les différentes parties de visages (e.g., yeux, nez, bouche) ne sont pas représentées explicitement. Les deux modèles sont similaires en ce qu'ils prédisent tous deux une intégration non linéaire des parties de visages. Ils s'opposent à un modèle dans lequel les différentes parties de visages sont utilisées de manière indépendante et sont intégrées de manière linéaire (Sekuler, Gaspar, Gold, & Bennett, 2004; Gold, Mundy, & Tjan, 2012).

Le domaine de la reconnaissance des expressions faciales d'émotion, quoique moins touché par cette vague de recherche opposant les modes de traitement configural/holistique et « par partie », n'y a toutefois pas échappé. L'un des paradigmes les plus populaires pour mettre en évidence l'utilisation automatique d'information configurale ou holistique durant l'identification des visages est l'effet composite (Young, Hellawell, & Hay, 1987). Dans ce paradigme, des visages composites sont créés en juxtaposant un haut et un bas de visage provenant de deux identités différentes. Lorsque deux visages présentés simultanément ou séquentiellement sont formés du même haut mais de bas différents, les hauts de visages ont tendance à être perçus comme différents. Toutefois, lorsque les mêmes stimuli sont présentés en désalignant le

haut et le bas du visage, le taux d'erreurs sur les jugements de haut de visage diminue. Cet effet est considéré par plusieurs comme une preuve que « le tout est plus grand que la somme de ses parties », puisqu'un percept complètement différent est créé lorsque le haut et le bas du visage peuvent être intégrés. L'effet composite est donc fréquemment utilisé comme un index du traitement configural ou holistique (e.g., Gauthier, Klaiman, & Schultz, 2009; Goffaux & Rossion, 2006; Michel, Rossion, Han, Chung, & Caldara, 2006; Rossion & Boremanse, 2008). Le paradigme de l'effet composite a été appliqué dans le domaine de la reconnaissance des expressions faciales (Calder, Young, Keane & Dean, 2000; Chiller-Glaus, Schwaninger, Hofer, Kleiner, & Knappmeyer, 2011). De manière similaire aux résultats trouvés en identification de visages, la perception de l'émotion exprimée par le haut du visage est influencée par l'émotion exprimée par le bas du visage. Sur la base de ces résultats, il a été proposé que le traitement des expressions faciales, tout comme celui de l'identité, reposerait sur un traitement configural de l'information contenue dans un visage.

Tout ce prélude sur l'effet composite vise à mettre l'emphase sur l'importance de ne pas mésinterpréter nos résultats. Bien sûr, contrairement à l'idée préconçue que la région des yeux est cruciale pour la reconnaissance des expressions faciales, nos résultats indiquent que c'est en fait la région de la bouche qui est la plus importante. Toutefois, nos résultats montrent également un lien statistiquement significatif entre la reconnaissance correcte des

expressions et l'utilisation de l'information contenue dans la région des yeux. Nos résultats ne sont donc pas incongruents avec l'existence de l'effet composite : le haut et le bas du visage sont tous deux utilisés pour reconnaître les expressions faciales, et peuvent interagir pour créer le percept final. Des analyses supplémentaires seraient toutefois nécessaires pour établir si les deux régions étaient utilisées conjointement par nos participants.

De plus, il est important de noter que la conclusion que la région de la bouche revêt une importance supérieure pour la reconnaissance des expressions faciales s'applique uniquement aux émotions de base (i.e., plus neutralité et douleur), puisque seules ces dernières ont été incluses dans les expériences de la présente thèse. Il demeure donc possible que la région des yeux ait un rôle spécial à jouer dans la reconnaissance des émotions complexes, telles que le sentiment de culpabilité ou l'embarras par exemple. Baron-Cohen, Wheelwright et Jolliffe (1997) ont comparé le taux de reconnaissance d'expressions d'émotions de base et d'expressions d'émotions complexes lorsque seulement les yeux, seulement la bouche, ou tout le visage est présenté aux participants. Leurs résultats sont incongruents avec les nôtres en ce qui a trait aux expressions de base : le taux de reconnaissance est supérieur lorsque le visage complet est présenté, mais la présentation isolée de la région des yeux ou de la région de la bouche mène à des taux de reconnaissance similaires. Pour les émotions complexes, leurs résultats indiquent que les expressions sont aussi bien reconnues lorsque le visage complet est présenté que lorsque la région des

yeux est présentée de manière isolée, et que le taux de reconnaissance dans ces deux conditions est supérieur à la condition où la bouche est présentée de manière isolée. Un autre groupe de chercheurs s'est intéressé à la contribution relative des yeux et de la bouche pour la reconnaissance d'expressions d'émotions complexes, mais présentées de manière dynamique (Back, Jordan, & Thomas, 2009). Ils ont trouvé que l'importance de ces deux régions dépend de l'émotion exprimée, mais que la région des yeux est plus souvent critique pour le taux de reconnaissance que la région de la bouche. Ces deux ensembles de résultats suggèrent donc que la région des yeux est plus importante que la région de la bouche pour la reconnaissance des expressions faciales d'émotions complexes. Néanmoins, puisque les résultats de Baron-Cohen et al. (1997) sont incongruents avec les nôtres pour les émotions de base, il serait intéressant d'utiliser une méthode où les régions (ex. yeux vs. bouche) ne sont pas prédéfinies de manière arbitraire, par exemple la méthode des Bulles, pour contre-vérifier ces résultats. En effet, tel que mentionné à la section 1.3.1, une méthode où les régions comparées sont définies arbitrairement peut mener à des biais dans les résultats. Il est donc possible que l'utilisation d'une méthode moins biaisée au niveau de l'espace de recherche mène à des conclusions différentes.

4.3 Objectif 2 : Vérifier le lien entre les différences individuelles au niveau de l'habileté à reconnaître les expressions faciales et les stratégies visuelles pour extraire l'information

Les résultats de la première étude de cette thèse montrent que pour discriminer toutes les expressions faciales de base entre elles, la région de la bouche et la région des yeux sont toutes deux utiles. Toutefois, la région de la bouche est plus utilisée que la région des yeux, et commence à être utilisée de manière plus précoce que la région des yeux, probablement parce qu'elle contient plus d'information pour cette tâche. Dans l'Introduction, il est mentionné que l'une des raisons probables pour l'idée préconçue que la région des yeux est particulièrement importante pour la reconnaissance des expressions faciales d'émotions repose sur l'observation systématique d'une sous-utilisation de la région des yeux par des population cliniques souffrant de difficultés avec cette tâche. Dans le contexte des résultats obtenus dans la première étude de cette thèse, l'observation susmentionnée implique que les populations cliniques utilisent *encore moins* que la population normale la région des yeux. Cette observation soulève la question suivante : est-il possible que le niveau d'habileté pour reconnaître les expressions faciales soit lié à l'utilisation de la région des yeux? Le deuxième article de cette thèse avait donc pour but d'examiner comment les différences individuelles au niveau de l'habileté à reconnaître les expressions faciales dans la population normale sont liées aux stratégies d'extraction de l'information visuelle dans cette tâche.

4.3.1 Modulation quantitative vs. qualitative des stratégies visuelles

Les résultats obtenus dans la deuxième étude de la présente thèse montrent un lien entre les stratégies visuelles et l'habileté à reconnaître les expressions faciales. L'utilisation de l'œil gauche et de la bouche est positivement corrélée avec l'habileté à reconnaître les expressions faciales. En d'autres mots, à mesure que le niveau d'habileté augmente à travers les participants, l'utilisation de la bouche et de l'œil gauche augmente elle aussi. De plus, une analyse de la corrélation entre l'habileté dans cette tâche et l'information utilisée à différents moments dans le temps montre que les différentes régions du visage ne sont pas utilisées aux mêmes moments par les meilleurs et les moins bons participants. Au début de la présentation du stimulus, l'habileté et l'utilisation de l'œil gauche corrélaient positivement, indiquant que les meilleurs participants ont un biais pour utiliser cette région en début de traitement. Au contraire, l'habileté et l'utilisation de la bouche et de l'œil droit corrélaient négativement, indiquant que les participants les plus faibles ont un biais pour utiliser ces deux régions en début de traitement. L'utilisation de la bouche et de l'œil droit commence à être positivement corrélée à l'habileté autour de 133 ms, alors que l'utilisation de l'œil gauche est négativement corrélée à l'habileté entre 300 et 367 ms.

Ainsi, les résultats sont incongruents avec l'idée que l'habileté à reconnaître les expressions faciales soit strictement liée à l'utilisation de la région des yeux. Une telle hypothèse prédisait en effet une corrélation de

l'habileté avec l'utilisation des deux yeux, et une absence de corrélation avec la région de la bouche. Au contraire, nos résultats indiquent une absence de corrélation entre l'utilisation de l'œil droit et l'habileté, de même que la présence d'une corrélation positive entre l'utilisation de la bouche et l'habileté. La sous-utilisation de la région des yeux observée chez les populations cliniques ne reflète donc pas un continuum dans lequel pire une personne est pour reconnaître les expressions faciales, moins elle utilise la région des yeux. Au contraire, elle reflète probablement la survenue d'un changement qualitatif dans les stratégies des patients.

Les résultats montrent également que la stratégie des meilleurs participants ne se distingue pas uniquement de celle des moins bons participants par une meilleure utilisation des régions informatives du stimulus (i.e., par des différences quantitatives). Si le contraire était vrai, l'habileté devrait également corrélérer avec l'utilisation de l'œil droit. L'observateur idéal présenté dans le premier article montre en effet que les deux yeux sont informatifs pour discriminer les expressions faciales entre elles, et la stratégie moyenne de tous les participants indique que l'œil droit est utilisé. L'absence de corrélation entre l'habileté et l'œil droit, lorsque considérée dans le contexte de la stratégie moyenne et de l'informativité relative des différentes régions du visage, indique que bien que cette région soit utilisée par les humains, son utilisation n'est pas liée à leur habileté pour discriminer les expressions faciales. La modulation du moment auquel les différentes régions du visage sont utilisées en fonction de

l'habileté confirme d'ailleurs la présence de différences qualitatives dans les stratégies d'extraction de l'information visuelle.

À cet effet, l'utilisation précoce de l'œil gauche par les meilleurs participants, de même que la corrélation positive entre l'habileté et l'utilisation de l'œil gauche, sont intrigants. Étant donné que la région de la bouche est la plus informative, l'observation d'une utilisation de la bouche très tôt dans la stratégie moyenne faisait beaucoup de sens. Une utilisation précoce de l'œil gauche a déjà été observée durant le traitement de l'identité des visages (Vinette, Gosselin, & Schyns, 2004). Il est possible que les individus qui présentent des habiletés supérieures pour la reconnaissance des expressions faciales présentent également un talent particulier avec d'autres types de jugements liés au traitement des visages, et que la stratégie visuelle qu'ils utilisent se soit développée de manière à être optimale pour effectuer la plupart de ces jugements. Il serait certainement intéressant de vérifier si l'habileté à effectuer un jugement spécifique sur les visages se généralise à d'autres types de jugements et d'examiner comment, au sein d'un même individu, les stratégies sont modulées par le type de jugement effectué. D'autre part, la corrélation entre l'utilisation de l'œil gauche et l'habileté à reconnaître les expressions faciales, lorsqu'associée à une absence de corrélation entre l'utilisation de l'œil droit et l'habileté, soulève la possibilité que des différences hémisphériques puisse en partie expliquer nos résultats.

4.3.3 Différences hémisphériques en reconnaissance d'expressions faciales

Il existe un débat en ce qui a trait aux différences hémisphériques avec les émotions (voir Borod et al., 1998; Canli, 1999). L'hypothèse de « l'hémisphère droit » propose que le traitement des émotions est pris en charge de manière prédominante par l'hémisphère droit. L'hypothèse de la « valence » propose que l'hémisphère gauche traite de manière prédominante les émotions positives (e.g., joie et surprise), alors que l'hémisphère droit traite de manière prédominante les émotions négatives (e.g., colère, dégoût, peur, tristesse). De nombreuses études appuient l'une ou l'autre des théories, mais aucun consensus n'a été atteint jusqu'à maintenant. Les deux hypothèses s'entendent toutefois sur un point : les deux hémisphères contribuent différemment au traitement des émotions.

Le design de la deuxième étude de cette thèse ne permet pas de se prononcer sur les deux hypothèses dominantes quant aux différences hémisphériques dans le traitement des émotions. Toutefois, il est possible que des différences hémisphériques expliquent en partie l'observation d'une latéralisation au niveau de la corrélation entre l'utilisation des yeux et l'habileté à reconnaître les expressions faciales. Une explication qui pourrait à première vue sembler logique serait de proposer que l'hémisphère droit des meilleurs participants est plus efficace que celui des moins bons participants pour cette tâche. En effet, cet hémisphère est considéré dominant dans la tâche de reconnaissance d'expressions faciales, du moins avec les expressions faciales

négatives, lesquelles étaient prédominantes dans cette étude. La structure du système visuel fait en sorte que l'information située dans l'hémichamp gauche des participants, du moins celle située en région parafovéale de la rétine, était prioritairement traitée par l'hémisphère droit. Un hémisphère droit plus efficace pourrait donc avoir mené à une meilleure utilisation de l'œil gauche des stimuli. Par ailleurs, il existe un débat concernant la projection cérébrale de l'information située dans la région fovéale de la rétine. Bien que le modèle classique de la rétine propose que l'information située sur la fovéa soit représentée de manière bilatérale au niveau du cortex visuel primaire, un modèle dans lequel la fovéa est représentée de manière unilatérale au niveau du cortex visuel primaire a vu le jour au cours des dernières décennies (i.e., modèle *split fovea*; Ellis & Brysbaert, 2010; Shillcock, Ellison, & Monaghan, 2000). Selon ce modèle, l'information située dans la portion gauche de la fovéa serait prioritairement traitée par l'hémisphère droit, et vice versa. Si cette hypothèse est vraie, et que la latéralisation au niveau de l'utilisation des yeux s'explique par un hémisphère droit plus efficace chez les meilleurs participants, on aurait dû obtenir une utilisation latéralisée de la région de la bouche aussi, laquelle est située au centre du stimulus. L'hypothèse d'un hémisphère droit plus efficace explique donc mal pourquoi les deux côtés de la bouche—plutôt que seulement le côté gauche de la bouche—étaient corrélés avec l'habileté des participants.

Une seconde possibilité serait que des différences d'informativité existent entre l'œil gauche et l'œil droit. Ce n'est bien sûr pas le cas au sein de nos stimuli, puisque l'image miroir des photographies était présentée pour la moitié des essais. Toutefois, certaines études indiquent que les deux côtés d'un visage ne sont pas égaux dans leur transmission de l'expression faciale d'émotion (Campbell, 1978; Chaurasia & Goswami, 1975; Sackeim & Gur, 1978) : le côté gauche du visage est plus expressif que le côté droit du visage lorsque l'expression faciale est posée. Une analyse plus détaillée de l'expressivité relative des côtés gauche et droit du visage montre que la supériorité du côté gauche se retrouve seulement pour le bas du visage, alors que pour le haut du visage le côté droit est plus expressif que le côté gauche (Asthana & Mandal, 1997; Richardson, Bowers, Bauer, Heilman & Leonard, 2000; voir cependant Ross & Pulusu, in press). L'œil qui se retrouve à gauche sur nos stimuli représente en fait l'œil droit d'une personne exprimant une émotion. Ainsi, placé dans ce contexte, nos résultats indiquent une corrélation positive entre l'habileté des participants à reconnaître les expressions faciales et l'utilisation de l'œil qui est habituellement le plus expressif dans les expressions faciales posées.

Les résultats de la seconde étude pourraient donc refléter une stratégie dans laquelle les propriétés naturelles des expressions faciales sont mieux prises en compte par les meilleurs participants. Si c'est le cas, alors on devrait s'attendre à ce que seule l'utilisation du côté droit de la bouche sur les images

de classification soit corrélée avec l'habileté des participants à reconnaître les expressions faciales. Les images de classification 2D (i.e., celles qui ne prennent pas en compte la dimension temporelle, ne vont pas en ce sens). Toutefois, une décortication de l'utilisation des côtés gauche et droit de la bouche à travers le temps indique que la corrélation négative entre l'habileté et l'utilisation de la bouche s'explique principalement par l'utilisation du côté gauche de la bouche sur le stimulus (i.e., côté droit de la personne qui exprime; voir Fig. 1). Au contraire, la corrélation positive retrouvée avec la bouche s'explique principalement par l'utilisation du côté droit de la bouche sur les stimuli (i.e., côté gauche de la personne qui exprime). En d'autres mots, les meilleurs participants ont un biais pour utiliser le côté le plus expressif de la bouche, tout comme ils ont un biais pour utiliser l'œil le plus expressif. Tel que mentionné ci-haut, cette différence d'expressivité n'était pas présente dans les stimuli eux-mêmes, puisque ces derniers étaient présentés en image miroir pour la moitié des essais. Toutefois, elle se retrouve probablement de manière naturelle dans les expressions faciales posées, et une hypothèse pour expliquer nos résultats serait de proposer que les meilleurs participants ont développé une stratégie qui prend mieux en compte ces différences d'expressivité.

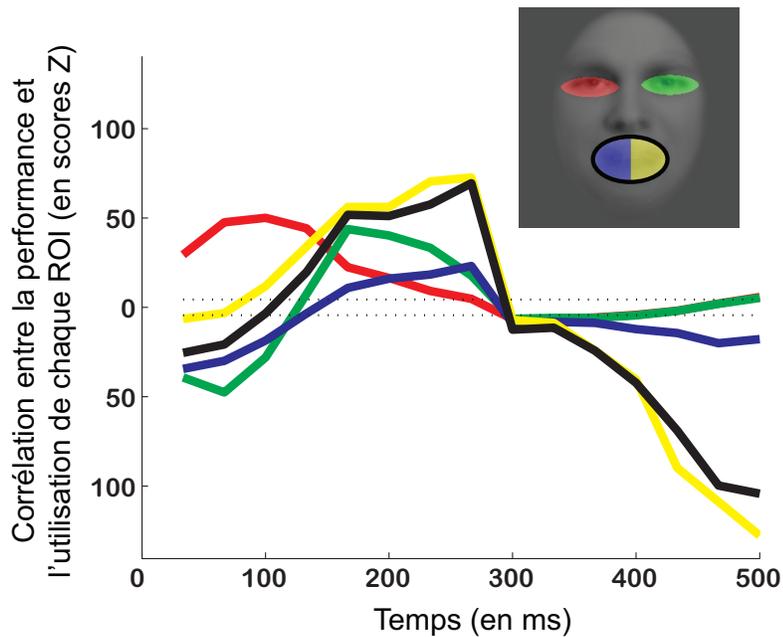


Figure 1. Résultat de l'analyse par région d'intérêt (ROI) à travers le temps. Les deux lignes pointillées représentent les seuils de signification ($p < 0.001$, avec correction de Bonferonni pour les comparaisons multiples). Chaque ligne représente la corrélation entre l'habileté des participants et l'utilisation d'une ROI. Le niveau d'utilisation d'une ROI est mesuré en calculant la moyenne des scores Z à l'intérieur de la ROI pour chaque participant.

Bien que cette hypothèse soit séduisante parce qu'elle permet d'expliquer pourquoi l'œil gauche corrèle avec l'habileté alors que l'œil droit non, elle doit être considérée avec précaution. En effet, une étude récente suggère que l'expressivité supérieure de l'œil droit survient seulement avec les expressions faciales posées, et que l'œil gauche serait quant à lui plus expressif lorsque l'expression faciale survient de manière spontanée (Ross & Pulusu, sous presse). De manière intuitive, il est tentant de croire que les expressions

faciales spontanées surviennent beaucoup plus fréquemment que les expressions faciales posées dans l'environnement. Si c'est le cas, on devrait s'attendre à ce que le développement des stratégies visuelles pour la reconnaissance des expressions faciales soit contraint par les propriétés des expressions faciales spontanées plutôt que posées. Aucune étude n'a, à notre connaissance, comparé la fréquence des expressions faciales spontanées et posées dans l'environnement. Toutefois, plusieurs études ont montré que les expressions faciales spontanées sont plus ambiguës que les expressions faciales posées (Carroll & Russell, 1997; Motley & Camden, 1988; Scherer & Ellgring, 2007; Wagner, Macdonald, & Manstead, 1986). Une des hypothèses avancées pour expliquer cette ambiguïté est l'existence de *display rules* (Ekman, 1972), qui consistent à activer ou à inhiber certains muscles faciaux afin de masquer ou d'exagérer l'apparence de l'expression faciale reflétant l'émotion réellement ressentie par un individu. Ces *display rules* sont différents d'une culture à l'autre puisqu'ils sont modulés par les normes sociales—e.g., quelle expression faciale est appropriée dans un contexte spécifique. Dans l'environnement naturel, l'ambiguïté des expressions faciales spontanées ne pose pas problème puisqu'il est possible de se servir du contexte dans lequel l'expression est survenue pour interpréter cette dernière. La question à se poser alors est la suivante : est-ce que, comme être social, l'humain tente de décoder l'expression faciale que l'individu tente d'exprimer, et ce même si elle ne reflète pas nécessairement l'émotion ressentie par la personne? Ou bien tente-il plutôt de

décoder l'expression qui reflète l'émotion ressentie par l'individu, même si cette expression est ambiguë et que le contexte à lui seul pourrait permettre d'inférer l'émotion? Bien que davantage de recherche sera nécessaire pour répondre à cette question, les résultats de la deuxième étude de cette thèse pointent plutôt vers la première option. En effet, les régions faciales corrélées avec l'habileté des individus à reconnaître les expressions faciales sont celles qui contiennent le plus d'information dans les expressions faciales posées (i.e., celles que l'individu veut réellement communiquer).

4.4 Limites et perspectives futures

4.4.1 Les émotions complexes

Les études de la présente thèse se sont toutes deux limitées à l'examen des stratégies visuelles utilisées pour reconnaître les expressions faciales d'émotions de base (plus les expressions de neutralité et de douleur). L'idée que la région des yeux est cruciale pour la reconnaissance des expressions faciales est sérieusement remise en question par la première étude de cette thèse en ce qui a trait aux émotions de base. Toutefois, à la question « Les yeux sont-ils le miroir de l'âme? », nous pouvons seulement répondre qu'ils ne sont en tout cas pas le miroir des émotions de base. Tel que mentionné à la section 4.2.5, il a été suggéré que les émotions complexes seraient plus facilement reconnues grâce à la région des yeux (Baron-Cohen, Wheelwright et Jolliffe, 1997). L'étude ayant mené à cette conclusion avait toutefois utilisé une méthode biaisée dans sa

définition des régions utiles pour reconnaître les expressions faciales (voir section 1.3.2). Il serait donc intéressant d'examiner l'utilisation de l'information visuelle durant la reconnaissance d'émotions complexes à l'aide d'une méthode d'échantillonnage non biaisée telle que la méthode des Bulles. De manière similaire, l'interprétation des résultats de la deuxième étude est limitée aux émotions de base. Il est possible que l'information visuelle corrélée à l'habileté pour reconnaître les expressions faciales d'émotions complexes soit différente de celle corrélée à l'habileté pour reconnaître les expressions faciales d'émotions de base.

4.4.2 L'impact de divers facteurs cognitifs, affectifs et sociaux sur les stratégies visuelles

En lien avec ce dernier point, plusieurs questions demeurent en suspens suite à cette thèse concernant les caractéristiques des meilleurs participants. Par exemple, est-ce que les individus présentant des habiletés supérieures à la moyenne en reconnaissance d'expressions faciales de base sont également supérieurs en reconnaissance d'expressions d'émotions complexes? Sont-ils meilleurs en général avec les visages, par exemple en identification ou en catégorisation du genre? Sont-ils plus empathiques, ou présentent-ils des aptitudes sociales supérieures à la moyenne? Malheureusement, la deuxième étude de cette thèse se limitait à mesurer l'habileté dans la tâche de reconnaissance d'expressions de base lorsque ces dernières sont présentées à

travers un masque de bulles. Il serait certainement intéressant non seulement d'avoir une mesure indépendante de la performance pour reconnaître les expressions faciales de base, mais également de mesurer la performance des participants dans diverses tâches cognitives ou tests psychologiques. Par exemple, des mesures de l'empathie (Interpersonal Reactivity Index; Davis, 1980), de la sociabilité (The Revised Cheek and Buss Shyness and Sociability scale; Cheek & Buss, 1981) et de la personnalité (NEO-PI-3; McCrae, Costa & Martin, 2005) pourraient être incluses. L'utilisation de ces diverses mesures—ou même d'une combinaison de ces mesures—comme régresseur pour les masques de bulles durant la tâche de reconnaissance d'expression faciales de base pourrait nous informer sur l'impact de divers facteurs sur les stratégies visuelles.

4.4.3 Expressions faciales posées vs. spontanées

La majorité des études qui se sont intéressées au décodage des expressions faciales—les études de la présente thèse comprises—ont utilisé des expressions faciales posées. Les résultats des études présentées dans cette thèse suggèrent que la stratégie visuelle utilisée par les participants n'est pas uniquement le reflet de l'information contenue dans le stimulus. En effet, les individus testés utilisaient davantage la région de la bouche et moins la région des yeux que ce qui était prédit par l'information contenue dans le stimulus. De plus, l'information visuelle corrélée avec l'habileté des participants ne reflétait

pas uniquement un biais informatif au sein du stimulus. Au contraire, une des hypothèses soulevées était qu'elle reflète plutôt un biais informatif dans les expressions faciales retrouvées de manière naturelle dans l'environnement. Il serait intéressant d'étudier plus en profondeur les caractéristiques des expressions spontanées et les stratégies visuelles associées à la reconnaissance de ces dernières. La comparaison des stratégies utilisées pour reconnaître les deux types d'expressions pourrait nous renseigner sur les mécanismes sous-jacents au développement des stratégies visuelles pour la reconnaissance des expressions faciales. Par exemple, si nos stratégies visuelles sont développées à partir de l'observation d'expressions faciales spontanées, les divergences entre l'information utilisée pour reconnaître les expressions faciales spontanées et l'information contenue dans le stimulus devraient être moins grandes que celles que nous avons retrouvées dans les deux études de cette thèse avec des expressions faciales posées.

4.4.4 Mouvements oculaires et expressions faciales

Les deux études présentées dans cette thèse visaient à examiner l'information visuelle *utilisée* par les participants pour reconnaître les expressions faciales d'émotions. Ces données ne nous renseignent toutefois pas sur les patrons de mouvements oculaires liés à l'extraction de cette information visuelle. En effet, tel qu'expliqué à la section 4.2.4, une dissociation existe entre la localisation des fixations oculaires et la localisation de l'information

réellement traitée. Jusqu'à maintenant, les études ayant enregistré les patrons de mouvements oculaires durant la reconnaissance des expressions faciales ont inclus dans leurs analyses toutes les fixations survenant entre l'apparition du stimulus et la réponse du participant (e.g., Jack et al., 2009), ou survenant durant une période de plusieurs secondes suivants l'apparition du stimulus (e.g., 2.5 secondes dans l'étude d'Eisenberg & Alpers, 2011). Une telle procédure est potentiellement problématique puisque quelques centaines de millisecondes sont amplement suffisantes pour traiter une expression faciale (Adolphs, 2002; Eimer & Holmes, 2002). De plus, une étude sur les mouvements oculaires durant l'identification de visages a permis de montrer que deux fixations oculaires sont suffisantes pour cette tâche : en effet, la performance n'est pas supérieure lorsque plus de deux fixations sont permises que lorsque le nombre de fixations est contraint à deux (Hsiao & Cottrell, 2008). Bien que le nombre de fixations suffisant pour reconnaître une expression faciale soit inconnu pour l'instant, il est raisonnable de penser que ce nombre se situe autour de la même valeur qu'en identification de visage. Les études qui se sont jusqu'à maintenant intéressées aux mouvements oculaires en reconnaissance d'expressions faciales ont donc inclus dans leurs analyses des fixations oculaires qui survenaient après la fin du traitement du stimulus, et qui ne reflétaient pas les mécanismes visuo-cognitifs sous-jacents à ce traitement.

Une étude visant à mesurer les patrons de mouvements oculaires survenant durant la reconnaissance d'expressions faciales statiques et

dynamiques a été menée au courant de la présente thèse. L'un des objectifs de cette étude consistait à inclure dans les analyses seulement les fixations oculaires survenant durant les 500 premières millisecondes du traitement afin d'éviter le problème mentionné ci-haut. Des analyses préliminaires indiquent que lorsque cette procédure est utilisée, les patrons de mouvements oculaires sont relativement congruents avec l'information utilisée. La carte des probabilités moyennes de fixations sur les différentes régions du visage à travers toutes les expressions faciales (i.e., les mêmes stimuli que pour les deux études de la présente thèse furent utilisés) est présentée sur la Figure 5. Pour rappel, l'inclusion de toutes les fixations survenant durant les quelques secondes suivant l'apparition du stimulus menait à l'observation d'une distribution en forme de triangle inversé sur les régions du visage (i.e., les fixations étaient distribuées sur les yeux et la bouche et ce, de manière approximativement égale; Eisenberg & Alpers, 2011; Jack et al., 2009). Au contraire, lorsque seules les fixations survenant durant les 500 premières millisecondes sont incluses, la densité est maximale au centre du visage. Ce résultat est congruent avec celui de Hsiao & Cottrell (2008) en identification de visage, qui montre que les deux premières fixations surviennent habituellement autour du nez. Il est aussi congruents avec l'hypothèse que les premières fixations oculaires sur un stimulus sont toujours dirigées vers le centre de gravité de celui-ci (Findlay, 1982; Findlay & Gilchrist, 1997). Une étude récente a d'ailleurs obtenu des résultats en accord avec cette hypothèse durant le traitement de visages (Bindemann, Scheepers, & Burton, 2009). Les auteurs

ont montré que peu importe le point de vue selon lequel un visage est présenté (e.g., de face, de profil), les fixations survenant durant les 250 premières millisecondes sont irrémédiablement dirigées vers le centre du visage, même si pour certains point de vue cet endroit ne contient pas ou très peu d'information pour la tâche. Après les 250 premières millisecondes toutefois, les fixations commencent à se distribuer ailleurs sur le visage, avec une proportion de fixations remarquablement plus élevée dans la région des yeux. Ce patron de mouvements oculaires a été retrouvé par les auteurs dans deux contextes différents : durant une tâche de discrimination du genre et durant une tâche d'observation passive.

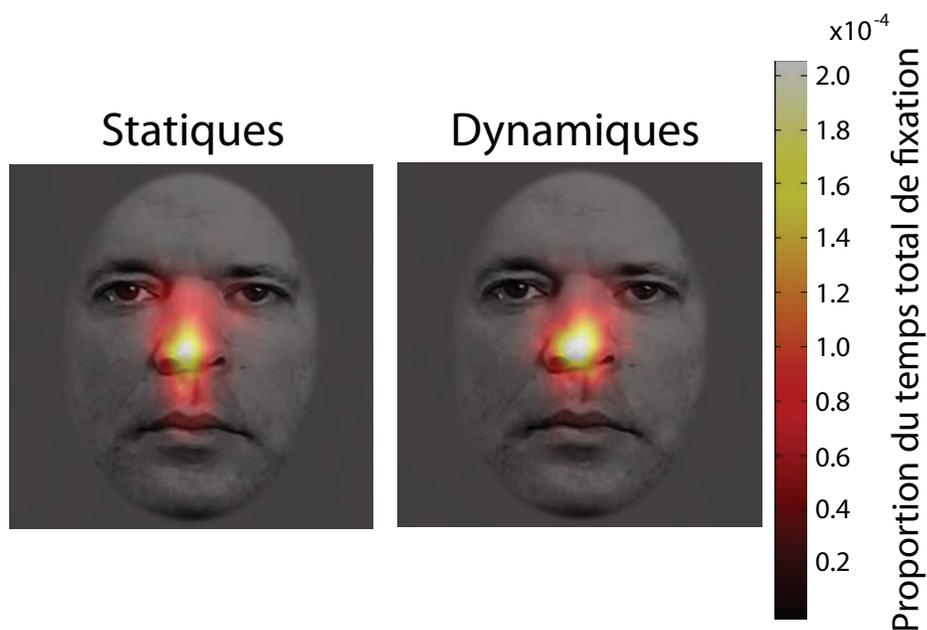


Figure 2. Cartes de la proportion moyenne des fixations oculaires sur chaque pixel du visage à travers les participants. Toutes les fixations survenant durant les 500 premières millisecondes suivant l'apparition du stimulus sont incluses.

Les pixels dont la probabilité de fixation était supérieure à 0 sont présentés en couleur.

Les données présentées dans la première étude de cette thèse indiquent que la bouche est la région la plus utilisée pour discriminer les expressions faciales de base entre elles. Les résultats présentés à la Figure 2 du présent chapitre montrent que durant les 500 premières millisecondes après l'apparition du stimulus, la proportion de fixation la plus élevée se situe au centre du visage. Ce résultat est probablement dû à la propension des premières fixations à se situer sur le centre de gravité d'un stimulus. Néanmoins, on peut se demander si la proportion des fixations se situant sur la bouche et sur les yeux vont dans le sens des résultats trouvés dans la première étude de cette thèse. Pour répondre à cette question, une analyse ROI similaire à celle décrite au Chapitre 2 a été effectuée pour comparer la proportion des fixations oculaires dirigée vers la région des yeux à celle dirigée vers la région de la bouche. Les résultats indiquent que la proportion de fixations est 2,7 fois plus élevée sur la région de la bouche que sur la région des yeux avec les expressions faciales statiques; et 5,2 fois plus élevée sur la région de la bouche que sur la région des yeux avec les expressions faciales dynamiques. Ainsi, nos résultats montrent que la distribution égale des fixations sur la région des yeux et la région de la bouche observée dans les études précédentes (Eisenberg & Alpers, 2011; Jack et al., 2009) n'est pas observée lorsque seules les fixations survenant durant les 500 premières millisecondes sont considérées. De plus, nos résultats, combinés à

ceux de Bindemann et al. (2009), montrent que la tâche influence les patrons de mouvements oculaires relativement tôt dans le traitement. Bindemann et al. (2009) trouvaient, dans l'intervalle survenant entre 250 et 500 ms après l'apparition du stimulus, une proportion très élevée de fixations oculaires dirigées vers la région des yeux durant une tâche de discrimination de genre ou d'observation passive. Nos résultats indiquent que durant les 500 premières millisecondes suivant l'apparition du stimulus, la proportion de fixations oculaires dirigées vers la région de la bouche est plus élevée que celles dirigées vers la région des yeux lorsque la tâche consiste à discriminer des expressions faciales d'émotions.

Les résultats de la deuxième étude de cette thèse montrent que l'information utilisée pour discriminer les expressions faciales entre elles est modulée par l'habileté des participants dans cette tâche. On peut donc se demander si les patrons de mouvements oculaires sont également modulés par l'habileté des participants à discriminer les expressions faciales. Pour répondre à cette question, une analyse similaire à celle présentée au Chapitre 3 a été effectuée : une corrélation entre les cartes de fixations oculaires et la performance des participants a été calculée. Le résultat de cette analyse pour les expressions faciales statiques et dynamiques est présenté à la Figure 3 du présent chapitre. Les patrons de mouvements oculaires sont liés à l'habileté à reconnaître les expressions faciales, et cette modulation semble congruente avec l'information utilisée. En effet, les fixations oculaires corrélées à la

performance sont légèrement décalées vers la portion gauche du visage. Pour rappel, l'utilisation de l'œil gauche (et non de l'œil droit) est corrélée avec l'habileté des participants. Dans le cas de la corrélation entre l'habileté et les fixations oculaires, la moyenne des scores Z dans le quadrant supérieur gauche est 3.3 et 6.0 fois plus élevée que celle dans le quadrant supérieur droit pour les expressions faciales statiques et dynamiques, respectivement.

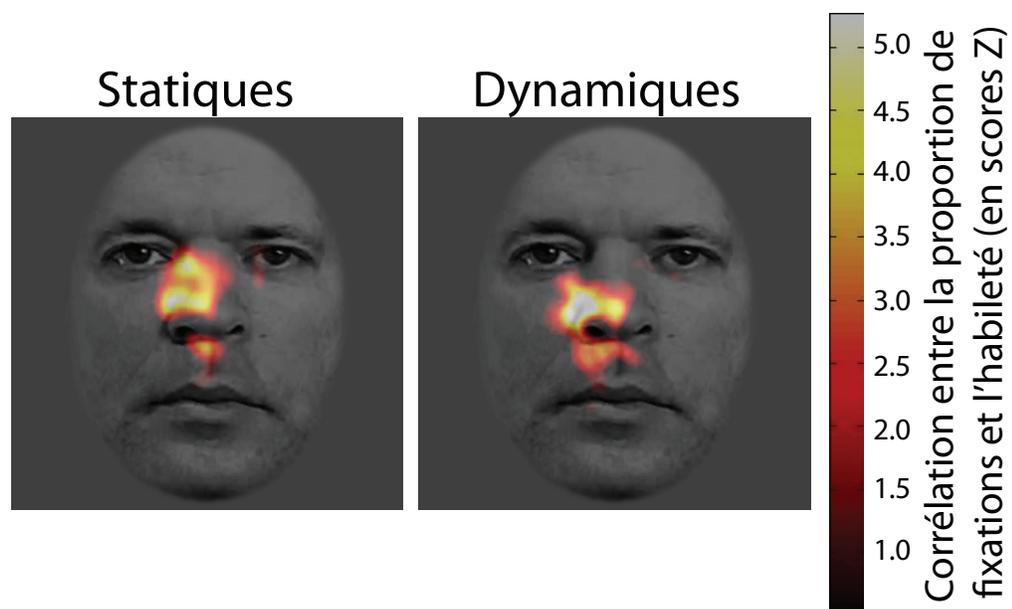


Figure 3. Cartes des corrélations entre les fixations oculaires et l'habileté des participants. Seules les corrélations positives sont présentées. Aucune corrélation négative avec un score Z inférieur à -1 n'a été obtenue.

Les patrons de mouvements oculaires observés dans notre étude sont donc relativement congruents avec les stratégies de traitement de l'information visuelle révélées avec la méthode des Bulles. Plusieurs questions demeurent néanmoins en suspens et nécessiteront des analyses ou des expériences

supplémentaires pour être répondues. Par exemple, tel que mentionné précédemment, aucune étude n'a vérifié jusqu'à maintenant combien de fixations oculaires sont suffisantes pour reconnaître les expressions faciales de base. Une expérience similaire à celle menée par Hsiao & Cottrell (2008), mais cette fois dans le domaine des expressions faciales, pourrait nous éclairer sur cette question. Une fois cette donnée en main, il serait intéressant de vérifier la distribution des fixations oculaires en incluant seulement les fixations qui contribuent réellement à la prise de décision. On peut également se demander si le nombre de fixations suffisant pour reconnaître les expressions faciales varie en fonction de l'habileté pour cette tâche, et si cette variation introduit des modifications dans les régions de visages traitées en priorité. En lien avec cette question, il serait intéressant de vérifier à quel moment des divergences surviennent dans les fixations oculaires des meilleurs et des moins bons participants. Tel que mentionné ci-haut, certaines études montrent que la position de la première fixation oculaire est toujours située sur le centre de gravité du stimulus, peu importe la tâche à accomplir. La première fixation semble donc être contrôlée par des mécanismes visuo-moteurs qui ne sont pas influencés par les processus de haut niveau spécifiques à une tâche. Est-ce que les meilleurs participants arrivent à « reprendre le contrôle » plus rapidement sur ces mécanismes de bas niveau? Ou bien la divergence se situe-elle au niveau de la région traitée une fois cette reprise de contrôle effectuée? La mise en lien des patrons de mouvements oculaires, de l'information utilisée, et de la modulation de ces deux variables par l'habileté des participants nous aidera

certainement à mieux comprendre les mécanismes en jeu durant la reconnaissance des expressions faciales d'émotion.

4.4.5. Conclusion

L'aptitude à reconnaître les expressions faciales des autres est cruciale pour le succès des interactions sociales. Les résultats de la présente thèse ont permis de montrer que la stratégie visuelle utilisée pour cette tâche implique un traitement privilégié de la région de la bouche, et que l'habileté à identifier les expressions faciales module cette stratégie. Ces résultats ouvrent la porte à une pléiade de questions. Par exemple, est-ce que les stratégies révélées sont optimales pour les expressions faciales spontanées? Quels facteurs visuels, cognitifs et sociaux caractérisent les individus qui présentent de meilleures habiletés avec la reconnaissance des expressions faciales? Comment les patrons de mouvements oculaires et l'information utilisée sont-ils liés? Bref, nous sommes encore loin de bien comprendre les stratégies visuelles utilisées pour la reconnaissance des expressions faciales de même que les facteurs qui les modulent et qui influencent leur développement. Mais comme disait Louis Pauwels : « La science progresse en indiquant l'immensité de l'ignoré ».

Références

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, *12*, 169-177.
- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P. G. & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68-72.
- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*, 669-672.
- Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, *16*, 403-410.
- Arizpe, J., Kravitz, D., Yovel, G., & Baker, C. (2012). Start position strongly influences fixation patterns during face processing: Difficulties with eye movements as a measure of information use. *PLoS ONE*, *7*(2): e31106. doi:10.1371/journal.pone.0031106.
- Back, E., Jordan, T. R., Thomas, S. M. (2009). The recognition of mental states from dynamic and static facial expressions. *Visual Cognition*, *17*, 1271-1286.
- Barker, J. et al. (2012). Dynamic and static facial expressions of emotion are recognized with equal efficiency. *Journal of Vision*, doi: 10.1167/12.9.966

- Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind*. Boston: MIT Press/Bradford Books.
- Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes" ? Evidence from normal adults and adults with autism or Asperger syndrome. *Visual Cognition*, 4, 311-331.
- Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 373-379.
- Bassili, J. N. (1979). Emotion Recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, 37, 2049-2058.
- Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, 23, 223-229.
- Besel, L. D. S., & Yuille, J. C. (2010). Individual differences in empathy: The role of facial expression recognition. *Personality and Individual Differences*, 49, 107-112.
- Bindemann, M., Scheepers, C., & Burton, M. A. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision*, 9(2) : 7, doi : 10.1167/9.2.7.
- Blais, C., Jack, R., Scheepers, C., Fiset, D., & Caldara, C. (2008). Culture shapes how we look at faces. *PLoS ONE*, 3(8) :e3022.

- Borod, J. C., Obler, L. K., Erhan, H. M., Grunwald, I. S., Cicero, B. A., Welkowitz, J., Santschi, C., Agosti, R. M., & Whalen, J. R. (1998) Right hemisphere emotional perception: evidence across multiple channels. *Neuropsychology, 12*, 446–458
- Bould, E., Morris, N., & Wink, B. (2008). Recognizing subtle emotional expressions: The role of facial movements. *Cognition & Emotion, 22*, 1569-1587.
- Bruce, V. and Young, A. (1986) Understanding face recognition. *British Journal of Psychology, 77*, 305–327.
- Bukach, C. M. Bub, D. N. Gauthier, I. Tarr, M. J. (2006). Perceptual expertise effects are not all or none: Spatially limited perceptual expertise for faces in a case of prosopagnosia *Journal of Cognitive Neuroscience, 18*, 48–63.
- Bukach, C. M., LeGrand, R., Kaiser, M., Bub, D. N., & Tanaka, J. W. (2008). Preservation of mouth region processing in two cases of prosopagnosia. *Journal of Neuropsychology, 2*, 227-244.
- Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception & Psychophysics. 72*, 1444-1449
- Caldara, R., Fiset, D., Blais, C., Schyns, P., Scheepers, C., & Meyers, E. (2007). Clarifying the nature of facial identity and facial expression representations with an acquired case of prosopagnosia. *Journal of Vision*, doi : 10.1167/7.9.628.
- Caldara, R., Rossion, B., Mayer, E., Smith, M., Gosselin, F. & Schyns, P. G. (2005). Does prosopagnosia take the eyes out from faces? Evidence for a

- defect in the use of diagnostic facial information in a single-case. *Journal of Cognitive Neuroscience*, *17*, 1-15.
- Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology : Human Perception and Performance*, *26*, 527-551.
- Campbell, R. (1978). Asymmetries in interpreting and expressing a posed facial expression. *Cortex*, *14*, 327-342.
- Canli, T. (1999) Hemispheric asymmetry in the experience of emotion: a perspective from functional imaging. *Neuroscientist*, *5*, 201–207.
- Carroll, J. M., Russell, J. A. (1997). Facial expressions in Hollywood's portrayal of emotion. *Journal of Personality and Social Psychology*, *72*, 164-176.
- Chaurasia, B. D., and Goswami, H. K. (1975). Functional asymmetry in the face. *Acta Anatomica*, *91*, 154-160.
- Cheek, J.M., & Buss, A.H. (1981). Shyness and sociability. *Journal of Personality and Social Psychology*, *41*, 330-339.
- Chiller-Glaus, S. D., Schwaninger, A., Hofer, F., Kleiner, M. & Knappmeyer, B. (2011). Recognition of emotion in moving and static composite faces. *Swiss Journal of Psychology*, *70*, 233-240.
- Coleman, J. (1949). Facial expressions of emotion. *Psychological Monographs*, *63*, 1–296.

- Corden, B., Chilvers, R., & Skuse, D. (2008). Avoidance of emotionally arousing stimuli predicts social-perceptual impairment in Asperger's syndrome. *Neuropsychologia*, *46*, 137–147.
- Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Attention, Perception, & Psychophysics*, *58*, 602-612.
- Cunningham, D. W., Kleiner, M., Bühlhoff, H. H., & Wallraven, C. (2004). The components of conversational facial expressions. Proceedings of the 1st Symposium on Applied perception in graphics and visualization, doi:10.1145/1012551.1012578.
- Darwin, C. (1965). The expression of the emotions in man and animals. Chicago: University of Chicago Press. (Original work published 1872).
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85.
- De Gelder, B., van Honk, J., & Tamietto, M. (2011). Emotion in the brain: of low roads, high roads and roads less travelled. *Nature Review Neuroscience*, *12*, doi:10.1038/nrn2920-c1.
- Deubel, H. & Schneider, W. X. (2003). Delayed saccades, but not delayed manual aiming movements, require visual attention shifts. *Annals of the New York Academy of Sciences*, *1004*, 289-296.
- De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 545-559.

- De Valois, K. K., De Valois, R. L., & Yund, E. W. (1979). Responses of striate cortex cells to grating and checkerboard patterns. *The Journal of Physiology*, *291*, 483-505.
- Diamond, R. & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107-117.
- Douglas-Cowie, E., Cowie, R., & Schröder, M. (2000). A new emotion database - considerations, sources and scope. *Proceedings ISCA ITRW on Speech and Emotion*, 39-44.
- Dunlap, K. (1927). The role of the eye muscles and mouthmuscles in the expression of emotions. *Genetic Psychology Monographs*, *2*, 199-233.
- Eimer, M. & Holmes, A. (2002). An ERP study of the time course of emotional face processing. *NeuroReport*, *13*, 427-431.
- Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes : Scanning emotional facial expressions. *Emotion*, *11*, 860-865.
- Ekman, P. (1972). Universals and cultural differences in facial expression of emotion. In J. Cole (Ed). *Nebraska Symposium on Motivation*, Lincoln :University of Nebraska Press.
- Ekman, P. (1982). Methods for measuring facial action. In Scherer, K. R. & Ekman, P. (Eds.), Handbook of methods in nonverbal behavior research. New York : Cambridge University Press, p. 45-135.
- Ekman, P. & Bressler, J. (1964). In Ekman, P. *Progress report to National Institute of Mental Health*. Bethesda, Maryland.

- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124-129.
- Ekman, P. & Friesen, W. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto.
- Ekman, P., Friesen, W., Hager, J. (2002). *Facial Action Coding System* (2nd ed.). Salt Lake City, UT: Research Nexus eBook.
- Ekman, P. & Rose, D. (1965). In Ekman, P. *Progress report to National Institute of Mental Health*. Bethesda, Maryland.
- Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition : a meta-analysis. *Psychological Bulletin*, 128, 203-235.
- Ellis, A. W., & Brysbaert, M. (2010). Split fovea theory and the role of the two hemispheres in reading : A review of evidence. *Neuropsychologia*, 48, 353-365.
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, 99, 9602-9605.
- Findlay, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research*, 22, 1033–1045.
- Findlay, J. M., & Gilchrist, I. D. (1997). Spatial scale and saccade programming. *Perception*, 26, 1159–1167.

- Frois-Wittman, J. (1930). The judgement of facial expression. *Journal of Experimental Psychology*, *13*, 113–151.
- Furl, N., van Rijsbergen, N. J., Treves, A., Friston, K. J., & Dolan, R. J. (2007). Experience-dependant coding of facial expression in superior temporal sulcus. *Proceedings of the National Academy of Science*, *104*, 13485-13489.
- Gaspar, C. M., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research*, *48*, 2817-2826.
- Gauthier, I., Klaiman, C., & Schultz, R. T. (2009). Face composite effects reveal abnormal face processing in Autism spectrum disorders. *Vision Research*, *49*, 470-478.
- Godjin, R. & Theeuwes, J. (2003). Parallel allocation of attention prior to the execution of saccade sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 882-896.
- Goffaux, V. & Rossion, B. (2006). Faces are « spatial »—holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology : Human Perception and Performance*, *32*, 1023-1039.
- Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, *39*, 3537-3560.

- Gold, J. M., Mundy, P. J., & Tjan, B. S. (2012). The perception of a face is no more than the sum of its parts. *Psychological Science, 23*, 427-434.
- Gosselin, F. & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition. *Vision Research, 41*, 2261-2271.
- Grimm, M., Dastidar, D., & Kroschel, K. (2006). Recognizing emotions in spontaneous facial expressions. *Proceedings of the International Conference on Intelligent Systems and Computing (ISYC)*.
- Gross, T. F. (2008). Recognition of immaturity and emotional expressions in blended faces by children with autism and other developmental disabilities. *Journal of Autism and Developmental Disorders, 38*, 297-311.
- Grossman, E.D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron, 35*, 1167–1175.
- Hainline, L. (1978). Developmental changes in visual scanning of face and non-face patterns by infants. *Journal of Experimental Child Psychology, 25*, 90-115.
- Haith, M., Bergman, T., & Moore, M. (1977). Eye contact and face scanning in early infancy. *Science, 198*, 853-855.
- Hammal, Z., Gosselin, F., & Fortin, I. (2009). How efficient are the recognition of dynamic and static facial expressions? *Journal of Vision*, doi:10.1167/9.8.499.
- Hanawalt, N. (1944). The role of the upper and lower parts of the face as the basis for judging facial expressions: II. In posed expressions and “candid camera” pictures. *Journal of General Psychology, 31*, 23–36.

- Harms, M., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychological Review, 20*, 290-322.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science, 4*, 223–233.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51*, 59-67.
- Henderson, J.M., Williams, C. C., Falk, R. J. (2005) Eye movements are functional during face learning. *Memory and Cognition, 33*, 98–106.
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology, 40*, 129–141.
- Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science, 19*, 998-1006.
- Humphreys, K., Avidan, G. and Behrmann, M. (2007). A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. *Experimental Brain Research, 176*, 2, 356-373.
- Humphreys, K., Minshew, N., Leonard, G. L., & Behrmann, M. (2007). A fine-grained analysis of facial expression processing in high-functioning adults with autism. *Neuropsychologia, 45*, 685-695.

- Hunt, W. A. (1941). Recent developments in the field of emotion. *Psychological Bulletin*, 38, 249-276.
- Howell, R. J. & Jorgenson, E. C. (1970). Accuracy of judging emotional behavior in a natural setting: A replication. *Journal of Social Psychology*, 81, 269-270.
- Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing : The core of social cognition. *Neuroscience and Behavioral Reviews*, 33, 843-863.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P., & Caldara, R. (2009). Cultural Confusions Show That Facial Expressions are Not Universal. *Current Biology*, 19, 1543-1548.
- Jack, R., Caldara, R., & Schyns, P. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotions. *Journal of Experimental Psychology: General*, 141, 19-25.
- Jack, R. E., Garrod, O. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expression of emotion are not culturally universal. *Proceedings of the National Academy of Science*, 109, 7241-7244.
- Janik, S.W., Wellens, A.R., Goldberg, M.L., Dell'Osso, L.F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual & Motor Skills*, 47, 857– 858.
- Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, 6, 766-774.
- Johnson, M. H. & Morton, J. (1991) *Biology and Cognitive Development: The*

Case of Face Recognition (Blackwell, Oxford).

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. L. A. Baddeley (Ed.), *Attention and Performance IX*. Hillsdale, NJ: Erlbaum.

Kätsyri, J., & Sams, M. (2008). The effect of dynamics on identifying basic emotions from synthetic and natural faces. *International Journal of Human-Computer Studies*, *66*, 233-242.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage*, *18*, 156-168.

Kumar, D., & Srinivasan, N. (2011). Emotion perception is mediated by spatial frequency content. *Emotion*, *11*, 1144-1151.

LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, *13*, 1023-1033.

Landis, C. (1924). Studies of emotional reactions : II. General behavior and facial expression. *Journal of Comparative Psychology*, *4*, 447-509.

Langner, O., Becker, E. S., & Rinck, M. (2009). Social Anxiety and anger identification: Bubbles reveal differential use of facial information with low spatial frequencies. *Psychological Science*, *20*, 666-670.

Lanzetta, J. T., & Kleck, R. E. (1970). Encoding and de- coding of nonverbal affect in humans. *Journal of Personality and Social Psychology*, *16*, 12-19.

- Lee, J., Gosselin, F., Wynn, J. K., & Green, M. F. (2011). How do schizophrenia patients use visual information for decoding facial emotion? *Schizophrenia Bulletin*, *37*, 1001-1008.
- Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, *4*, 501-506
- Lischke, A., Berger, C., Prehn, K., Heinrichs, M., Herpertz, S. C., & Domes, G. (2012). Intranasal oxytocin enhances emotion recognition from dynamic facial expressions and leaves eye-gaze unaffected. *Psychoneuroendocrinology*, *37*, 475-481.
- Maffei, L. & Fiorentini, A. (1973). The visual cortex as a spatial frequency analyser. *Vision Research*, *13*, 1255-1267.
- McCrae, R. R., Costa, P. T., Jr., & Martin, T. A. (2005). The NEO-PI-3: A more readable Revised NEO Personality Inventory. *Journal of Personality Assessment*, *84*, 261-270.
- Maurer, D. (1985). Infants' perception of facedness. In Field, T. & Fox, M. (Eds). *Social perception in infants*, Norwood, NJ: Ablex.
- Matsuzaki, N. & Sato, T. (2008). The perception of facial expressions from two-frame apparent motion. *Perception*, *37*, 1560-1568.
- Mesquita, B., and Frijda, N.H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin*, *112*, 179-204.
- Michel, C., Rossion, B., Han, J., Chung, C., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, *17*, 608-615.

- Moeller, S. K., Ewing Lee, E. A., Robinson, M. D. (2011). You never think about my feelings: Interpersonal dominance as a predictor of emotion decoding accuracy. *Emotion, 11*, 816-824.
- Motley, M., & Camden, C. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Speech Communication, 52*, 1–22.
- Naab, P. J., & Russell, J. A. (2007). Judgments of emotion from spontaneous facial expressions of New Guineans. *Emotion, 7*, 736-744.
- Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Cognitive Brain Research, 12*, 225-231.
- Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research, 39*, 3824-3833
- Nummenmaa, T. (1964). The language of the face (Jyvaskyla studies in education, psychology, and social research). Jyvaskyla, Finland.
- Nusseck, M., Cunningham, D. W., Wallraven, C., & Bühlhoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision, 8*, 1-23
- Pelphrey, K.A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B.D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders, 32*, 249–261.
- Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala : from

a « low road » to a « many roads » of evaluating biological significance.

Nature Reviews Neuroscience, 11, 773-783.

Plutchik, R. (1962). *The emotions : facts, theories, and a new model*. New York: Random House.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.

Puce, A., Allison, T., Bentin, S., Gore, J.C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 18, 2188–2199.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.

Richardson, C. K., Bowers, D., Bauer, R. M., Heilman, K. M., & Leonard, C. M. (2000). Digitizing the moving face during dynamic displays of emotion. *Neuropsychologia*, 38, 1028–1039.

Ross, E. D., & Pulusu, V. K. (in press). Posed versus spontaneous facial expressions are modulated by opposite cerebral hemispheres. *Cortex*.

Rossion, B. & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation : Evidence from the face composite illusion. *Journal of Vision*, 8, 1-13. doi : 10.1167/8.4.3

Rossion, B., Kaiser, M.D., Bub, D., & Tanaka, J.W. (2009). Is the loss of diagnosticity of the eye region of the face a common aspect of acquired prosopagnosia? *Journal of Neuropsychology*, 3, 69-78.

Ruckmick, C. A. (1921). A preliminary study of the emotions. *Psychological*

Monographs, 30, 30-35.

- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*, 102-141.
- Sackeim, H. A., and Gur, R. C. (1978). Lateral asymmetry in intensity of emotional expression. *Neuropsychologia, 16*, 473-481.
- Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cognitive Brain Research, 20*, 81-91.
- Schyns, P. G., Bonnar, L. & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science, 13*, 402-409.
- Shillcock, R., Ellison, T. M. & Monaghan, P. (2000). Eye-fixation behavior, lexical storage and visual word recognition in a split processing model. *Psychological Review, 107*, 824-851.
- Sebe, N., Lew, M. S., Cohen, I., Sun, Y., Gevers, T., & Huang, T. S. (2004). Authentic Facial Expression Analysis. *Proceedings of the 6th International Conference on Automatic Face and Gesture Recognition, 517-522.*
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14*, 391-396.
- Simon, D., Craig, K. D., Gosselin, F., Belin, P., & Rainville, P. (2008). Recognition and discrimination of prototypical dynamic expressions of

- pain and emotions. *Pain*, *135*, 55-64.
- Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present and future. *Trends in Cognitive Sciences*, *9*, 16-20.
- Smith, M., Cottrell, G., Gosselin, F. & Schyns, P. G. (2005). Transmitting and decoding facial expressions of emotions. *Psychological Science*, *16*, 184-189.
- Smith, F. W., & Schyns, P. G. (2009). Smile through your fear and sadness. Transmitting and identifying facial expression signals over a range of viewing distances. *Psychological Science*, *20*, 1202-1208.
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007a). Abnormal use of facial information in high-functioning autism. *Journal of Autism and Developmental Disorders*, *37*, 929–939.
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007b). Analysis of face gaze in autism using “Bubbles. *Neuropsychologia*, *45*, 144–151.
- Sullivan, S., & Ruffman, T. (2004) Emotion recognition deficits in the elderly. *International Journal of Neuroscience*, *114*, 403–432.
- Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *Journal of Gerontology: Psychological Sciences*, *62B*, 53-60.
- Suzuki, A., Hoshino, T., & Shigemasu, K. (2006). Measuring individual differences in sensitivities to basic emotions in faces. *Cognition*, *99*, 327-353.
- Tamietto, M. & de Gelder, B. (2010). Neural bases of the non-conscious

- perception of emotional signals. *Nature Reviews Neuroscience*, *11*, 697-709.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and whole in face recognition. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *46*, 225-245.
- Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion : Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research*, *1284*, 100-115.
- Tsuchiya, N., Kawasaki, H., Oya, H., Howard, M. A., & Adolphs, R. (2008). Decoding face information in time, frequency and space from direct intracranial recordings of the human brain. *PLoS ONE*, *3*(12): e3892. doi:10.1371/journal.pone.0003892.
- Vinette, C., Gosselin, F. & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes! *Cognitive Science*, *28*, 289-301.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, *6*, 624-631.
- Wagner, H. L. (1990). The spontaneous facial expression of differential positive and negative emotions. *Motivation and Emotion*, *14*, 27-43.
- Wagner, H. L., Lewis, H., & Ramsay, S., Krediet, I. (1992). Prediction of facial displays from knowledge of norms of emotional expressiveness.

Motivation and Emotion, 16, 347–362.

Wagner, H., MacDonald, C., & Manstead, A. (1986). Communication of individual emotions by spontaneous facial expressions. *Journal of Personality and Social Psychology, 50, 737–743.*

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C. Arguin, M., Tanaka, J., Bub, D. & Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36, 122-135.*

Yarbus, A. L. (1967). *Eye Movements and Vision*, New York: Plenum. (Originally published in Russian 1962).

Young, M. E., & Brunet, P. M. (2011). It's about time: The influence of sociability—but not shyness—on the categorization of facial expressions of emotion. *Journal of Individual Differences, 32, 65-73.*

Young, A. W., Hellawell, D., & Hay, D. C., (1987). Configurational information in face perception, *Perception, 16, 747 – 759.*

Yik, M., Meng, Z., & Russell, J. A. (1998). Adults' freely produced emotion labels for babies' spontaneous facial expressions. *Cognition & Emotion, 12, 723–730.*