



**HAL**  
open science

# Micro-expressions audio-visuelles dans la communication expressive : enjeux pluri-culturels

Yuko Sasa

► **To cite this version:**

Yuko Sasa. Micro-expressions audio-visuelles dans la communication expressive : enjeux pluri-culturels. Sciences de l'Homme et Société. 2012. dumas-00709396

**HAL Id: dumas-00709396**

**<https://dumas.ccsd.cnrs.fr/dumas-00709396>**

Submitted on 18 Jun 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# **Micro-expressions audio-visuelles dans la communication expressive : enjeux pluri-culturels**

**Nom : SASA  
Prénom : Yuko**

**UFR LLASIC - SCIENCES DU LANGAGE**

---

Mémoire de master 2 recherche - 30 crédits - Mention Sciences du Langage

Spécialité : Industries de la Langue (IDL)

Parcours : Traitement Automatique de la Langue Écrite et de la Parole (TALEP)

Sous la direction de Véronique AUBERGÉ

Année universitaire 2011-2012





# **Micro-expressions audio-visuelles dans la communication expressive : enjeux pluri-culturels**

**Nom : SASA  
Prénom : Yuko**

**UFR LLASIC - SCIENCES DU LANGAGE**

---

Mémoire de master 2 recherche - 30 crédits - Mention Sciences du Langage

Spécialité : Industries de la Langue (IDL)

Parcours : Traitement Automatique de la Langue Écrite et de la Parole (TALEP)

Sous la direction de Véronique AUBERGÉ

Année universitaire 2009-2010

## Avant-propos

Le langage est un phénomène naturel de l'humain qui suscite depuis très longtemps de nombreuses interrogations, tantôt sur son origine, tantôt sur sa multiplicité et la complexité de ses formes. Le mythe de la Tour de Babel<sup>1</sup> est un exemple des traces socioculturelles révélatrices de cette réflexion permanente autour du langage. La récurrence des efforts a alors donné la parole à des statues ou à des automates et cela depuis au moins le début de l'Antiquité (Pettorino et al, 2010)

Au fil des années, cette réflexion, reste l'une des questions fondamentales, voire existentielles, adressée par l'homme. Elle a été abordée par les courants religieux, socioculturels, philosophiques puis scientifiques. Elle transparaît dans certains thèmes philosophiques, la rhétorique, les règles du « bien parler », l'articulatoire, la dialectologie, l'ethnologie...et bien d'autres encore, jusqu'à ce que la discipline scientifique des Sciences du Langage rassemble, dans un contour qui semble évoluer à l'écoute des autres disciplines scientifiques auxquelles elle fait appel, ces différentes approches de manifestations, descriptions, représentations et modélisations de cet « objet » complexe qu'est le langage.

Ainsi cette discipline commence à s'affirmer, notamment avec les débuts de la linguistique moderne dont nous considérons F. de Saussure comme le fondateur. Progressivement des noyaux de sous-disciplines se sont créés à différents niveaux de granularité. Bien que les apports soient nombreux, ces disciplines ont un point de vue assez constant mais également rigide sur les langues. Nous pouvons d'ailleurs résumer cette vision très succinctement en ces termes : des unités s'assemblent pour en former de plus grandes à différents niveaux de finesse. Sur ces unités viennent se greffer des facteurs socioculturels et méta langagiers (Saussure, 1995).

Le Traitement Automatique de la Langue et de la Parole est à la fois un champ qui fait évoluer l'instrumentation du langage (qui s'inscrit dans l'une des traditions des

---

<sup>1</sup> « Tout le monde se servait d'une même langue et des mêmes mots. Comme les hommes se déplaçaient à l'orient, ils trouvèrent une vallée au pays de Shinéar et ils s'y établirent. Ils se dirent l'un à l'autre : Allons ! Faisons des briques et cuisons-les au feu ! La brique leur servit de pierre et le bitume leur servit de mortier. Ils dirent : Allons ! Bâtissons-nous une ville et une tour dont le sommet pénètre les cieux ! Faisons-nous un nom et ne soyons pas dispersés sur toute la terre ! Or Yahvé descendit pour voir la ville et la tour que les hommes avaient bâties. Et Yahvé dit : Voici que tous font un seul peuple et parlent une seule langue, et tel est le début de leurs entreprises ! Maintenant, aucun dessein ne sera irréalisable pour eux. Allons ! Descendons ! Et là, confondons leur langage pour qu'ils ne s'entendent plus les uns les autres. Yahvé les dispersa de là sur toute la face de la terre et ils cessèrent de bâtir la ville. Aussi la nomma-t-on Babel, car c'est là que Yahvé confondit le langage de tous les habitants de la terre et c'est de là qu'il les dispersa sur toute la face de la terre ». (Genèse, 11, La Bible de Jérusalem). Extrait tiré du site :<http://crdp.ac-paris.fr/parcours/fondateurs/index.php/category/babel> (consulté le 28/04/2012)

sciences du langage qui est l'expérimentation), et bien entendu une ingénierie qui permet de développer des applications par un outillage informatique. Ces applications mises au service des humains communicants, permettent en retour d'évaluer « écologiquement » la modélisation théorique qui sous-tend ces outils, et de s'assurer de la pertinence socio-culturo-économique de ces applications.

Ainsi, l'ingénierie des langues peut par exemple apporter un éclairage différent du problème de l'ambiguïté qui peut être un verrou pour certaines des sous-disciplines des sciences du langage : en « remettant » le langage dans une des réalités (nouvelles) de son usage, « l'usager », acteur direct de sa langue, pointe parfois sur un traitement plus global au bout duquel l'ambiguïté (intrinsèquement modulaire dans les modélisations linguistiques) n'est plus visible.

Dans les applications de la synthèse vocale par exemple, il a été montré que des éléments « extralinguistiques », expressifs ou stylistiques, en particulier véhiculés par la « prosodie », sont d'une grande pertinence pour le succès de l'interaction, ancrant la personnalité du sujet, ses motivations et ses attentes, dans son rôle sociétal et sa culture, alors qu'une place mineure leur est donnée dans la linguistique. Ces « micro-éléments marginaux » pourraient être plus fondamentaux que ce que l'on a montré d'eux en sciences du langage jusqu'ici. Ils pourraient même être à l'origine de la naissance de la communication parlée (Vanpé, 2012). Mais alors comment étudier ces objets sans qu'ils soient « cachés » par les « macro-éléments » qui implémentent en particulier les fonctions lexico-morpho- syntaxiques ? A quoi peuvent servir ces micro-éléments et comment peut-on les observer, les analyser et les modéliser ?

Les questions qui sont posées dans ce préambule ne trouveront bien sûr pas de réelles réponses ici, mais le fait de les poser est bien l'enjeu de ce mémoire qui traite des « micro-expressions dans la communication parlée », ces objets considérés encore souvent comme « marginaux ».

En particulier dans ce mémoire, je porterai mes propos à travers une étude biculturelle qui mettra en opposition la perception de « micro-expressions de la parole » par des Français et des Japonais. En effet, si des différences communicatives sont notables sur cette analyse comparative, ne serait-ce pas là un indice que ces objets relèvent bien de

la linguistique puisqu'ils répondraient à la lecture d'objets codés, portant une valeur communicative différente ?

J'espère donc que ce mémoire participera à ce champ émergeant, qui reprend celui plus ancien de la phonostylistique par exemple, et s'insère dans l'« Affective Computing »<sup>2</sup> ou les signaux sociaux<sup>3</sup>. En particulier, nous souhaitons nous inscrire dans une perspective nouvelle du traitement de la langue parlée envisageant par exemple de nouvelles pistes de modélisation de la parole interactionnelle, pour sa simulation, son apprentissage L2 ou encore dans des technologies d'aide à la personne.

Bonne lecture à tous !

---

<sup>2</sup> Ce domaine est largement décrit dans les conférences ACII

<sup>3</sup> Cette thématique est principalement abordée par les travaux présentés sur le site SSPNet : <http://sspnet.eu/> (consulté en septembre 2011)

## Remerciements

Je tiens tout d'abord à remercier Véronique Aubergé, ma directrice de mémoire qui me proposa un sujet passionnant, me donna l'opportunité de retravailler avec ma langue maternelle, et enfin qui fut d'une grande aide dans les moments difficiles.

De même, merci à tous mes enseignants m'ayant accompagné durant ces deux années d'études (voire plus pour certains) ainsi que les anciens étudiants de la filière qui ont su être de bons conseils dans l'élaboration de ce mémoire.

Un grand merci également à Albert, pour sa patience et son aide sans qui je n'aurais jamais réussi à créer et analyser le matériel de mon mémoire.

Merci à toute ma famille qui a supporté ma mauvaise humeur, et surtout à ma mère qui a pris le temps de me relire sur mes productions japonaises.

Mes remerciements à tous mes camarades de classe, une promotion formidable avec laquelle j'ai pu partager beaucoup. Spéciale dédicace pour Lucie avec qui nous nous sommes serrées les coudes pour venir à bout de nos mémoires.

A Yan et à Manon qui ont su être des présences rassurantes durant la rédaction de mon mémoire.

Et enfin, quelques mots en japonais pour tous les participants de mon expérience :

実験にご参加して頂いた皆様、誠にありがとうございました。

# Sommaire

<b>PARTIE 1 ETAT DE L'ART .....</b>	<b>11</b>
CHAPITRE 1 - PRESENTATION DU LABORATOIRE.....	12
CHAPITRE 2 – SURVOL DE LA COMMUNICATION EXPRESSIVE .....	14
<i>Réflexions philosophiques sur les émotions</i> .....	14
<i>Les principales théories des émotions dans le domaine de la psychologie</i> .....	18
<i>Conclusion du Chapitre 2</i> .....	24
CHAPITRE 3 - LES TRAVAUX PRECEDENTS A L'ORIGINE DU MEMOIRE .....	25
<i>Les motivations d'études sur les micro-expressions et terminologie</i> .....	29
<i>Premières observations et annotations de corpus</i> .....	30
<i>Conclusion du Chapitre 3</i> .....	34
<b>PARTIE 2 EXPERIENCES PERCEPTIVES.....</b>	<b>35</b>
CHAPITRE 4 – TEST DE PERCEPTION DE MICRO-EXPRESSIONS PAR DES SUJETS JAPONAIS.....	36
<i>Une manipulation répondant aux perspectives d'études antérieures</i> .....	36
<i>Protocole et outil expérimental</i> .....	37
<i>Contexte de recueil de données</i> .....	45
<i>Profils des sujets japonais</i> .....	48
<i>Conclusion sur le chapitre 4</i> .....	57
<b>PARTIE 3 ANALYSE DES RESULTATS EXPERIMENTAUX .....</b>	<b>58</b>
CHAPITRE 5 – ANALYSE DES RESULTATS JAPONAIS .....	59
<i>Matrices de confusions et analyses statistiques</i> .....	59
<i>Les degrés de confiance des réponses des sujets</i> .....	68
<i>La perception culturelle des micro-expressions</i> .....	72
<i>Conclusion du chapitre 5</i> .....	73
CHAPITRE 6 – ANALYSE COMPARATIVE FRANCO-JAPONAISE.....	75
<i>Analyse comparative sur les taux de bonnes réponses générales</i> .....	75
<i>Explication de l'analyse comparative étiquette par étiquette</i> .....	76
<i>Observations sur les stimuli du sujet T</i> .....	77
<i>Observations sur les stimuli du sujet S</i> .....	78
<i>Observations sur les stimuli du sujet M</i> .....	80
<i>Observations sur les stimuli du sujet J</i> .....	81
<i>Observations sur les stimuli du sujet N</i> .....	83
<i>Observations sur les stimuli du sujet R</i> .....	84
CHAPITRE 7 – LES OBSERVATIONS DE SYNERGOLOGIE .....	87
<i>Aperçu de la synergologie</i> .....	87
<i>Recueil d'analyse de synergologie</i> .....	90
<i>Résultats des questionnaires</i> .....	93
<i>Conclusion du Chapitre 7</i> .....	102
<b>PARTIE 4 PERSPECTIVES ET ENJEUX APPLICATIFS .....</b>	<b>103</b>
CHAPITRE 8 – BILAN DE L'ETUDE ET PERSPECTIVES .....	104
<i>Les problèmes et difficultés mis en évidence</i> .....	104
<i>Les perspectives d'amélioration des problèmes soulignés</i> .....	106

## Introduction

Dans le domaine du Traitement Automatique de la Langue (TAL), qui est une des nombreuses branches des Sciences du Langage, les traitements de l'écrit et de la parole relèvent de communautés spécifiques, même si bien entendu, elles se rencontrent souvent. Pour ce mémoire, j'ai choisi de traiter des objets qui se manifestent dans la modalité acoustique. Ces éléments n'ont toutefois pas forcément de représentation écrite (seules les interjections sont orthographiées en éléments pré-lexicaux). (Vanpé et Aubergé, 2012).

Pourtant, la frontière entre l'écrit et l'oral, principalement motivée par leurs usages, est de plus en plus fine. L'écrit, qui était avant tout un moyen de conserver des informations véhiculées par la parole, voit de grosses évolutions dans ses codes et ses modalités de propagation. Nous sommes d'ailleurs témoins de l'émergence de nouveaux genres textuels dits « oralisés » tels que nous les voyons dans les SMS, les mails et les réseaux sociaux qui ont une fonction communicative très marquée. De part ces nouvelles habitudes cognitivo-socio-culturelles, je pense qu'il est d'autant plus important de travailler sur ces éléments car ils correspondent aux demandes et aux besoins réels des technologies actuelles. De plus, les aspects cognitifs et sociaux qui s'ajoutent à ces interactions humaines sont des sujets qui me passionnent tout particulièrement.

La possibilité d'accéder à des informations qui sont non lexicalisées bien qu'elles jouent un rôle primordial dans l'interaction, peut parfois permettre de désambiguïser certains propos ou d'ajuster le déroulement de l'interaction. Par exemple, si une personne soupire au cours d'une interaction qui perdure, nous avons la possibilité de l'écourter ou de changer de sujet. Ce soupir qui n'est donc nullement lexicalisé a un rôle communicatif important. Or ce type d'objets, nous en retrouvons énormément dans une interaction. D'ailleurs si nous adoptons le point de vue d'un seul locuteur, en supposant que le temps de parole est équilibré, le temps de production lexicalisée ne représenterait au grand maximum que la moitié du temps de l'interaction. Néanmoins nous continuons à communiquer entre nos tours de parole (« talking listener ») et ponctuons également nos propos lexicaux d'autres informations aussi bien visuelles (à tout âge) (Colletta et al, 2009) qu'auditives. De plus, sur nos productions lexicalisées elles-mêmes, nous retrouvons toujours les indices du non lexical tels que la prosodie.

Il y a donc un aspect important du langage qui doit être étudié dans le non lexical ou parfois dit, peut-être à tort, non verbal. Dans ce mémoire, nous nommerons ces éléments : les « micro-expressions de la parole », (Magno et al, 1997 ; Matsumoto 2008 ; Loyau et al, 2006). Dans le cadre de cette étude, ces objets désigneront uniquement les parties non lexicales, une focalisation volontaire pour simplifier le problème dont le bruit produit par la phonologie et la syntaxe complique les tâches d'observation.

Mais comment peut-on définir les « micro-expressions de la parole » ? Quelles dimensions culturelles et donc linguistiques caractérisent ces objets ? Comment interviennent ces paramètres ?

A terme, nous voulons montrer une nouvelle perspective d'observation de la langue et du langage, tout en illustrant les enjeux applicatifs s'appuyant sur les micro-expressions.

Ce mémoire étudiera donc cette problématique en quatre parties. Une première sera dédiée à un état de l'art qui essaiera de définir ce que sont les « micro-expressions de parole » par la description de leurs différents aspects, à travers des points de vue disciplinaires variés. Une seconde partie présentera les manipulations expérimentales qui nous ont permis d'observer ces objets par le biais de protocoles de reconnaissance de micro-expressions en contraste Français/Japonais. Le choix du japonais en contraste avec le français s'avère intéressant du fait de l'usage important de la qualité de voix et des micro-expressions, en particulier d'interjections, très riches et très fréquentes dans cette langue. (Campbell, 2007a ; Campbell, 2007b ; Matsumoto, 1988 ; Ishi, 2012 ; Maekawa, 2004). Sur les données de cette expérience, sera mise en œuvre une analyse intégrative des données s'appuyant à la fois sur des aspects perceptifs interculturels (objectivés par des calculs statistiques) ainsi que des aspects acoustiques et visuels (en particulier synergologiques pour le visuel). Ensuite nous verrons comment la synergologie (une approche empirique récente de « décodage » des micro-expressions faciales) qui souhaite se confronter à la falsification scientifique, traite et analyse les stimuli que nous avons donnés à percevoir aux sujets naïfs. Enfin nous terminerons par une dernière partie qui mettra en exergue les perspectives d'études ultérieures ainsi que des enjeux applicatifs dérivant de l'étude de « micro-expressions de parole ».

# **Partie 1**

## **Etat de l'art**

## Chapitre 1 - Présentation du laboratoire<sup>4</sup>

Ce mémoire utilisera les termes de « micro-expressions de la parole » ainsi que de « micromots » pour désigner les objets que nous voulons observer puisqu'il n'existe pas de véritable consensus terminologique. Ce sont les études à venir qui permettront de savoir ce qu'ils représentent véritablement ainsi que la façon dont nous pourrions les appeler, les décrire et enfin comprendre la manière dont ils fonctionnent. Cependant nous ne partons pas de rien.

Les travaux de ce mémoire s'inscrivent directement dans la continuité d'un certain nombre d'études menées au sein du Laboratoire GIPSA-LAB de Grenoble. Ce chapitre 1 décrira donc brièvement le laboratoire dans lequel j'ai travaillé cette année.

Le GIPSA-LAB (Grenoble Image Parole Signal Automatique) est une unité de recherche de Grenoble créée en 2007 s'appuyant sur la mixité des travaux du CNRS (Centre National de la Recherche Scientifique), du groupe Grenoble-INP (Institut Polytechnique de Grenoble, composé de 6 écoles d'ingénieurs), de l'université Joseph Fourier et de l'université Stendhal. Cette unité de recherche est conventionnée avec l'Inria (Institut national de recherche en informatique et en automatique créé en 1967 sous le nom de IRIA et devient INRIA en 1979 pour enfin changer d'identité graphique en 2011), l'Observatoire de Grenoble et l'université Pierre Mendès France. Les travaux qu'il engendre s'inscrivent dans le développement des systèmes automatiques ainsi que le traitement de la parole, du signal et des images. GIPSA est également reconnu au niveau international pour ces travaux de recherches fondamentales et finalisées des signaux et des systèmes complexes à visée stratégique dans les domaines de l'énergie, de l'environnement, de la communication, des systèmes intelligents, de la santé et de l'ingénierie linguistique, d'où son fort lien avec ses partenaires du milieu industriel.

Le laboratoire est composé de 3 départements : « Automatique », « Images et Signal », « Parole et Cognition » parmi lesquels plus de 300 personnes travaillent et où environ 120 thèses sont actuellement en cours au sein de 12 équipes de recherche. Parmi ces équipes, nous avons 160 permanents dont 68 enseignants-chercheurs et 33 chercheurs

---

<sup>4</sup> Cette présentation est très largement inspirée du site officiel du laboratoire GIPSA et reprend les descriptifs qu'on y trouve. La source est consultable en ligne : <http://www.gipsa-lab.grenoble-inp.fr/>

CNRS. Enfin nous trouvons aussi des chercheurs invités, une vingtaine de post-doctorants, des ATER (Attaché Temporaire d'Enseignement et de Recherche). A cela se rajoutent une soixantaine de stagiaires de divers masters dont je fais parti.

Ce mémoire est élaboré au sein du département « Parole et Cognition » de ce laboratoire, de même que les travaux qui seront présentés par la suite. Ce dernier revendique la compétence cognitive et humaine de la communication langagière en s'appuyant sur de nombreuses disciplines telles que le traitement de l'information, la géolinguistique, les neurosciences intégratives ou encore la psychologie cognitive. Ce département est lui-même composé de 4 équipes : GAMA (Gestes phonatoires, Analyse et Modélisation Aéroacoustique), MAGIC (Machines-parlantes, Gestes oro-faciaux, Interaction face-à-face et Communication augmentée), PCMD (Parole, Cerveau, Multimodalité et Développement) et enfin SLD (Systèmes Linguistiques et Dialectologie) au sein de laquelle je me suis investie cette année.

Cette équipe SLD s'intéresse plus particulièrement à l'observation, la description et la modélisation de la linguistique de la parole en situation selon 4 thématiques : la description des systèmes linguistiques à tradition orale ; l'émergence et la dynamique des systèmes linguistiques ; la modélisation comportementale et interculturelle des affects ainsi que la modélisation de la prosodie ; et enfin l'acquisition des systèmes linguistiques. Les approches scientifiques sur lesquelles elle s'appuie et qui nous intéressent particulièrement pour ce mémoire sont notamment : la phonétique expérimentale, la linguistique de terrain, le recueil de corpus (avec la mise en place de plateformes expérimentales type E-Wiz pour recueillir la parole expressive et de la microgestualité), la prosodie comparative ou encore la communication expressive audiovisuelle. Par ailleurs, bien que cela n'ait pas un lien direct avec les travaux de ce mémoire, l'équipe s'attèle à la réalisation de bases de données orales multilingues ou d'atlas linguistiques et mène également ses recherches sur la modélisation de systèmes phonologiques.

Ici, nous nous attacherons plus particulièrement aux éléments de la communication expressive faisant suite à un recueil spécifique de corpus que nous décrirons dans la suite de notre état de l'art.

## **Chapitre 2 – Survol de la communication expressive**

Les micro-expressions sont des formes visuo-acoustiques non langagières dans la communication parlée mais qui sont empreintes de valeurs socio-affectives. Elles sont ainsi contrôlées volontairement ou plus instinctivement.

Ce chapitre 2 sera donc consacré à un très rapide survol des principales réflexions et courants théoriques qui se sont intéressés aux concepts d'émotions et d'affects, des notions qui représentent la valeur informative des micro-expressions. Ce champ s'ouvre désormais à de plus en plus de disciplines et de domaines et cette présentation n'illustrera donc qu'une infime partie des travaux qui ont été réalisés. En revanche, les précurseurs restent les mêmes malgré l'effervescence de l'intérêt qu'on porte à ce domaine. L'exploration et les nouvelles perspectives commencent à peine à émerger, mais s'appuient encore bien largement sur les observations des prédécesseurs, ce que nous allons essayer d'illustrer dans ce chapitre.

### ***Réflexions philosophiques sur les émotions***

Il n'est pas étonnant qu'en travaillant sur les micro-expressions et les émotions nous soyons amenés à revenir sur de vieux concepts philosophiques puisque l'émotion est un sujet qui est resté au cœur des réflexions et des débats, s'étant très tôt vu opposée au pilier de fondation de la pensée philosophique : la raison. Nous verrons dans cette section que les débats philosophiques ont fait évoluer la manière dont nous pouvions appréhender les émotions mais également la perception et le fonctionnement de l'être humain qui est au centre de nos intérêts quand nous travaillons sur les micro-expressions. De fait, cette partie, s'est très largement inspirée de l'article de (C. Solomon, 2000).

### **L'évolution de la grande opposition entre raison et émotion**

Les considérations sur les émotions sont apparues très tôt, dès l'Antiquité avec les « présocratiques » puis Socrate et son élève Platon. Pourtant les émotions sont souvent refoulées en arrière-plan au détriment de la raison. A l'époque, la dichotomie entre émotion et raison s'est vu attribuée une métaphore mettant en jeu la relation entre un maître et son esclave. En effet, la raison était semblable au maître sage et ferme, ayant le contrôle de son esclave, sujet lui à de malfaisantes impulsions émotionnelles dont il fallait éviter tout débordement. Cette métaphore met en évidence deux aspects des émotions. Le

premier est l'infériorité du rôle des émotions. Celles-ci seraient plus bestiales, primitives et moins intelligentes d'où la nécessité de les contrôler à tout prix. C'est d'ailleurs ce qui justifiait l'esclavage aux temps des athéniens (Aristote, 2001). Le second aspect souligné par la dichotomie est la nature même de la raison, quant à savoir si l'émotion en est une composante inférieure ou si l'émotion et la raison sont deux éléments distincts, antagonistes qui composent « l'âme ». Quelques théories se sont essayées à assimiler ces deux notions en une seule entité mais de manière générale, c'est la raison qui l'emporte sur les émotions. Ce qui est sûr c'est que malgré le fait que certains ne veulent s'atteler à comprendre ce que sont les émotions qu'en commençant par en donner une définition, nous ne pouvons en donner qu'une prémisse aux termes des observations faites ci-dessus.

### **Les premières définitions et notion d'éthique**

L'une des premières définitions de l'émotion est donnée par Aristote (384-322 av. J.C) dont la contribution principale sur le sujet a été la taxonomie des émotions. En voici sa définition : « les causes qui font varier les hommes dans leur jugement et ont pour conséquence la peine et le plaisir comme la colère, la pitié, la crainte et toutes les autres émotions de ce genre ainsi que leurs contraires » (Aristote, 2007, II, 2, 1378a). Ce qui est surtout intéressant de noter chez Aristote, c'est le fait qu'il aborde toujours les émotions en leur associant un contexte ayant des considérations éthiques. En effet, pour lui les émotions seraient dotées d'une grande force morale qui peut être entretenue et même activée par la raison. Ainsi nous nous mettons en colère car nous aurions par exemple perçu une injustice qui motive notre morale à laquelle nous pouvons répondre impulsivement. Nous retrouvons cette proximité avec l'éthique chez les Stoïciens puis également au Moyen Âge dans le contexte du christianisme où l'on associait les émotions à des notions de désirs qui préoccupaient les hommes de foi.

### **Les émotions et le corps**

Les premières considérations reliant le corps aux émotions prennent leur origine dans les médecines antiques, avec notamment la formalisation de la théorie des humeurs par Hippocrate (Thivel, 2009), largement reprise dans la période médiévale. Ces lectures ont notamment inspiré Descartes (1596-1650), le « père de la philosophie moderne », grand rationnel, défenseur de la raison et de l'autonomie de l'esprit. Pour lui qui place la raison au-dessus de toutes les considérations « bestiales », le corps et l'esprit sont séparés. Ainsi la raison réside dans l'esprit et les contractions de l'estomac qui peuvent être l'une

des manifestations physiologiques d'une émotion appartiennent au corps. Cependant, même à l'époque, l'interaction qui pouvait y avoir entre corps et esprit était indéniable. Descartes décrit d'ailleurs dans son œuvre *Les passions de l'âme* qu'il existe un contact entre les deux par l'intermédiaire de la glande pinéale qui se situe à la base du cerveau.

### **La source des émotions passant du corps vers l'esprit**

L'un des problèmes avec les théories de Descartes était de considérer que les émotions ne sont pas seulement issues d'une manifestation physique mais qu'elles prennent également source dans des événements cognitifs tels que la perception, les désirs ou les croyances. Ainsi Descartes suggère que les émotions sont à la fois une perception du corps mais également de l'âme, pour expliquer la source des désirs, ainsi que de certaines perceptions tels que les rêves. Donc nous avons un passage de la physiologie vers des considérations cognitives qui auraient un rôle essentiel dans le développement de la sagesse. Les émotions sont alors considérées comme des « passions » pour reprendre la terminologie de Descartes, passions qu'il ne néglige plus dès lors et qui l'incitent même à orienter ces études de telle façon qu'il finit par décrire 6 passions primitives qu'il considère indispensable pour le « bien vivre ».

### **Les considérations philosophiques cognitives des émotions**

Spinoza (1632-1677) qui est un contemporain de Descartes avait lui une toute autre conception que le mathématicien-physicien. En effet, ce dernier s'inscrit dans la lignée des Stoïciens et considère que les émotions sont une forme de pensée percevant le monde et qu'une mauvaise interprétation de ce dernier nous rendrait malheureux. Nous pouvons donc dire que Spinoza est, avec les Stoïciens, l'un des premiers à proposer une théorie cognitive des émotions. Il définit notamment deux types d'émotions : celles qui sont passives et vous laissent des traces de douleurs, de frustrations et d'énervements, et celles qui sont actives qui permettent de nous prendre en main. De plus, bien qu'il existe une dualité entre esprit et corps, ces deux aspects formeraient une unité indissociable.

D. Hume (1711-1776) qui s'est fait remarquer pour son franc-parler et ses rébellions intellectuellement contre les dogmes de la religion et l'irrationalité au temps des Lumières, suit les traces des Stoïciens en reprenant la notion d'éthique. Il vient même à remettre en cause la raison, ses fonctions et ses capacités quant à la motivation qu'elle peut apporter à nos comportements moraux. Outre sa remise en question de la raison, il présente des tentatives d'une théorie de l'esprit. Celle-ci décrit des boucles cognitives entre les

« impressions » causées par les émotions et les « idées » qui en sont à la fois source et conséquence d'un enchevêtrement de relations causales mettant en jeu d'autres « impressions ».

### **Une description favorable et précise des émotions sans théorisation**

Suite aux remises en cause des capacités de la raison, les philosophes commencent à placer les émotions à un niveau central de leur réflexion. Ainsi Kant (1724-1804) et Nietzsche (1844-1900), sans jamais vraiment théoriser le concept, décrivent avec une grande précision ce que peuvent être les émotions vis-à-vis de la raison. Le premier montre par ses critiques de la raison, l'importance des émotions dans l'échange et la manière d'appréhender la beauté des choses, alors qu'il est très inspiré par l'art et l'esthétique. Son disciple Hegel pousse même cette réflexion jusqu'à créer une « logique des passions ». Nietzsche quant à lui, connu pour son scepticisme, pense que les émotions permettent de véritablement regarder le monde alors que la raison n'en permet qu'une vision suspicieuse. Son apologie pour l'émotion étant tellement « passionnée », elle provoqua la peur de perdre la raison des autres philosophes. Dès lors, cela value un nouveau refoulement de l'émotion arrière-plan.

### **La philosophie moderne et les émotions**

Suite à ces réflexions, les pensées ont pris des chemins différents de part et d'autre de l'Atlantique et de l'Angleterre. Au 20ème siècle, du côté de l'Amérique du nord, il y eu un grand engouement pour la logique et la science qui ont porté le problème des émotions jusqu'aux mains des physiologistes pour finalement prendre son essor dans la psychologie. En Europe, l'émotion a su récolter plus d'attention par la lecture de certains auteurs tels que F. Brentano, E. Husserl, M. Scheler, M. Heidegger, P. Ricouer ou encore J-P. Sartre qui donne une dimension existentielle à l'émotion. Mais là encore l'intérêt pour les émotions s'est essoufflé et les réflexions se sont fermées au seul domaine de psychologie. Ce repliement sur la psychologie est l'une des plus grandes critiques que l'on fait actuellement, de même que les aspects trop interprétatifs de ces objets dont la description manque bien souvent d'objectivation. Ainsi les travaux actuels essaient de s'appuyer aussi bien sur la physiologie, la neurologie ou l'anthropologie pour assurer une théorisation plus globale et plus ouverte des émotions.

## ***Les principales théories des émotions dans le domaine de la psychologie***

Les émotions étant très étudiées, de nombreuses théories et réflexions existent et celles-ci peuvent être elles-mêmes un objet d'étude en soi. Nous ferons donc une brève introduction qui nous permettra de mieux appréhender les grands courants existants dans le domaine en montrant comment la psychologie organise et aborde le problème des émotions. Pour cela nous nous appuyerons sur une synthèse qui en est fait dans *Connaître les émotions humaines* (De Bonis, 1996).

### **Avant de rentrer dans les grands courants des émotions**

De Bonis évoque tout d'abord le problème de la définition des émotions qui est récurrente et permanente au fil des années. Chacun donne sa propre définition avec ses propres critères sans pouvoir véritablement faire preuve de consensus générale. On y évoque notamment une expertise de Plutchik sur les définitions de l'émotion données entre 1884 et 1977. Il aurait alors trouvé 28 définitions et cela uniquement en se basant sur les domaines de la psychologie et de la psychiatrie. De même, un tour d'horizon similaire aurait été réalisé entre 1971 et 1981 pour aboutir au dénombrement de 92 définitions (Kleinginna & Kleinginna, 1991). Pour ces auteurs, il semblerait que malgré la multiplicité de ces définitions, le terme d'émotions n'est que le reflet des grands courants de la psychologie, à savoir le fonctionnalisme, le béhaviorisme, le mentalisme, le cognitivisme et les théories basées sur la physiologie. La différence se ferait sur le fait que certaines mettent plus ou moins en avant les dimensions suivantes : le continuum hédonique (lié à la motivation), les effets adaptatifs, la délimitation des situations stimulantes et les réponses associées, le rôle des processus d'appréciation et l'étiquetage, la description des réactions physiologiques parallèles et enfin la remise en question du concept même d'émotion.

D'après De Bonis, W. James semblait quant à lui critiquer ses contemporains disant que ceux-ci se contentaient de faire de simples descriptions sans se préoccuper de ce que l'on pouvait déduire de ces entités, une idée qui sera partagée avec d'autres (Mandler, 1984 ; Frijda, 1986). De fait, il existe actuellement une grande variété de taxonomies que l'on voudrait utiliser comme entrée pour accéder à l'architecture de la représentation émotionnelle. Mais je trouve cela assez réducteur de se limiter à faire une simple liste d'émotions pour en comprendre leur fonctionnement. Il existe ainsi de nombreux lexiques et des dictionnaires dont on peut citer des auteurs comme Davitz. Averill a quant à lui

établi des atlas rassemblant diverses émotions (Averill, 1975) et il existe par ailleurs de la même manière, des cartographies cognitives des émotions.

La majorité des théories visent finalement dans un premier temps, à avoir une approche classificatoire des émotions. Là encore chacun propose sa propre répartition et pour cela on puise des critères dans les théories de l'évolution, de la cognition ou du langage. Ainsi, l'une des premières réflexions sur le sujet est partie d'une métaphore des émotions avec les couleurs. Une classification possible de ces couleurs se fait notamment par la différence que nous faisons entre couleurs primaires et secondaires. D'autres répartitions prennent également en compte les nuances de saturation et l'intensité. Ces observations sont également appliquées pour la description des émotions.

Au niveau expérimental, les théories s'appuient essentiellement sur des protocoles de description, classification, dénomination ou discrimination des émotions à partir de tests de perception. L'une des méthodes qui est principalement utilisée pour faire cela est de partir de termes que des experts proposent. Une observation fréquente est de voir que les éléments négatifs sont décrits avec de plus grandes précisions que les éléments positifs. Cela a notamment permis de mettre en évidence la notion de valence, qui est elle-même un objet de formalisation dans le domaine du TAL comme le montre par exemple Polanyi et Zaenen sur les « valence shifters » (Shanahan et al, 2006, pp.1-9).

Retenons que d'une manière générale, les théories des émotions diffèrent en ce qu'elles s'intéressent à comprendre le comment (manière dont elles se manifestent, soit la description), le combien (mesure de différents paramètres et comptage) ou le pourquoi (essayer de comprendre quelle est l'origine des émotions). Enfin, certaines théories essaient d'expliquer localement ce concept (soit faire la description d'une seule émotion) ou d'adopter un point de vue plus large (ex: cas des études taxonomiques des émotions).

### **Le courant évolutionniste**

Le premier à avoir étudié les émotions sur des bases scientifiques est certainement l'auteur de *On the origin of species*, qui décrit la théorie de l'évolution naturelle. En effet, pour Darwin, les émotions expliquent directement sa théorie évolutionniste car il s'agit de phénomènes évolués correspondant à des fonctions de survie pour permettre aux hommes de s'adapter à leur environnement (Darwin, 1872). Il met ainsi en avant l'hypothèse de l'universalité des émotions. Pour illustrer cela, il effectua de nombreuses études au sein de peuples de différentes cultures de tous les continents en s'intéressant tout particulièrement

aux fonctions communicatives et adaptatives des individus. Le principe de ses études consistait à montrer des mimiques de personnes, récoltées durant ces voyages. Ces expressions produites dans certaines situations soulignent 16 interrogations posées à des compatriotes de Darwin, quand à savoir quel état émotionnel elles représentaient. Les réponses ont alors permis d'émettre l'hypothèse de l'universalité des émotions.

P. Ekman qui est le grand héritier de ce courant émet un certain nombre de critiques pour en retirer des améliorations qu'il utilise pour ses propres études (Ekman, Sorenson et Friesen 1969). Il s'appuie pour cela sur l'interculturalité et utilise, contrairement à Darwin, des questions discriminantes (plutôt que des réponses oui/non). Ces interrogations sont adressées à des natifs pour illustrer l'universalité des émotions, notamment en interrogeant des personnes qui ne sont pas empreintes de la culture véhiculée par les médias pour se soustraire au biais de l'apprentissage de ces émotions. Il met alors en évidence 6 émotions principales (appelés communément les Big Six), visibles sur des patterns spécifiques d'expressions faciales qui sont universellement reconnues : la joie, la surprise, la tristesse, la peur, le dégoût, la colère. A cela se rajoute de multiples émotions secondaires qui dérivent de ces émotions primaires. De la même manière Plutchik décrit 8 émotions de base (Plutchik, 1977), Izard et Tomkins en comptent neuf par exemple, en s'appuyant sur des études dont les méthodes sont similaires à celles d'Ekman. (Izard, 1980 ; Tomkins, 1980).

Depuis le début du 20<sup>ème</sup> siècle, de telles expériences basées sur l'observation des émotions en contexte bi ou pluriculturel se comptent par centaine. De la conclusion de ces multiples observations d'expressions faciales, nous aboutissons finalement au modèle du « Facial Action Coding System » ou FACS (Ekman & Friesen, 1975), s'appuyant sur découpage précis des muscles faciaux du visage pour permettre la reconnaissance des émotions.

Ekman observe par le biais du FACS que bien qu'il y ait une large reconnaissance et une universalité des émotions, il existe des différences culturelles dans la perception de ces objets. En effet, il remarque que ce qui changent dans la perception, ce sont surtout les modes d'expressions des émotions ou encore les conditions de déclenchement et non les émotions en eux-mêmes. A la suite d'Ekman, Matsumoto note cette différence culturelle également au niveau de l'expression de certaines émotions en lien avec le contexte social ce qui s'illustre par exemple par des différences de réactions face à des stimuli émotionnels identiques chez des Américains et des Japonais. Enfin au niveau de la perception, il

existerait un facteur d'intensité qui permettrait d'expliquer les différences culturelles (Ekman et Matsumoto, 1989).

### **Le courant physiologique**

A la différence des théories précédentes qui s'intéressent à la nature même des émotions, les théories physiologiques observent le mode d'expressions de ces émotions. Au 19<sup>ème</sup> siècle, James et Lange se sont posés les mêmes questions presque simultanément. Nous appelons ainsi ces considérations la « théorie James-Lange ». Celle-ci prône que toute émotion passe tout d'abord par une expression physiologique et corporelle. Pour James, l'expression cognitive des émotions n'apparaîtraient qu'après son expression physiologique. Il considère donc que ce ne sont pas les centres neuraux mais le système autonome qui intervient dans la boucle émotionnelle. Ainsi c'est le fait de percevoir des stimuli par les capteurs proprioceptifs qui induit les émotions qui auraient ainsi un rôle adaptatif vis-à-vis de son environnement (James, 1984). Lange est lui plus nuancé car il considère que le cognitif a sa part dans la boucle émotionnelle, en considérant qu'il existe des patrons spécifiques permettant des réponses adaptées face au type de stimuli impliqués (Lange, 1985).

Cette théorie est largement critiquée par A. Damasio qui pense que ni James, ni Lange n'aborde les aspects évaluatifs des émotions (Damasio, 1995). D'ailleurs, ces théories physiologiques servent de point d'ancrage aux technologies telles que les détecteurs de mensonges dont la pratique pose de gros problèmes éthiques à l'heure actuelle.

Enfin, même s'il s'insère dans le courant physiologique, W. Cannon critique la théorie de James-Lange en s'appuyant sur sa théorie centrale thalamique des émotions. En effet, Cannon considère que face à un stimulus de l'environnement, nous avons une conscientisation émotionnelle et une boucle cognitive qui s'active au niveau central. C'est cette activation qui génère les réactions physiologiques et non l'inverse. Pour argumenter ces propos, ce dernier a effectué des expériences sur des animaux dont les systèmes végétatifs ont été coupés, qui rendent compte de réponses émotionnelles même en l'absence de stimuli périphériques (Cannon, 1927).

### **Les théories cognitivistes**

Ce courant de pensée apparaît au début des années 60 avec M. Arnold et R. Lazarus. Le cognitivisme est l'un des courants dominants dans les théories des émotions

avec un nombre important d'auteurs qui se placent dans la lignée de ces réflexions. Ces auteurs introduisent une notion importante : « *appraisal* ». Cet *appraisal* illustrerait les représentations cognitives ainsi que les interprétations que nous faisons et gardons d'un événement, en remarquant notamment son caractère plaisant ou déplaisant. Ces processus sont « direct, immédiat, non réflexif, non-intellectuel et automatique » (Arnold, 1960). Nous jouons alors sur la distance vis-à-vis de cet événement. Cette notion sous-tend donc l'idée qu'il y aurait une phase de préparation à une action évaluant les différents paramètres qui interviennent. Les émotions sont alors différenciées en fonction de nos souvenirs de situations semblables, notre personnalité, ou encore notre environnement social. C'est cette idée que nous retrouvons chez Scherer qui qualifie ce processus « d'évaluation » (Scherer, 2001). Ce qui est d'ailleurs intéressant avec lui, c'est le fait qu'il s'intéresse également aux paramètres acoustiques des comportements affectifs, ce qui est encore très rare dans le domaine puisque les principaux travaux s'appuient sur la réalisation visuelle de ces événements. Par ailleurs, N. Frijda qualifie le phénomène d'*appraisal* par la « préparation à l'action » (Frijda, 1986 et 1987). C'est également de cette évaluation que Damasio, cité plus haut parlait quand il critiquait les théories physiologistes et lui parle de « prise de décision » (Damasio, 1995). En effet, il illustre ses propos avec le cas de Phineas Gage (1848) atteint d'un traumatisme crânien sans altération des capacités intellectuelles qui montre que le comportement d'un individu peut changer du fait des modifications perceptives qu'il a de son environnement. Ces modifications induisent donc des prises de décisions différentes d'avant son accident ce qui semblait s'apparenter pour ses proches à une modification de sa personnalité. Cela suppose que la signification émotionnelle d'une situation est fortement dépendante de la subjectivité de chacun et que nos croyances ainsi que notre arrière-plan culturel influencent notre évaluation, d'où la diversité de nos réactions émotionnelles face à une certaine situation.

### **Le constructivisme social des émotions**

Ce courant soutient que toutes nos attitudes affectives ne sont que le résultat d'une construction socioculturelle de l'homme. C'est notamment ce que présente J. Averill dans ses études. Les émotions seraient des réponses à une norme sociale pour s'adapter aux conditions qui s'imposent à une situation socioculturelle donnée. Cette théorie est notamment motivée par le fait que les réponses émotionnelles ne sont pas toujours les mêmes d'une culture à une autre et s'oppose donc aux théories physiologiques qui suggèrent que ces réponses émotionnelles sont issues d'une perception entièrement

biologique (Averill, 1980). Ainsi au niveau méthodologique, ce courant préconise de se focaliser sur les différences du comportement émotionnel plutôt que sur les similitudes, c'est-à-dire des idées favorables aux classifications émotionnels comme le fait Ekman. Pour comprendre comment les sociétés s'approprient ces émotions, de nombreuses études se sont intéressées au langage car pour certains c'est dans le lexique d'une langue que nous retrouvons la manière dont nous percevons notre environnement, ce qui est une conception très linguistique du monde (Wierzbicka, 1999). Ces études des émotions par les langues est la méthode empirique essentiellement utilisée dans ce courant socioconstructiviste.

### **Interdisciplinarité, sciences affectives et nouvelles réflexions**

Comme nous l'avons vu à la fin de la partie sur les conceptions philosophiques des émotions, la psychologie ouvre ses portes à de nouvelles disciplines pour traiter la problématique des émotions. Ce sont ces nouvelles disciplines qu'on appelle les « sciences affectives » qui s'intéressent à cette nouvelle perception pluridisciplinaire des comportements humains. Certains auteurs s'intéressent par exemple aux relations entre neurosciences et émotions pour mettre en évidence les zones corticales qui interviennent lors d'une apparition d'un stimulus émotionnel (Davidson et al, 2003). Cela est désormais possible grâce aux avancées actuelles de l'imagerie médicale. En tant que nouvelles sciences, nous avons également des considérations génétiques qui rentrent désormais en jeu (Goldsmith, 2003). En effet, cela permet de montrer les enjeux développementaux des processus cognitifs chez l'humain mais fait également le pont avec le milieu médical. Ce domaine est à la fois le lieu d'observation de phénomènes pathologiques qui permettent d'étudier par contraste les comportements affectifs comme les troubles de l'autisme (Insel, 2003) ou des problèmes de dépressions (Garlow et al, 2003). Ces observations peuvent alors permettre de mieux comprendre les phénomènes de la psychologie affective. Par ailleurs, le domaine médical peut aussi être concerné car les émotions peuvent avoir une conséquence dans la progression du cancer (Giese-Davis et Spiegel, 2003). Ainsi la compréhension des processus cognitifs pourrait permettre d'apporter des solutions à certains problèmes pathologiques. Enfin la musique ou l'art sont également de nouveaux lieux d'investigation en lien avec les émotions (Gabrielsson, 2003).

## ***Conclusion du Chapitre 2***

Les considérations philosophiques avaient placé les émotions au cœur du débat qui les opposaient à la raison. D'abord soumises puis s'affirmant de plus en plus, elles semblaient parfois avoir des fonctions qui surpassaient la raison, à tel point qu'elles ont été mises de côté par peur des conséquences. Ce contexte a favorisé la prise en main de ces émotions par la psychologie qui s'est renfermée sur ses observations et a de nouveau marginalisé ces concepts. La psychologie a ainsi investi et établi de grands courants dont la pertinence est toujours d'actualité, en revanche elle se trouve confrontée à nécessité d'ouverture vers d'autres réflexions et disciplines pour continuer à évoluer. Cette pluridisciplinarité a désormais lancé un nouvel engouement pour les émotions car de nouvelles possibilités technologiques sont apparues pour investir les méthodes d'observation d'objets émotionnels et affectifs.

Nous venons donc de voir dans le chapitre 2 ce que sont les émotions, or ces objets sont aussi ce que nous pouvons considérer comme le « signifié » des micro-expressions. Mais comme tout objet de communication, les micro-expressions sont également caractérisées par une forme ou un « signifiant » si nous reprenons les termes de F. de Saussure, le père de la linguistique moderne. Les disciplines des sciences du langage proposent ainsi de nombreuses solutions pour aborder le « signifiant », son « référent » et les manières dont ces éléments se manifestent dans la communication (Saussure, 1995).

Ces observations se font à des niveaux de granularité très variés et alors que les micro-expressions ne sont pas toujours véritablement considérées comme des faits de langage car non lexicalisés, ces disciplines donnent des pistes d'investigation pour ces micro-expressions. Il y a donc là tout un état des lieux à mener sur les moyens que les disciplines de sciences du langage (et notamment la pragmatique) utilisent pour caractériser ces micro-expressions et leurs usages qui sont un aspect important à considérer. Ce domaine étant lui-même très vaste, nous ne l'aborderons pas précisément dans le cadre de ce seul mémoire, toutefois nous essaierons d'évoquer certaines de ces notions dans le chapitre 3 à partir des premières observations effectuées par des travaux antérieurs. En effet, le principal souci de ce mémoire étant de comprendre comment l'information portée par les micro-expressions persiste au filtre culturel, c'est moins les formes du mode de communication mais la valeur informative des micro-expressions qu'il importait de bien définir.

## Chapitre 3 - Les Travaux précédents à l'origine du mémoire

Avant toute réflexion sur les micromots en eux-mêmes, il est important de se demander où est-ce que ces objets apparaissent et comment les recueillir afin de pouvoir les observer. Les micro-expressions de la parole sont des événements qui surviennent dans toute interaction conversationnelle écologique humaine, qui quand elles apparaissent de manière involontaire, sont décrites comme des « émotions ». C'est donc une parole « naturelle » à une situation sociale qu'il faut pouvoir recueillir afin de trouver des micro-expressions. Cette question méthodologique est d'ailleurs bien posée par Schröder qui traitait le phénomène des rires et de l'amusement dans la parole, qui font partie intégrante de nos micro-expressions (Schröder, 1998).

Par ailleurs, O. Rosec<sup>5</sup> affirme qu'actuellement, dans les synthèses de parole industrielles les plus avancées, des comédiens jouent le rôle social d'une personne pour permettre leur recueil de corpus. En effet, ces professionnels veulent volontairement insérer certaines caractéristiques notamment prosodiques et stylistiques dans leur synthèse pour se rapprocher au plus d'une parole « naturelle ». Or cela ne peut se faire qu'en contrôlant les productions des acteurs en les remplaçant dans un contexte interactionnel maîtrisé.

Nous voyons donc que le recueil de tels objets fait parti d'un enjeu applicatif primordial, mais nous devinons qu'il est également très complexe. Le chapitre 3 illustrera donc les notions importantes qui sont à connaître et à mettre en œuvre pour travailler sur l'élaboration d'un corpus de micro-expressions ainsi que les solutions qui ont été apportées au GIPSA, puisque les études de mon mémoire reprendront un corpus constitué au sein de ce laboratoire, dans le but de trouver des méthodologies pouvant répondre aux problèmes soulignés ci-dessus.

### Les différentes méthodes de recueil des micro-expressions

Dans la constitution de tout corpus, mais surtout d'objets émotionnels ou attitudeaux non lexicalisés, trois paramètres seraient mise en jeu (Aubergé et al, 1995) : l'opposition du recueil *in vivo* / *in vitro* ; le degré de contrôle que l'expérimentateur impose

---

<sup>5</sup> Chercheur issu de l'entreprise Orange, désormais dans l'unité de recherche de l'entreprise Voxygen. Il a présenté des méthodes de synthèses vocales utilisées dans les entreprises actuelles lors de sa conférence «*La synthèse de la parole : principes et enjeux*» ayant eu lieu à la Maison des Langues du campus universitaire de Grenoble, le 6 Janvier 2012.

pour s'assurer de la présence de paramètres spécifiques dans le corpus (à des granularités linguistiques et interactionnelles différentes), et l'opposition entre attitude « actée » et « authentique ».

Cette dernière opposition est la différence entre un recueil « naturel » soit (quasi) écologique avec peu ou sans contrôle de la part des expérimentateurs et un recueil « contrôlé » qui passe essentiellement par le jeu d'un acteur mimant un certain état émotionnel tel qu'il est imposé par les conditions expérimentales (Schröder, 1998 p33-36).

F. Loyau oppose quant à elle les méthodologies *in vivo* et *in vitro*. Dans le premier cas, le corpus est enregistré à partir de situations quasi naturelles, hors cadre expérimental apparent. Ces méthodes sont opposées aux manipulations *in vitro* où un contexte expérimental contraignant s'impose aux sujets enregistrés. Dans cette opposition, elle observe notamment dans un corpus *in vivo* de pilotes de chasse que la naturalité est importante mais que la qualité du corpus est fortement liée à la motivation des locuteurs pour les tâches dans lesquelles ils sont impliqués. Ainsi les méthodes *in vitro* permettraient de mieux contrôler l'environnement pour véritablement observer les phénomènes attendus. Donc il faut un véritable compromis entre ces méthodes de recueil, l'authenticité et le degré de contrôle (Loyau, 2007 p57-74).

De nombreuses études ont ainsi mis en place des méthodologies de recueil variées pour répondre aux divers besoins d'observation, et bien que nous ne puissions être exhaustifs dans leur présentation, en voici un bref aperçu.

Très tôt, nous avons remarqué l'importance d'intégrer des dimensions affectives dans les corpus, en réalisant des enregistrements d'acteurs lisant des textes chargés de diverses situations émotionnelles (C.Williams et N.Stevens, 1972). Dès lors, de nombreuses expériences ont été menées pour aboutir à un inventaire de 250 études recueillies sur une dizaine d'années, portant sur les méthodes d'induction d'émotions, classifiées en cinq groupes de procédure distincts (Gerrards-Hesse et al, 1994). Nous voyons par ailleurs des corpus *in vivo* qui présentent comme dans le cadre du projet PHYSTA (Principled Hybrid System Theory and Applications) des ressources audiovisuelles de deux types : des discussions de sujets parlants de thématiques sensibles suscitant diverses émotions et des extraits de programmes télévisés faisant un focus sur les réactions spontanées du public (Douglas-Cowie et al, 2000). Mais il semblerait qu'il soit plus intéressant de stimuler les émotions pour produire une expression humaine naturelle

plutôt que de simuler ces émotions (Campbell, 2000). De fait, bien qu'elles soient moins nombreuses, certaines méthodes essaient d'induire des contrôles relativement élaborés dans un cadre *in vivo*. De fait certains corpus sont élaborés à partir de tâches informatiques ou des jeux vidéos (Kaiser et al, 1994 ; Johnstone et al, 1999), pour inciter des émotions authentiques et analyser ces dernières en termes d'expressions faciale. Cela est fait automatiquement chez Kaiser, basé sur une analyse physiologique et acoustique chez Johnstone ou encore Scherer.

### Recueil de corpus E-Wiz avec le protocole Sound Teacher

Le compromis trouvé au GIPSA est l'utilisation d'une méthode *in vitro* rendue écologique dans le cadre d'une IHM contrôlée par un protocole de type « Magicien d'Oz » (Aubergé et al, 2005). Ce protocole de recueil d'expressions authentiques est mis en place à l'aide de la plateforme E-Wiz (Expressive Wizard of Oz). Le principe de cette manipulation est de « piéger » un sujet en lui faisant faire une tâche qui n'intéresse pas véritablement l'expérimentateur, mais servant à provoquer des réactions authentiques qui auraient pour source : les difficultés d'une tâche, la satisfaction obtenue lors d'une réponse correcte, le doute que peut susciter certaines questions, etc.

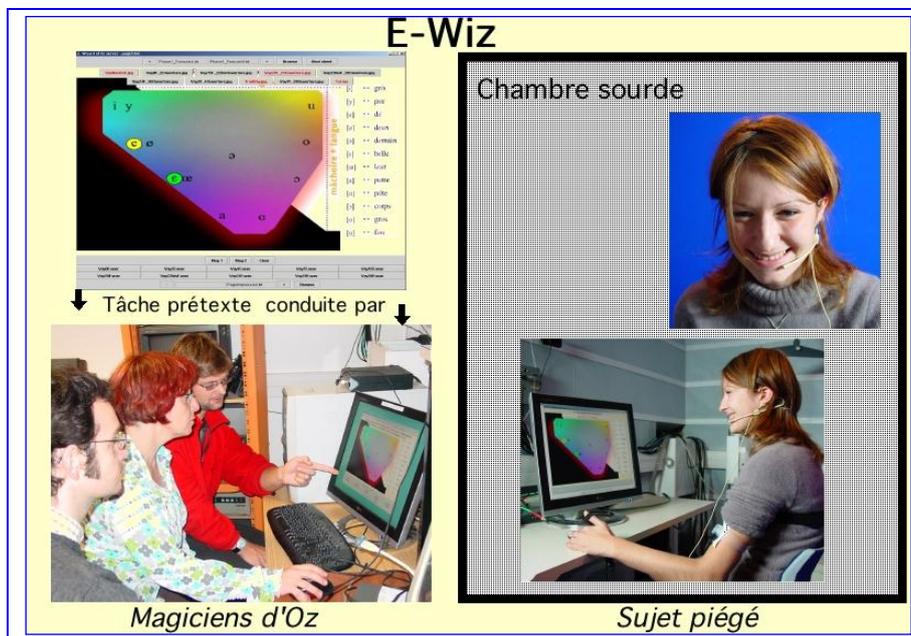


Figure 1 : Schématisation du protocole E-Wiz

Cette plateforme E-Wiz a été développée en langage Java, à l'ICP (Institut de la Communication Parlée). Il est composé d'une part d'un éditeur qui permet de concevoir

graphiquement des scénarios avec une obtention automatique des scripts correspondants. D'autre part des applications de type client/serveur permettent de générer les scripts précédemment créés pour effectuer les enregistrements de corpus. Lors de cette étape d'enregistrement, les expérimentateurs ont la possibilité de gérer le comportement de l'interface que le locuteur utilise pour induire certaines attitudes chez le sujet et donc de récupérer les productions langagières et les micro-expressions qui nous intéressent. En effet, l'IHM mime une communication par commande vocale qui justifie totalement la présence de micro pour l'obtention d'enregistrements de bonne qualité. Les éléments lexicalisés correspondent aux termes utilisés dans le cadre de la commande vocale, qui sont notamment « page suivante » pour continuer le processus et 5 termes monosyllabiques servant à donner une réponse (« rouge », « sable », « vert », « jaune », « brique »). A cela peut s'ajouter des productions lexicales libres du sujet lorsqu'il fait des commentaires sur le logiciel « qu'il évalue ». Nous considérerons que tous les autres éléments (majoritaires) rentrent dans ce que nous appelons les micro-expressions de la parole.

Actuellement, deux scénarios sont disponibles : Top Logic dont nous ne détaillerons pas ici la description car elle nous intéresse moins dans le cadre de ce mémoire, et Sound Teacher sur lequel nous nous attarderons plus longuement. Ce dernier scénario est présenté comme un nouveau logiciel révolutionnaire qui permet d'améliorer l'apprentissage de la phonétique des langues étrangères. Les sujets interrogés sont amenés à s'entraîner dans le but de reconnaître des valeurs parmi quatre paramètres du conduit vocal (l'ouverture, la position avant/arrière, la centralisation et l'arrondissement des lèvres) à l'issue d'écoutes de voyelles.

Globalement, la tâche des locuteurs se découpe en 4 étapes de difficulté croissante. La première consiste à la reconnaissance de voyelles françaises qui étant simple, permet d'induire une rétroaction très positive induisant les réactions conséquentes chez le sujet. Puis dans la seconde étape cette tâche se complique un peu en passant par l'apprentissage de voyelles proches du système vocalique français. Là encore la rétroaction est positive et elle permet de faire une transition avec une troisième étape se complexifiant davantage. La rétroaction devient alors négative et induit un nouveau type de productions expressives. Dès lors, cela permet l'introduction d'une quatrième étape qui permet de revenir sur les voyelles du français en prétextant que le logiciel a pu altérer les capacités de reconnaissance du sujet et nécessite donc une vérification. En diminuant les contrastes perceptifs des voyelles présentées, cette dernière étape permet de voir les réactions des

sujets face à l'annonce du faible score qu'ils ont obtenu. L'enregistrement audiovisuel de toute cette pseudo-séance d'évaluation du système d'apprentissage permet donc le recueil d'un corpus expressif authentique.

### ***Les motivations d'études sur les micro-expressions et terminologie***

A l'aide du scénario Sound Teacher d'E-Wiz, F. Loyau décrit la constitution d'un corpus audiovisuel à partir de 17 sujets français (Loyau, 2007). Cela lui a notamment permis d'effectuer les premières observations sur les micro-expressions et de mettre au point quelques précisions terminologiques non négligeables que nous présenterons dans cette section.

#### **Le Feeling of Thinking**

F. Loyau associe les événements recueillis dans ce corpus à la notion de « Feeling of Thinking » qu'elle définit comme des icônes minimales de « tous les gestes, les mouvements du buste, de la face, les bruits de bouche produits par le sujet quand il ne parle pas lors d'une situation de communication et qui correspondent aux manifestations visuelles et sonores de ses états mentaux et affectifs», définition qui collerait parfaitement à nos micro-expressions (Loyau, 2007 p117-165). Nous retrouvons également la dénomination de « prosodie audiovisuelle » pour désigner ce concept. Cette notion est directement dérivée et élargie de celle du « Feeling of Knowing » (Reder & Ritter, 1992). Ce « Feeling of Knowing » serait le sentiment d'avoir connaissance d'un événement que nous aurions gardé en mémoire sans véritablement pouvoir y accéder au moment où nous produisons l'objet qui l'illustre. C'est notamment ce concept que l'on retrouve lorsque nous faisons face au phénomène du mot sur le bout de la langue.

Retenons simplement ici que le « Feeling of Thinking » sous-tend l'idée que les micro-expressions s'apparentent à un mécanisme cognitif mettant en jeu la mémoire et la perception que chacun a de ses propres connaissances.

#### **Quelques dichotomies à noter sur les micro-expressions**

Une précision terminologique importante est celle du paradigme « signaux » / « indices » (Loyau, 2007). Les premiers seraient des objets récurrents chez tous les individus observés, les seconds des spécificités de chaque personne. Cette différence n'est pas anodine car la fonction communicative est au centre de nos micro-expressions. En effet, les indices que nous appelons communément des tics ne semblent pas être

intrinsèquement significatifs contrairement aux signaux. En revanche la manière dont apparaissent ces indices (notamment la fréquence et la durée des intervalles entre leur apparition) peut illustrer l'état émotionnel dans lequel se trouve une personne, ce qui sera important de considérer dans les perspectives d'observations de micro gestes (Carlier et Graff, 2006 cités par Loyau, 2007).

De même, la distinction entre « verbal » / « non verbal » est importante à souligner dans le cadre de la description des micro-expressions. Au lieu de considérer que tout ce qui est lexicalisé fait partie du verbal et tout le reste du non verbal, nous considérons que les micromots et les microgestes, qui sont volontairement produits dans un but communicatif, font partie du verbal alors que ceux qui sont involontaires font partie du non verbal. Cela n'enlève en rien l'importance du non verbal puisque ce dernier est la manifestation de ce que nous appelons les émotions, qui elles-mêmes donnent un contexte à l'interaction en cours. C'est par exemple le cas des phénomènes physiologiques qui surviennent involontairement quand nous changeons d'état émotionnel (ex: augmentation de la sudation face à une situation de peur). Ce sur quoi nous porterons de l'intérêt pour ce mémoire est entre autre le passage d'un contrôle involontaire vers des structures contrôlées qui illustreraient peut-être la première phase de l'acquisition du langage.

### ***Premières observations et annotations de corpus***

Dès lors que nous avons défini comment recueillir les micro-expressions et que nous avons mis l'accent sur quelques aspects terminologiques qui me paraissent indispensables de connaître, la question suivante est de savoir ce qu'il faut exactement regarder dans la masse d'informations sélectionnées ; ce qui est pertinent ou non de remarquer et ce qui est significatif pour la fonction communicative. Ces réflexions ont alors fait l'objet d'un premier travail d'annotations basée sur l'éthologie (Vanpé, 2011 pp. 112-133).

### **Des méthodes d'observations basées sur l'éthologie**

L'éthologie prend ses origines auprès de trois hommes K. von Frisch, K. Lorenz et N. Tinbergen qui obtiennent le prix Nobel de médecine et de physiologie en 1973, grâce à leurs travaux sur le comportement animal. Le premier travailla sur la « danse des abeilles », le second est lui connu pour sa théorie des instincts, l'« imprinting » ou « l'empreinte ». Enfin le troisième étudia divers animaux tels que les laridés (oiseaux marins) ou encore des poissons comme l'épiloche. Par l'observation de ces différentes

espèces, naît l'éthologie définie comme « the biological study of behavior ». Cette définition amène Lorenz et Tinbergen à s'opposer aux théories behavioristes qui prônent que l'instinct est un phénomène acquis au contact de l'environnement, en affirmant que ces comportements font partis de l'inné (R.W. Burkhardt, 2005).

Depuis, de nombreux travaux s'appuient sur cette méthode et au GIPSA, A. Vanpé a adopté une méthode éthologique pour éviter les écarts vis-à-vis de la science puisque l'objet de l'analyse est très abstrait et facilement subjectivable à mauvais escient. Dans les études comportementales et affectives, cette méthode est très souvent utilisée. Le groupe de Recherche Geste et Voix d'Aix-en-Provence du laboratoire « Parole et Langage » cite ainsi de nombreux travaux tels que ceux de Eibl-Eibesfeld (1979), Cosnier (1987), Cyrulnik (1989 puis 1991), Feyereisen & De Lannoy (1997) ou encore Montagner (1978). (Guaïtella, 1998). Cette pratique consiste à adopter les méthodes d'observation similaires à celles utilisées dans les études de comportement animal en remplaçant l'humain au centre des observations en tant qu'objet biologique, sans pré-requis subjectifs.

### **Les différentes annotations du corpus**

A partir du corpus réalisé sous protocole E-Wiz, les sujets qui ont participé à l'expérience se sont auto-annotés eux-mêmes sur les icônes gestuelles, faciales et vocales. L'expérience place les sujets dans une situation de subjectivité réflexive, un parti pris que nous utiliserons également pour les tests de perception qui seront réalisés dans ce mémoire.

Dans l'impossibilité de recueillir des données purement objectives sur de tels éléments, il est plus judicieux de faire preuve de subjectivité naïve qui soit en adéquation avec le ressenti des sujets puisque c'est exactement ce que nous rechercherons à connaître et récupérer afin de déduire les modalités de réponses ou les attentes des sujets. Nous verrons d'ailleurs que cette rétroaction adaptée est en partie un autre enjeu applicatif pour les nouvelles technologies de la parole.

Dans un second temps, des « experts », qui sont cette fois les sujets du test de perception, annotent ces données sans avoir connaissance des étiquettes issues des auto-annotations effectuées précédemment, pour ne pas amplifier la subjectivité de l'observation. Ce travail d'étiquetage a permis d'avoir une référence des éléments observables dans les moments hors tour de parole pour avoir un aperçu global des icônes expressives produites lors d'une situation d'interaction personne-machine simulant une activité didactique.

### **Les perspectives d'études mis en évidence par A. Vanpé**

Les éléments observés par A. Vanpé sont loin d'être uniquement des réactions involontaires, puisqu'il existerait de nombreux signaux qui sont contrôlés. Une étude psychoacoustique fine de ces objets est l'une des pistes d'investigation proposée. D'autre part, elle observe une grande variation interpersonnelle dans les choix d'utilisation et la nature de ces icônes, en particulier vocales. Ces événements auraient également une dynamique qui semble jouer un rôle dans l'organisation temporelle de ces objets dans l'interaction. Elle émet également l'hypothèse de l'existence d'un continuum de ces microévénements sur une échelle passant d'objets émotionnels non contrôlés vers des objets de plus en plus contrôlés, introduisant une progressive lexicalisation. Enfin, elle conclut sur l'importance de la multimodalité dans l'étude des comportements expressifs.

Afin de répondre à ces problèmes, deux études ont été proposées et ont mis en évidence une dimension culturelle des micro-expressions pour l'une et une dimension socio-affective de ces éléments pour l'autre.

#### **Une étude perceptive culturelle des micro-expressions**

R. Signorello s'intéresse à « l'identification perceptive de juges français et italiens sur des microévénements acoustiques et visuels non langagiers », soit nos micro-expressions. L'expérience (Signorello et al, 2011) s'appuie sur des stimuli choisis à partir du corpus de Loyau et Vanpé, produits par 6 locuteurs français. Ces éléments sélectionnés constituent un corpus de 24 stimuli ordonnés selon un ordre croissant de contrôle prosodique, ainsi que leur distance avec les sons permettant de former les entrées lexicales du français, soient les sons phonologiques. L'étude teste la perception en visuel seul, audio seul et audiovisuel de 15 juges de nationalités françaises ou italiennes.

Le résultat de l'expérience montre une première tendance des juges à identifier la source des stimuli comme étant produite par des Français et cela quel que soit la nationalité des juges. Le second point est qu'en visuel, les locuteurs font un choix d'attribution d'une culture, dès le début de l'expérience, en s'appuyant sur l'aspect des locuteurs et restent sur leurs impressions durant toute la durée de l'expérience, que cette décision soit vraie ou fausse. Par opposition, en audio et en audiovisuel, la décision des juges varie en fonction du stimulus observé. L'observation des réponses sur l'ensemble montre néanmoins que pour les deux modalités de perception, il existe une frontière à partir de laquelle le contrôle introduit un « point de stabilité », c'est-à-dire que l'information culturelle portée par les

indices langagiers est suffisamment importante pour discriminer la langue du locuteur. Après cette frontière, les réponses et la confiance que ceux-ci apportent convergent vers la langue des locuteurs qui produisent les stimuli. Enfin, cette expérience montre que l'information visuelle et l'information auditive ne sont pas additives dans l'élaboration de l'information.

L'une des perspectives citées dans cette étude serait de regarder si nous observons le même type de phénomènes pour des locuteurs de langues plus éloignées que le français et l'italien pour définir si cette observation a un caractère universel ou non. De même, observerions-nous la même chose dans le cas inverse des stimuli de locuteurs italiens analysés par des locuteurs français ?

### **La dimension socio-affective dans un test de perception**

La seconde étude de perception concerne le travail mené par G. De Biasi. Ce dernier a repris comme dans l'étude précédente des stimuli du corpus d'A. Vanpé afin de réaliser des tests de perception sur la valeur émotionnelle que renvoient les micro-expressions. Il utilise ainsi les auto-annotations (33 sélections à partir de 6 locuteurs : T, S, M, J, N, R) recueillies par A. Vanpé comme choix d'étiquettes pour les sujets. Comme précédemment, cette expérience porte sur les trois modalités : auditive, visuelle et audiovisuelle sur 66 stimuli (2 par étiquette). Le nombre de participants était de 15 sujets (9 femmes et 6 hommes) (De Biasi, 2011).

L'un des premiers constats effectués autour de cette étude est la différence perceptive entre les objets acoustiques et visuels. Les micromots seraient porteurs d'« émotions » à savoir les ressentis innés et involontaires, alors que les microgestes à l'origine d'une perception visuelle seraient sujets à véhiculer des affects et attitudes sociaux qui sont le fait de contrôles volontaires. La résultante en serait que tout ce qui rentre dans le cadre du *Feeling of Thinking* est visible avant de pouvoir être entendu alors que les autres formes qui s'écartent des formes langagières auraient uniquement tendance à pouvoir être entendues.

Enfin, quel que soit le cas, sans la connaissance du contexte de productions des micro-expressions (ce qui a été le cas pour cette étude) les sujets semblent faire preuve de nombreuses confusions, ce qui nous permet de souligner l'importance du contexte d'énonciation dans la compréhension de notre propre perception concernant les micro-expressions.

De Biasi propose lui aussi un certain nombre de perspectives aux termes de son étude. Tout d'abord, il suppose qu'une analyse du degré de confiance des réponses de chaque juge permettrait de renforcer les constats qui ont été faits sur le caractère acquis/inné des objets émotionnels. Par ailleurs, il considère qu'une annotation par stimuli (c'est-à-dire en fonction des caractéristiques phonétiques plutôt que par étiquettes d'annotations) pourrait éventuellement permettre de faire des classements empiriques des différents types de micro-expressions produites. D'autre part, il suggère une création de méta-classes d'annotations pour chacune des modalités (auditive, visuelle et audiovisuelle) afin d'affiner les caractéristiques en lien avec la modalité. Enfin, il préconise d'ajouter une dimension interculturelle à cette étude en prenant des locuteurs de langues différentes (comme le japonais dont les caractéristiques langagières et culturelles sont très éloignées des langues européennes d'ors et déjà étudiées).

### ***Conclusion du Chapitre 3***

Le recueil de micro-expressions est un processus complexe nécessitant du temps est un protocole précis, mais les solutions sont nombreuses afin de répondre aux besoins spécifiques d'observation.

Par ailleurs, les quelques expériences du GIPSA présentées ci-dessus ont permis de mettre en évidence des objets particulièrement intéressants véhiculant une information communicationnelle basée sur l'état affectif ou émotionnel, à savoir les micro-expressions. Ces dernières semblent être le lieu d'observations du passage de faits non lexicaux vers des entités lexicalisées qui permettraient d'apporter des prémises d'idées sur l'une des questions existentielles de l'homme : la naissance et l'acquisition du langage chez l'humain.

Il reste dès lors à observer ces objets avec une plus grande finesse en y ajoutant une dimension interculturelle afin de pouvoir mettre en évidence des caractéristiques plus ou moins universelles et approfondir la fonctionnalité communicative de ces derniers.

La suite de ce mémoire, notamment les parties réalisation des manipulations expérimentales et analyses essaieront de prendre en considération les perspectives que les travaux antérieurs ont mis en évidence, en adoptant notamment une démarche empirique d'observations de données.

## **Partie 2**

### **Expériences perceptives**

## **Chapitre 4 – Test de perception de micro-expressions par des sujets japonais**

Ce mémoire s'intéresse à la perception des micro-expressions dans la modalité audio-visuelle et à la contribution perceptive spécifique de chacune des deux modalités dans le contraste franco-japonais. Mais pour commencer, nous nous intéressons au protocole expérimental mis en place pour mesurer la perception par les sujets japonais dans ce chapitre 4.

Ici il sera ainsi fait un rappel sur les perspectives attendues dans de précédentes études sur lesquelles nous nous sommes basées, afin de resituer le contexte de l'expérience. Puis nous décrirons le protocole et l'outil expérimental, tel qu'il a déjà été utilisé pour des sujets français, et tel que nous l'avons adapté spécifiquement aux sujets japonais. Nous ferons ensuite une description des conditions de recueil des données qui étaient relativement spécifiques dans le cadre de cette étude. Enfin nous finirons par décrire les sujets interrogés pour illustrer plus particulièrement leurs connaissances linguistiques (surtout françaises), paramètres qui pourraient introduire du bruit dans les réponses du test de perception.

### ***Une manipulation répondant aux perspectives d'études antérieures***

Le recueil de micro-expressions (cf. Chapitre 3 section 2) relève en soi d'un protocole complexe puisque nous choisissons de travailler sur des données expressives authentiques, ne relevant pas d'une simulation actée : le corpus E-Wiz basé sur le protocole SoundTeacher. Ainsi le sous-ensemble des stimuli qui est utilisé dans notre étude reprend les objets qui ont déjà été sélectionnés et mesurés perceptivement par des juges français (G. De Biasi, 2001). Pour rappel le corpus SoundTeacher a été auto-annoté par chacun des 17 sujets qui avaient en particulier produit ces micro-expressions. Pour 6 de ces sujets, une description phonétique a été annotée (Vanpé, 2011). Etant donné que nous nous situons dans une approche où nous ne souhaitons pas définir d'étiquettes reliées à une théorie spécifique dès le départ, les auto-annotations naïves ont été un moyen de nommer les différentes micro-expressions en faisant intervenir la subjectivité réflexive mais également naïve des sujets (Aubergé et al, 2006). En effet, comme nous avons une démarche purement empirique, la classification et la modélisation des objets sont ce que nous visons à terme donc nous ne pouvons pas faire un choix arbitraire et purement

subjectif d'étiquettes à cette étape de notre étude. Par contre, le fait d'avoir une description réflexive du sujet naïf permet d'avoir le ressenti de la personne, ce qui nous permet de rester dans une perspective communicative plus intéressante qu'un choix « expert » sans véritable fondement scientifique.

D'autre part, des expériences de perception sur les micro-expressions, notamment celles de G. De Biasi (cf. Chapitre3 section4 §5) et de R. Signorello (cf. Chapitre3 section4 §4) nous ont fournies des bases d'analyse. Pour R. Signorello, des hypothèses sur l'existence d'une influence des dimensions socioculturelles d'une personne dans sa perception des micro-expressions ont été illustrées. Pour G. de Biasi, une étude contrastée permettrait de mettre en exergue des éléments qui ne peuvent être observés dans le cadre d'une seule origine ethnique. Les deux auteurs proposaient ainsi en termes de perspectives, des pistes d'études avec des locuteurs de cultures éloignées de celle des Français.

Pour répondre à cette proposition, ce mémoire présente un protocole expérimental qui s'appuie sur le même test de perception que celui effectué par G. de Biasi mais s'appliquant à des Japonais. La conservation du même test permettra ainsi de faire une étude comparative des réponses françaises et japonaises.

### ***Protocole et outil expérimental***

Cette section proposera d'une part la description du protocole expérimental en termes de synopsis et d'autre part l'outil qui a servi à la mise en place du test, avec notamment l'adaptation de l'interface en Japonais et les modifications algorithmiques qui en ont découlées.

#### **Les outils utilisés**

Le test de perception de cette étude est codé avec le logiciel Livecode version 5.0.2 de la société américaine Runtime Revolution (créé en 1997). Cet outil permet la création d'applications de faibles tailles, portables sur tous les systèmes d'exploitation principaux à savoir Windows, Mac et Linux, de même que des applications de Smartphones (Android et iPhone). La programmation s'appuie sur un langage utilisant des commandes sous formes de phrases pleines anglaises.

Le programme de G. de Biasi avait quant à lui été développé avec le logiciel Revolution, une ancienne version de Livecode. La reprise du programme n'a toutefois posé aucun problème puisque la syntaxe du langage de programmation n'a pas changé.

Durant cette année, j'ai développé le programme sur un jeton d'essai de Livecode. La version finale du programme a elle été compilée avec une version complète du logiciel à l'aide d'un chercheur du LIMSI (Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur) de Paris, A. Rilliard, avec qui j'ai eu l'occasion de travailler tout au long de ce mémoire.

Enfin le logiciel Livecode utilise par défaut le lecteur multimédia QuickTime Player de la société Apple pour la lecture des vidéos qui seront nos stimuli. Il a donc été nécessaire de prévoir l'installation de ce lecteur sur toutes les machines qui ont servi à effectuer la manipulation.

### **Description générale**

L'expérience qui est proposée dans ce mémoire est un test de perception de micro-expressions produites par des Français dont la reconnaissance est demandée à des sujets Japonais. Les stimuli qui sont utilisés sont tous identiques à ceux qui ont été sélectionnés dans les travaux de G. De Biasi afin de permettre la comparaison des données.

Ce test se décompose en trois parties:

- Une présentation de stimuli uniquement auditifs.
- Une présentation de stimuli uniquement visuels.
- Une présentation de stimuli dans la bimodalité audiovisuelle.

Les stimuli utilisés dans chacune des parties sont des vidéos en format .avi durant en moyenne 2 à 4 secondes. Le format a été choisi pour sa compatibilité avec Livecode et QuickTime ainsi que son rapport qualité/taille puisque nous voulions que l'application supportant le test soit facilement transférable mais d'assez bonne qualité. En effet, les micro-expressions étant principalement des souffles, des soupirs ou de petites vocalises sur de courtes durées, la qualité de l'audition est importante.

Pour la modalité visuelle, un cache a été placé devant la vidéo et pour la modalité auditive nous avons baissé le niveau sonore au maximum pour que ce dernier soit nul. Cela nous a donc permis de réduire le nombre de fichiers vidéos et d'éviter les traitements de modifications sur les fichiers d'origine.

Pour chaque modalité, nous avons sélectionné les stimuli de six locuteurs. La présentation des stimuli se fait locuteur par locuteur. Cependant l'ordre d'apparition des

stimuli pour un locuteur donné ainsi que l'ordre d'apparition des locuteurs eux-mêmes sont variés. En revanche, comme Livecode ne permet pas de mélanger aléatoirement les stimuli puisqu'il fonctionne sur des fichiers externes définissant l'ordre de présentation du corpus, nous avons créé 9 patterns<sup>6</sup> différents pour lister les fichiers de micro-expressions présentées.

Cette variation de présentation permet d'éviter les effets d'apprentissage. Néanmoins, nous avons une spécificité dans l'ordre de présentation pour la modalité auditive. En effet, parmi la liste des stimuli que G. De Biasi a choisie dans son test, nous avons des productions phonologiques, voire même pré-lexicalisées telles que des « euh », des « bah » ou encore des « hello ». Or nous partons de l'hypothèse que ces objets sont culturellement marqués puisqu'ils ont un formalisme écrit. Ainsi pour la modalité acoustique et les groupes de stimuli de chacun des locuteurs, nous avons tout d'abord présenté les stimuli non phonologiques (tels que les souffles) puis les objets lexicalisés. Cela nous permettra ainsi de vérifier si nous pouvons mettre en évidence avec les Japonais une frontière perceptive sur la dimension culturelle des micro-expressions, telle que l'avait suggérée R. Signorello dans son étude.

D'autre part, la présentation d'un stimulus n'est réalisée qu'une seule fois et suite à chaque écoute, le sujet choisit une étiquette parmi les auto-annotations fournies par les locuteurs des stimuli. Pour pouvoir étudier la perception interculturelle, les sujets choisissent également si le stimulus correspond pour eux à une production de natif ou d'étranger. A cela s'ajoute la définition du degré de confiance que chaque sujet évalue de sa propre réponse.

Les données validées par les sujets sont automatiquement transcrites dans un fichier .txt au format de type csv (dont les séparateurs sont des tabulations), facilement transposable dans un fichier Excel<sup>7</sup>.

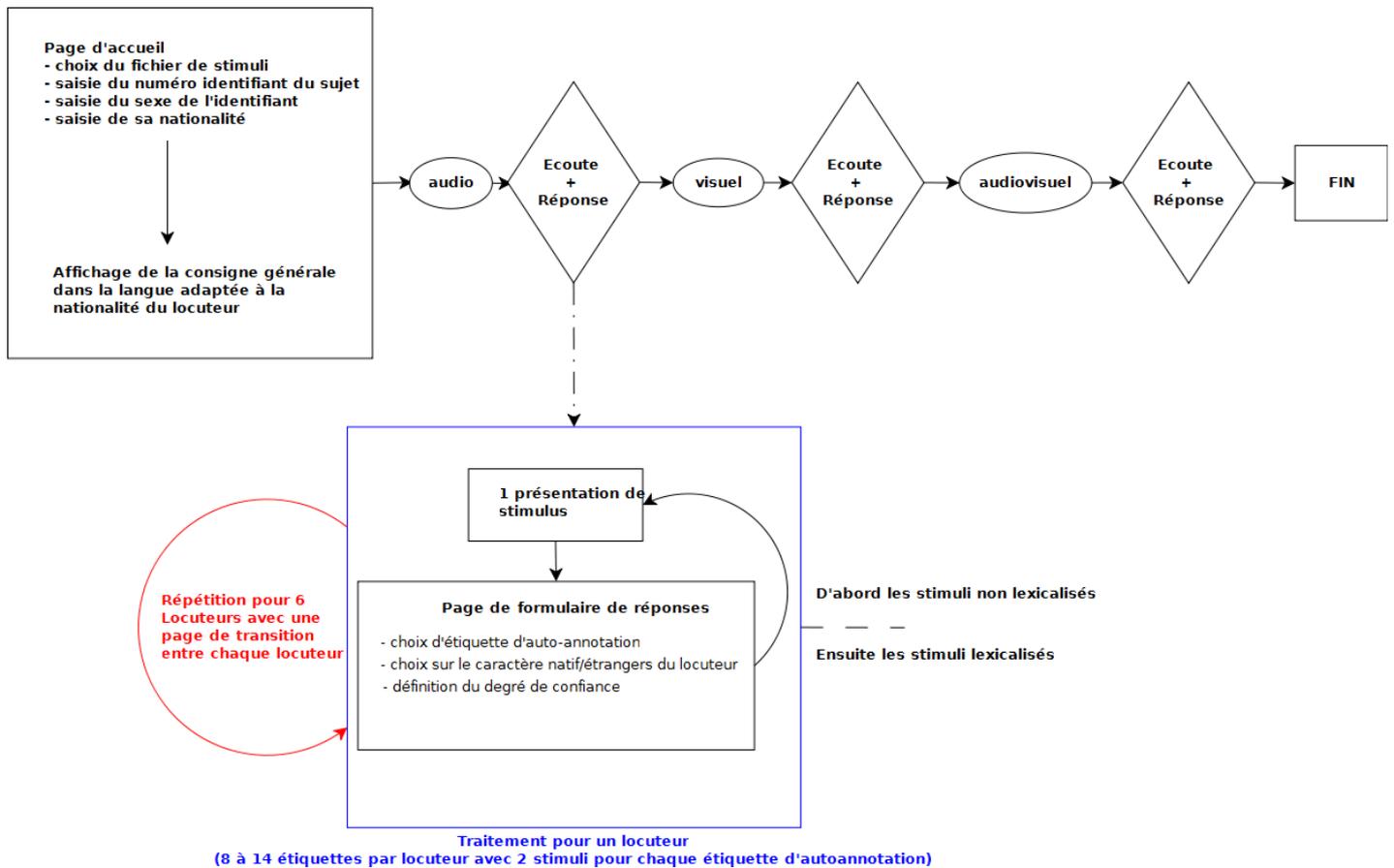
Pour ce test, nous avons sélectionné 4 à 8 étiquettes d'auto-annotations illustrées chacune par deux stimuli. Au total nous avons 66 étiquettes, tout locuteur confondu, soit un total de 396 stimuli traités (pour nos 3 modalités) par chacun de nos sujets japonais.

Un diagramme qui permet de résumer le déroulement du test de perception comme il a été décrit ci-dessus est présenté à la Figure 2.

---

<sup>6</sup> Annexe CD : testPerception\\ManipRevo\\listes

<sup>7</sup> Tous les résultats bruts au format.txt se trouvent en Annexe CD : FichiersTxt



**Figure 2 : Diagramme de modélisation du test de perception**

### Interfaces du test de perception

Suite à la description du test de perception, nous allons présenter l'interface correspondant aux étapes illustrées dans la Figure 2.

Tout d'abord, au lancement de l'application<sup>8</sup> que nous avons nommée « testM2 », une page d'accueil apparaît avec un bouton « Init » qui, quand on clique dessus, permet de faire l'initialisation du test. Cette initialisation consiste en la saisie de quatre éléments :

- Le fichier de test que nous souhaitons utiliser (choisi à partir des 9 fichiers patterns cités dans le paragraphe précédent).
- Le numéro d'identifiant du sujet qui effectue le test.
- Le sexe du sujet.
- Et sa nationalité (avec un choix possible entre Français et Japonais).

<sup>8</sup> Si vous êtes sous PC => Annexe CD : testPerception\\ManipRevo\\Windows\\TestM2.exe  
Si nous êtes sous Mac => Annexe CD : testPerception\\ManipRevo\\MacOSX\\TestM2.app

Une fois que la saisie est validée, un texte de présentation apparaît dans la langue maternelle du locuteur ainsi qu'un bouton « Commencer le test ». Cette étape correspond à la phase d'habituation du test. L'interface est présentée sur la Figure 3.

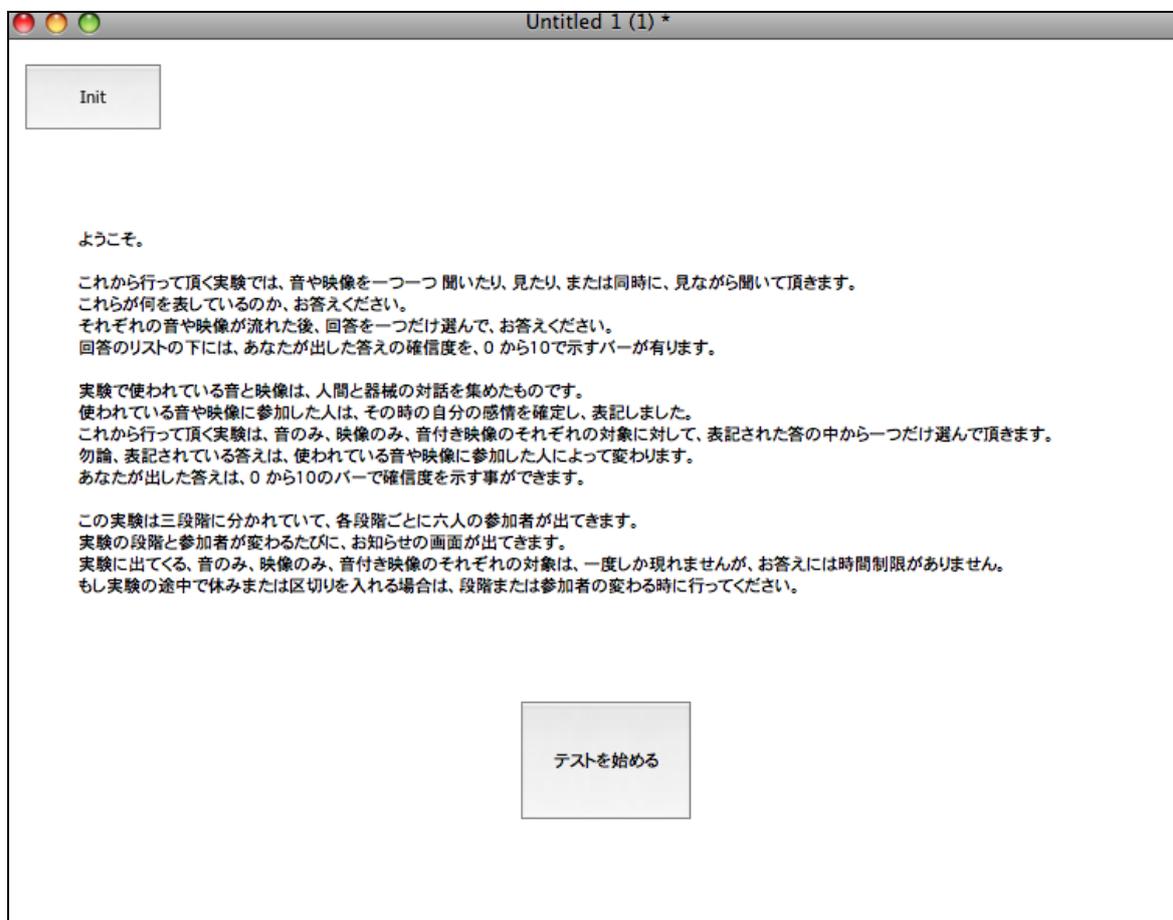


Figure 3 : Image de la page d'accueil du test après initialisation pour un sujet japonais

Une fois que nous avons commencé le test, une page de transition apparaît avec un message introduisant la modalité de présentation des stimuli (soit auditive, visuelle ou audiovisuelle), et/ou le changement de locuteur. Ce court message s'accompagne d'un bouton « suivant » qui permet de lancer le stimulus pour une unique écoute. Voici donc en Figure 4 l'interface d'une page de transition.



**Figure 4 : Exemple d'interface d'une page de transition du test de perception pour un japonais**

Enfin après écoute du stimulus, nous arrivons sur une page de formulaire où le sujet doit répondre sur 2 points, un point supplémentaire étant également réservé à la modalité acoustique (\*) :

- Une étiquette d'auto-annotation qu'il pense correspondre au stimulus qu'il a entendu et/ou vu.
- Un degré de confiance à la réponse apportée qu'il définit à l'aide d'une barre graduée.
- (\*) Dans le cas de la modalité acoustique, choisir si le stimulus entendu est plutôt natif ou plutôt étranger

Voici donc cette interface en Figure 5.

何かに集中している       びっくりしてもイライラする  
 集中しているけど詰まらない       イライラする  
 詰まらない       分からなくて当然だ!  
 意味が分からない

今の音を発したのは:       日本人       日本人ではない

あなたの答えに対する確信度:

0

次へ

**Figure 5 : Exemple d'interface de formulaire de réponses du test perceptif pour un stimulus de modalité auditive en japonais**

L'alternance entre l'écoute de stimulus et la page de formulaire, ponctuée de pages de transition, est continue jusqu'à l'apparition d'un message « Fin de la manip ». Tous les textes qui apparaissent dans le test de perception se trouvent récapitulés dans les deux langues (français et japonais) en Annexe 1.

### **Adaptation du test en Japonais et modification algorithmique**

A l'origine, le programme de G. de Biasi sur lequel je me suis appuyé était entièrement dédié aux Français. Or son programme présentait, dans l'initialisation du test, une option de choix de nationalité. Ainsi, plutôt que de complètement remplacer les textes français, j'ai décidé de faire une adaptation bilingue de ce programme qui gère à la fois une version française et une version japonaise. Le code de programmation pour cette adaptation est présenté en Annexe 2.

L'une des difficultés majeures pour cette adaptation était la gestion des caractères japonais, incompatibles dans la programmation elle-même, ce qui impliquait qu'on ne pouvait pas directement coder du texte japonais. Il fallait donc trouver une alternative pour remplacer les textes français et la solution fut d'utiliser des images textuelles de Livecode.

En effet, sur l'interface nous avons la possibilité de choisir d'insérer des images de texte, une option qui propose une large gamme de caractères et notamment les polices japonaises. Ainsi au lieu de faire des structures conditionnelles jouant sur l'apparition de tel texte dans telle condition, j'ai adapté l'affichage du message par l'affichage d'images. Pour être systématique, j'ai également transformé sous cette forme le texte français pour avoir une unité dans la structure algorithmique du programme.

Par ailleurs, les traductions elles-mêmes étaient relativement difficiles à élaborer, en particulier pour les étiquettes d'auto-annotation qui s'appuient sur du vocabulaire subjectif et vernaculaire de sujets naïfs. Pour pouvoir avoir des étiquettes qui soient les plus proches de la compréhension naïve qu'on pourrait en avoir dans un contexte français, j'ai consulté quatre Japonais non spécialistes et naïfs, entièrement bilingues français/japonais pour obtenir une traduction consensuelle qui ne soit pas littérale mais qui essaie de garder les nuances de ponctuation ou de vocabulaire familier à l'endroit où ces derniers étaient utilisés dans les étiquettes. Ainsi, nous restons proches d'une compréhension et d'une interprétation naïve qui soient accessibles aux Japonais.

Le fait d'utiliser des stimuli français sur des Japonais nous permet de poser l'hypothèse que les micro-expressions sont bien des objets linguistiques, culturellement marqués. Donc leur discrimination en termes de plus natif ou plus étranger devrait donner des informations sur la dimension culturelle de ces objets. Ainsi, nous avons ajouté une question supplémentaire portant sur cette perception, mais uniquement dans le cadre de la modalité acoustique. En effet, l'information visuelle étant trop discriminante (les visages des sujets étant clairement non japonais), nous n'avons pas posé cette question pour la modalité visuelle et audiovisuelle. Cette adaptation a également été réalisée pour la version française en vue d'éventuelles perspectives, qui pourraient être de mélanger des stimuli également produits par des Japonais pour avoir un véritable protocole bilingue.

Finalement sur l'interface finale, nous avons deux versions du test (une française et une japonaise) qui apparaissent en fonction de l'initialisation de l'application de test (en particulier après le choix de la nationalité des auditeurs). Sur ces étapes de paramétrage, il n'a pas été possible de faire une adaptation bilingue car les fenêtres ne supportaient que les caractères latins. Pour que les consignes soient universellement compréhensibles, les indications de saisie pour l'initialisation ont été données en anglais, une langue plus véhiculaire que le français pour les sujets japonais.

### ***Contexte de recueil de données***

La recherche de participants au test de perception a été une tâche difficile de mon étude. En effet, il était bien plus pertinent de trouver des Japonais dont le contact avec les langues européennes soit infime voire même inexistant pour le français. Le recueil de données a donc été réalisé à distance pour permettre l'accès à des Japonais n'ayant pas ou peu de contact direct avec la langue française afin de véritablement observer les différences culturelles de la perception des micro-expressions. Ce contexte a ainsi introduit des contraintes particulières qui m'ont poussée à créer des supports supplémentaires pour pallier le manque de présence des expérimentateurs auprès des sujets.

#### **Les contraintes de distance avec les sujets**

Les moyens de communication privilégiés pour la recherche de contact ont été la messagerie électronique, le téléphone et les sessions de visiophonie. Cependant il ne s'agit pas là du plus contraignant car la prise de contact reste une étape difficile de la recherche, même en présence des sujets.

En revanche, ce qui a été plus compliqué, était la gestion du niveau informatique des participants ainsi que l'adaptation aux outils qui étaient à leur disposition. En effet, contrairement au présentiel où il suffit qu'un programme marche pour une machine donnée, le protocole à distance impose que le sujet soit lui-même doté d'un ordinateur, ce qui implique donc d'anticiper la manipulation d'autant de machines différentes que de nombre de Japonais interrogés.

Par ailleurs l'installation et l'initialisation du programme, bien qu'il ne soit pas complexe, reste tout de même une tâche difficile et contraignante avant même d'effectuer le test. Il est donc nécessaire d'avoir des consignes claires et de nombreux supports pour que cette installation se passe au mieux.

Enfin le temps imparti pour la recherche de données et le recueil de données devait être plus important dans ces conditions car la mise en place et la communication sont plus lentes à distance.

#### **Les solutions apportées aux contraintes d'étude à distance**

Concernant la compatibilité du programme avec les ordinateurs des sujets, Livecode fut un choix judicieux car quasiment toutes les personnes interrogées n'ont pas eu de problème pour l'ouverture du programme. Il semblerait que certaines versions de

Windows Vista fonctionnent tout de même mal avec l'application, ce qui a donné lieu à quelques pertes de sujets potentiels.

La contrainte de temps de développement imposée par la version d'essai du logiciel m'a incité à programmer l'application dans un délai très court. En contrepartie, cela m'a permis d'allonger le temps de récolte des données. La réduction du temps de programmation à tout de même pour conséquence une perte de la robustesse de l'application puisqu'il était possible d'ajouter encore plus de contrôles sur le paramétrage de l'outil. En effet, quand le paramétrage n'est pas effectué de manière attendue, notamment quand le fichier de stimuli n'a pas correctement été choisi, l'application démarre puis s'arrête plutôt que de redemander un bon paramétrage. Quand le sujet oublie de répondre (par inattention ou par choix) nous avons un vide dans nos données alors qu'un programme plus robuste aurait pu forcer le choix des réponses ou aurait permis de proposer une possibilité de non réponse. Pour pallier ces problèmes de robustesse, j'ai mis en place des documents supports détaillés, des consignes spécifiques et un suivi personnalisé pour chaque sujet, cela ayant été possible grâce à l'augmentation de mon temps de récolte.

Voici donc les étapes du protocole général avec : la mise à disposition de l'outil, le contact avec les sujets, l'installation de l'application, le recueil des réponses et leur récupération. Tout d'abord, l'application a été mise en ligne pour que le dossier correspondant soit téléchargeable par un simple lien hypertexte. La plateforme utilisée pour cela est celle de l'outil de développement Dropbox créé en 2008.

Lors du premier contact (essentiellement par messagerie électronique) j'ai communiqué un document d'appel à participation<sup>9</sup> (cf. Annexe 3) et un mode d'emploi<sup>10</sup> (cf. Annexe 4). Le document d'appel à participation, après acceptation des modalités de l'expérience, donne comme consigne de renvoyer un mail de confirmation de participation. Une fois la confirmation réceptionnée, nous renvoyons un « numéro de fichier » à traiter (parmi les 9 qui définissent les patterns d'ordre d'apparition des stimuli), un « numéro d'identification de sujet » ainsi qu'une vidéo explicative<sup>11</sup> (disponible dans l'Annexe CD) créée à partir d'une capture vidéo de l'écran et un enregistrement sonore de mes propres explications.

---

<sup>9</sup> La version japonaise correspond à l'Annexe3A et sa traduction française à l'Annexe 3B

<sup>10</sup> La version japonaise correspond à l'Annexe4A et sa traduction française à l'Annexe 4B. Cette annexe reprend tous les détails de l'interface et décrit l'ensemble des étapes qu'a dû réaliser le sujet donc il complète la section « Protocole et outil expérimental ».

<sup>11</sup> Annexe CD : videoExpli\test.wmv

Cette vidéo a été réalisée à l'aide du logiciel libre CamStudio version 2.1, en licence GNU GPL. Elle a uniquement été produite sous l'interface Windows 7, sachant que la plupart des sujets étaient des utilisateurs PC. En contrepartie, les copies d'écran que j'ai réalisées pour le mode d'emploi ont été prises de l'interface d'un Mac pour avoir la plus grande couverture d'explication possible, sans avoir à faire plusieurs versions d'un même document, considérant le temps qui m'était imparti. La vidéo illustre ainsi les étapes de réception du mail avec les numéros utiles à l'initialisation du document, de téléchargement du dossier d'application, de sa décompression puis de l'initialisation jusqu'à l'affichage du test et l'exemple d'une première question. Les explications orales sont identiques à celles données dans le mode d'emploi et sont entièrement en japonais.

Une fois le test réalisé par les sujets, ses réponses sont directement transcrites en format .txt dans un dossier « resultats » présent dans le dossier<sup>12</sup> d'application téléchargé. Nous demandons donc le renvoi de ce fichier pour récupérer les réponses des sujets.

Par précaution, dans la version mise en ligne, nous avons protégé le dossier « stimuli » pour que les sujets n'aient pas accès aux stimuli en dehors de l'expérience pour récupérer des réponses pertinentes aux questions portant sur la provenance native ou étrangère du locuteur du stimulus ou ne pas avoir une influence externe lors du changement de modalité.

D'autre part, des consignes supplémentaires ont également été données par téléphone, mail ou visiophonie afin de s'assurer des conditions d'expérimentation. Les points sur lesquels j'ai assisté étant les suivants:

- Réaliser le test au calme (lieu non bruyé) à un moment où les sujets sont disponibles sur une durée assez longue, puisqu'en moyenne l'expérience dure une heure pour les Japonais.
- Réaliser le test seul et utiliser un casque et/ou bien augmenter le volume sonore de son ordinateur.
- Convertir le mode de saisie de caractères pour avoir le bon encodage des caractères latins, sans quoi l'initialisation de l'application ne se fait pas correctement.

---

<sup>12</sup> Ce dossier se trouve dans l'Annexe CD : testPerception

Enfin quand une application fonctionnait correctement sur une machine, j'ai essayé dans la mesure du possible de rassembler des sujets autour de celle-ci pour éviter les téléchargements et la mise en place de l'application supplémentaires sur d'autres ordinateurs. De même la sollicitation des sujets qui avaient déjà réalisé l'expérience permettait également de fournir une aide aux nouveaux participants, surtout pour l'étape d'initialisation de l'application (et d'uniquement celle-ci).

### ***Profils des sujets japonais***

Bien qu'il y eu une volonté de sélection de sujets pour que ces derniers aient le moins possible de connaissance de la langue et de la culture française (voire même plus généralement de langues ou de cultures européennes), les effets de mode des Japonais pour la France, ainsi que le protocole à distance n'ont pas permis de filtrer entièrement les sujets visés. En effet, lors du colloque « Francophonie au pluriel »<sup>13</sup> de 2001, il a été rapporté que même si le Japon est loin d'être un pays francophone, le français reste une langue relativement bien étudiée au niveau universitaire. Dans les établissements proposant cette formation qui accueillent environ 700 000 étudiants, 25% d'entre eux choisissent un cours de français langue seconde, alors que l'archipel est très fortement marqué par le monolinguisme.

Par ailleurs, un article du Monde<sup>14</sup> décrit que sur l'année 2006, alors que le nombre de touristes japonais avaient diminué de 3%, 650 000 Japonais sont venus visiter la France, et ce chiffre reste aussi important voire augmente chaque année.

### **Questionnaire**

Pour pallier le problème des sujets japonais qui ont des connaissances en français bien qu'ils se déclarent non francophones, nous avons décidé de faire un point sur les connaissances linguistiques ainsi que les habitudes culturelles (en lien avec le français), de tous les Japonais qui ont passés le test de perception. En effet, en fonction de ces connaissances il pourrait y avoir des influences sur les réponses de l'expérience.

---

<sup>13</sup> Colloque « Francophonie au pluriel » organisé par l'Année Francophonie Internationale en Sorbonne du 17 au 20 mai. Les informations sont tirées de l'article « Comment promouvoir le français en zone non-francophone » de Nobutaka MIURA. Article disponible en ligne à l'adresse : <http://www.potomitan.info/atelier/miura.html> (consulté le 01/05/2012)

<sup>14</sup> Article « Tourisme : vaste campagne française au Japon pour faire oublier les émeutes » daté au 25/04/2007 consulté en ligne : <http://japon.aujourdhuilemonde.com/tourisme-vaste-campagne-francaise-au-japon-pour-faire-oublier-les-emeutes> (01/05/2012)

Pour définir les paramètres nuisibles, j'ai réalisé un questionnaire linguistique<sup>15</sup> (cf. Annexe 5). Celui-ci a été rempli par chacun des participants du test de perception, suite à leur participation à l'expérience. Ce document s'intéresse notamment au niveau de français des sujets, aux études de français qui ont été éventuellement effectuées, à la proximité des sujets avec des personnes françaises de différentes pratiques linguistiques, à l'expérience de séjours en France, ainsi qu'aux loisirs tels que la musique, les films et les livres en français. Les réponses ont entièrement été anonymisées (avec l'identification utilisée lors du test de perception) puis convertie au format PDF. Ces résultats sont disponibles dans l'Annexe CD<sup>16</sup>. Nous décrivons dans la suite de cette section les sujets japonais en fonction de leurs réponses au questionnaire.

### **Description générale des sujets**

Au total 23 sujets ont participé à l'expérience dont deux étudiantes chinoises qui étaient en échange linguistique dans une université japonaise. Pour ne pas fausser les résultats, nous n'avons pas ajouté les réponses de ces deux sujets dans le corpus final de réponses des Japonais. Malgré tout, ces étudiantes qui n'étaient pas prévues dans nos manipulations expérimentales, se sont spontanément portées volontaires, donc nous avons conservé ce matériel en vue de perspectives à apporter à notre étude. En effet, nous pourrions étudier l'influence de la proximité culturelle de locuteurs dans la perception de micro-expressions et ce premier échantillon pourra peut-être donner un premier aperçu des tendances de perception chez des sujets plus proches de la culture japonaise que les Français.

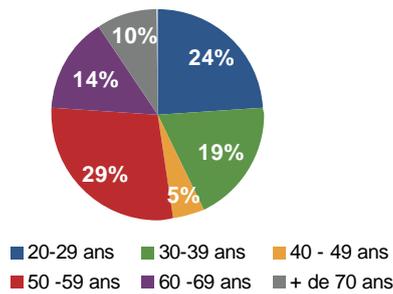
Concernant nos 21 sujets japonais, nous avons recherché des adultes, leur répartition par âge et par sexe est illustrée sur la Figure 6.

---

<sup>15</sup> Un exemplaire de ce questionnaire en bilingue est disponible en Annexe 5. En revanche, c'est une version monolingue qui a été remplie par les sujets de l'expérience.

<sup>16</sup> Annexe CD : Questionnaires

Diagramme de répartition des sujets par tranche d'âges



Graphique de répartition des sujets par tranche d'âges et par sexe

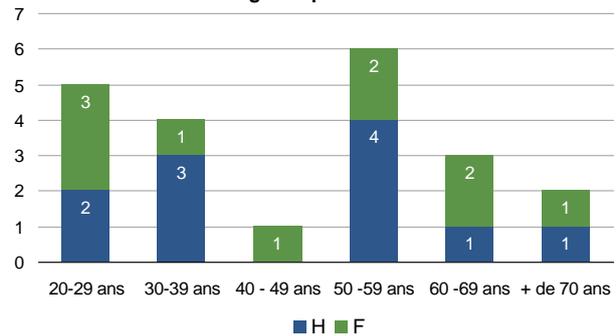


Figure 6 : Figures de répartition des sujets en fonction de leur âge (à gauche), en fonction de leur âge et leur sexe (à droite)

Notre échantillon est relativement équilibré au niveau de l'âge des locuteurs. En effet, la moitié du groupe représente des adultes moyens de 30 à 59 ans et nous avons également les extrêmes de la classe adulte avec les jeunes adultes (de 20-29ans) d'une part, et les personnes plus âgées de plus de 60 ans qui représentent chacun un quart de la population interrogée.

La répartition homme/femme est également équilibrée puisque nous avons obtenu la réponse de 11 hommes et de 10 femmes, répartis dans toutes les classes d'âges sauf celle de 40-49ans pour les hommes, mais qui est compensé par ceux de la classe 50-59ans.

### Connaissances et expériences linguistiques

Dans ce paragraphe, nous décrirons à la fois les pratiques en langues étrangères des sujets ainsi que leur connaissance et/ou apprentissage du français.

D'après le questionnaire, seul le quart de la population interrogée n'a pas d'expériences en langues étrangères (cf. Annexe 6A) et le reste pratique essentiellement l'anglais que l'illustre le Tableau 1. Quatre personnes affirment également utiliser ou avoir appris le français. Nous pouvons également observer la pratique de langues minoritaires comme le thaïlandais, le chinois, le tchèque ou encore l'italien dans notre échantillon.

Un groupe de sujets à majorité pratiquante de langues étrangères est une conséquence liée à notre protocole expérimental. En effet ce dernier a nécessité que les sujets sachent utiliser des ordinateurs et le niveau socioculturel qui sous-tend alors cette

connaissance s'accompagne par la pratique des langues, en particulier de l'anglais qui est appris comme première langue vivante étrangère au Japon, et cela dès l'école primaire.

Sujets	Langues étrangères étudiées par les sujets
01-M-JP	10 ans d'anglais
02-F-JP	bilingue anglais + 4 ans de français à l'université
02-M-JP	anglais
03-F-JP	thaïlandais - anglais - quelques mots en français
04-M-JP	anglais - chinois
05-F-JP	anglais
05-M-JP	anglais
06-F-JP	anglais
06-M-JP	anglais
08-F-JP	anglais - tchèque (cours de cuisine)
12-F-JP	3,5ans français + 3,5 ans italiens + anglais
13-M-JP	un peu d'anglais
14-M-JP	anglais
22-F-JP	français

**Tableau 1 : Illustration des langues pratiquées par les sujets ayant une expérience en langues étrangères**

D'après l'Annexe 6B, 57% des sujets soit 12 personnes affirment n'avoir aucune expérience de la langue française. Il s'agit de la population idéale que nous visions. En revanche, nous avons également 33% des sujets soit 7 personnes qui affirment avoir déjà appris le français mais ne le maîtrisent pas, alors que 10% des sujets soit 2 personnes maîtrisent le français mais ne l'utilisent pas dans leur pratique quotidienne. Ces personnes, bien qu'elles aient eu une initiation sur le français, ne se considèrent pas comme étant francophones. De fait, aucune personne n'est francophone déclarée, ni pratiquante du français dans son quotidien parmi les sujets interrogés.

Parmi les personnes qui se disent avoir appris le français, il faut compter les apprentissages informels de quelques mots ou phrases lors de séjours touristiques ou lors

des rencontres de courte durée avec des Français. D'ailleurs, quand nous demandons aux sujets s'ils ont appris le français, seuls 5 sujets ont répondu à l'affirmative pour un cadre formel d'apprentissage du français (cf. Tableau 2). Pour 10-F-JP, l'apprentissage a eu lieu il y a fort longtemps, et cette dernière n'en garde aucun souvenir.

Donc parmi tous les sujets, il nous faudra particulièrement rester attentif dans l'analyse des réponses des cinq apprenants de français car leur connaissance de la langue française pourrait éventuellement modifier les tendances des résultats du test de perception.

	Type de structure	Type de cours	Durée de la formation	niveau atteint
<b>01-M-JP</b>	cours de radio	cours de radio	3 / semaine - 1 an	salutations
<b>02-F-JP</b>	université	cours généralistes	4 ans	conversation quotidienne
<b>04-F-JP</b>	université	cours généralistes	5/ semaine pendant 2 ans + 3/ semaine pendant 2ans	compréhension globale d'un texte en lecture
<b>10-FJP</b>	lycée	cours généralistes	3 ans	plus aucun souvenir
<b>12-M-JP</b>	université française	1 cours / semaine de 1h30	2 ans	salutations

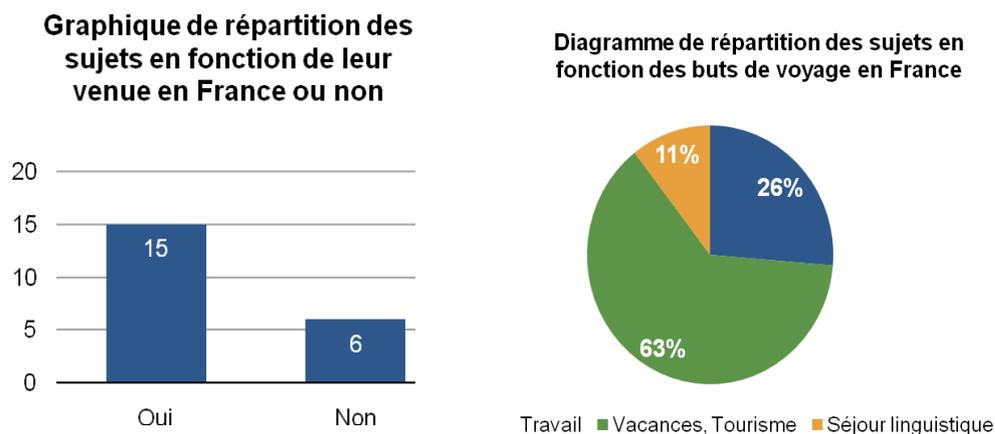
**Tableau 2 : Récapitulatif de la formation en français des sujets japonais dans un cadre d'apprentissage formel**

### Séjours en France

Ce facteur a bien entendu été le plus problématique dans la sélection des sujets pour notre test car comme nous l'avons vu en introduction de cette section, la France est l'une des destinations préférées des touristes japonais. De plus, il existe de nombreuses collaborations franco-japonaises ainsi que des jumelages entre villes françaises et japonaises, qui sont à l'origine de déplacements professionnels assez fréquents de Japonais en France.

Dans la Figure 7, nous voyons ainsi que seul un quart des sujets interrogés ne sont jamais venus en France, les autres se sont essentiellement déplacés pour le tourisme (63%

soit 12 personnes). Assez peu sont également venus pour le travail (26% soit 5 personnes) et enfin deux personnes ont effectués un séjour linguistique. Donc au final, beaucoup de personnes de notre échantillon se sont déplacées en France mais elles ne semblent pas avoir eu de nouvelles opportunités d'apprentissage linguistique lors de ces déplacements. En revanche, il faut voir si cette venue n'aurait pas de lourdes conséquences sur les liens qu'ils ont pu tisser avec des Français, jouant alors sur leur exposition à des stimuli de Français.



**Figure 7 : Illustration des expériences de séjours en France (à gauche) et répartition des buts des voyages effectués (à droite)**

D'après le Tableau 3, la majorité des personnes qui sont venues en France a fait entre 1 à 6 séjours, pour une durée pouvant aller de quelques jours jusqu'à deux semaines. En revanche, certains types de séjours sont plus fréquents et/ou beaucoup plus longs, notamment pour 6 sujets. Parmi eux, nous retrouvons 3 des 5 sujets ayant effectués une formation formelle de français. Il s'agit des sujets 01-M-JP, 02-F-JP et 12-M-JP. Pour les autres apprenants de la langue, sachant que les cours de français sont dispensés par des Japonais ayant étudiés le français et que les méthodes d'apprentissage du français posent un véritable problème au Japon comme en témoigne S. Naito (2002), l'apprentissage de la langue dans un milieu formel n'est pas ce qui peut véritablement changer la perception culturelle et interactionnelle des Japonais. Donc, plus que les apprenants, il s'agit surtout de personnes qui ont une expérience de voyage en France assez soutenue (mais pas seulement), ce qu'il est important de noter, mais bien entendu les personnes ayant fait des études dans ces conditions sont encore plus à surveiller puisque l'on peut supposer qu'elles sont motivées par un véritable intérêt pour la connaissance de la culture française.

Sujets	Fréquence de venue en France	Durée moyenne des séjours	Les villes et régions côtoyées
01-M-JP	7	2 mois pour le plus long séjour	Paris- Rhône Alpe - Sud
02-F-JP	1 à 2 fois par an depuis l'âge de 20 ans	3 jours à 1 mois	Paris - Provence - Alpes
02-M-JP	4	14	Alsace - Paris
03-F-JP	20	14	Paris - Nice-Alsace- Alpes
04-F-JP	3	1-2 semaines	Paris - Provence - Grenoble
05-F-JP	3	14	Rhône-Alpes - Sud
05-M-JP	1	4-5 jours	Paris- Grenoble
06-F-JP	6	1-2 semaines	Paris - Provence - Alpes
06-M-JP	2	quelques jours	Paris - Alpes
08-F-JP	2	3-6 jours	Paris
10-F-JP	3	7 jours	Paris- Lyon - Grenoble
12-F-JP	20	vécu 1,5 ans en France + 2 ans en Belgique + 1semaine*10	Paris-Alpes-Bruxelles
12-M-JP	3	3 mois	Paris - Tours - Grenoble
16-M-JP	??	14 jours	divers
22-F-JP	1	1 an et 10 mois	Paris - Franche Comté

**Tableau 3 :Récapitulatif de la durée, la fréquence et les lieux de séjours en France des sujets japonais**

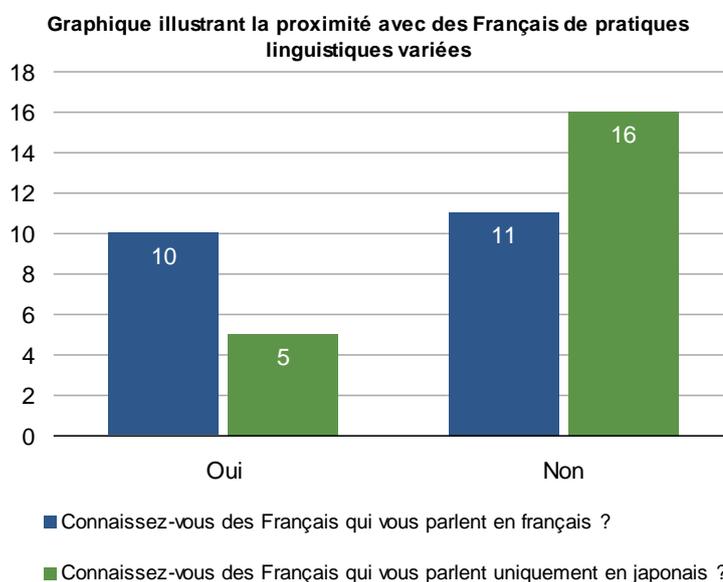
Parmi les non apprenants, trois autres personnes ont fait des séjours de telle manière qu'ils pourraient influencer la perception culturelle des micro-expressions. Il s'agit des locuteurs 03-F-JP, 12-F-JP et 22-F-JP. Si nous revoyons les profils de ces derniers sujets, 03-F-JP et 12-F-JP se considèrent comme apprenants informels ne maîtrisant pas la langue alors que 22-F-JP semble la maîtriser bien que son apprentissage fut informel. En revanche, le but de son voyage était avant tout linguistique d'où cette différence. Gardons donc cela à l'esprit pour voir si les autres critères permettraient de dire si elles auraient une éventuelle influence sur les réponses du test perceptif.

#### **La proximité avec des Français de pratique linguistique variée**

Les questions portant sur cette partie visent tout d'abord à montrer si les sujets ont l'habitude d'entendre des sons produits par des Français, qui peuvent alors par exemple être francisés si le Français n'est pas un bilingue natif (ou s'il s'agit d'un locuteur avec un accent marqué des sons de sa langue maternelle). Mais surtout, du point de vue bimodal et visuel, cela permet également de voir si les Japonais interrogés ont l'habitude de voir des personnes avec des attitudes françaises comme ce serait le cas dans les stimuli à reconnaître dans le test de perception. Dans le cas où ils connaissent des Français leur parlant japonais, il s'agit de savoir s'ils côtoient des interlocuteurs qui se sont accommodés de la culture japonaise ayant alors des attitudes qui vont de pair avec la langue japonaise,

même si les caractéristiques anatomiques qu'ils voient restent ceux de Français. Ces paramètres sont importants à noter puisqu'ils pourront influencer les réponses des sujets dans le test de perception. Voici donc le résumé des résultats obtenus dans la Figure 8.

Nous voyons ici que la part de personnes qui connaissent des Français leur parlant français est assez importante puisque cela représente approximativement la moitié des personnes. En revanche, seules 5 personnes connaissent des Français qui parlent japonais. Le fait qu'autant de monde connaissent des Français est surtout lié au fait que beaucoup de sujets sont allés en France. En effet, outre les personnes qui ont étudiées la langue, certains sujets ont eu l'occasion de faire de nouvelles rencontres avec des Français qui s'adressent à eux en français.



**Figure 8 : Illustration de la proximité des Japonais avec les Français de pratiques linguistiques variées**

Finalement, bien que les sujets aient eu l'occasion de côtoyer des Français natifs, cela est essentiellement dans le cadre de leur voyage et non le fait d'un contact permanent dans leur quotidien ou avec une fréquence soutenue (cela a été vérifié par contact direct avec les participants). Notons que ce sont les 3 apprenants remarquables plus hauts qui font également des séjours longs et/ou fréquents. En plus des critères que nous venons de rappeler, 01-M-JP et 12-M-JP connaissent tous les profils caractéristiques de Français que nous avons proposé dans le questionnaire. 02-F-JP est certainement pour le moment la personne la plus francophone sur le plan de son niveau de français, de sa formation, de ses

voyages ainsi que de sa connaissance de personnes françaises ne parlant que le français avec elle. L'apprenante 04-F-JP qui connaît une personne lui parlant en français est également à surveiller dans les analyses bien qu'elle n'ait pas forcément une maîtrise de la langue car elle a également plusieurs expériences de séjours en France. Nous retrouvons également les sujets 03-F-JP et 12-F-JP qui pourraient influencer les réponses du test car elles connaissent au moins un Français qui leur parle en français. Enfin le cas de 22-F est assez étrange car elle n'a aucun contact avec les Français malgré sa maîtrise de sa langue et son long séjour. Il faudra donc regarder si ses réponses se rapprochent de celles de Japonais ayant le profil idéal que nous recherchions, ou si elle ses tendances sont plus similaires à celles des personnes qui pourraient influencer différemment les réponses du test de perception.

### **La langue française dans les loisirs des sujets japonais**

En termes de stimulation acoustique et audiovisuelle, les chansons et les films nous permettent d'avoir une idée de l'habitué que les Japonais interrogés ont par rapport à des stimuli français. La lecture, elle, est une indication sur la maîtrise de la langue. Voici donc les réponses obtenues illustrées en Annexe 6C.

Dans l'ensemble, les Japonais interrogés ne maîtrisent pas la langue française bien que 7 sujets sur 21 disent lire des œuvres françaises. En effet, les livres lus par ces derniers sont quasiment tous des traductions japonaises d'auteurs français. La seule exception est le sujet 02-F-JP qui confirme sa grande maîtrise du français par la lecture d'œuvres originales. Nous pourrions donc la qualifier de francophone bien qu'elle ne considère pas l'être.

Concernant la musique, le nombre de personnes qui écoutent des chansons françaises sont assez nombreuses puisque 8 sujets sur 13 sont concernés par ce cas. Cependant, même s'ils ont une entrée acoustique des sons du français, la mélodie reste tout de même gênante pour l'accès prosodique de la langue.

En revanche, nous avons encore plus de personnes qui regardent des films français que de sujets qui n'en regardent pas et tous sans exception les visionnent en version originale sous-titrée. De plus, 6 personnes parmi les 11 regardant des films de ce type, affirment en avoir vus « beaucoup » et pour ceux qui ont précisé, ce nombre s'élèverait à plus d'une vingtaine de films. Cela signifie que globalement les sujets interrogés ne maîtrisent peut-être pas le français, mais sont souvent exposés à des stimuli français de part leurs voyages et leurs goûts filmographiques.

Finalement les sujets que nous avons interrogés sont répartis équitablement en termes d'âge et de sexe mais sont assez familiers aux stimuli français. Nous avons tout de même une dizaine de sujets qui répondent bien à nos critères de méconnaissance du français. Cependant, nous devons également rester attentifs aux réponses que donnent certains sujets.

#### ***Conclusion sur le chapitre 4***

Le test de perception effectué est assez classique dans son genre, bien qu'il n'existe pas grand nombre d'études s'appuyant sur des micro-expressions. Afin de répondre à des perspectives d'études antérieures, nous avons choisi d'étudier des sujets Japonais dont la culture sociétale et linguistique s'éloigne de celle du français.

Afin de mettre en évidence d'éventuelles influences sur la dimension culturelle de la perception, nous avons opté pour un protocole à distance avec l'avantage d'accéder à des Japonais non francophones vivant dans leur environnement culturel d'origine. Ce choix concéda également de nombreuses contraintes pour la mise en place et le choix des sujets.

Les Japonais qui se déclarent comme non francophones ont finalement une exposition assez importante à des stimuli français. Il est donc intéressant de compléter les études par un questionnaire linguistique qui nous permet de mieux appréhender les sujets que nous observons et de rester vigilants aux résultats que nous obtenons dans le test.

Ainsi, bien que nous sommes amenés à traiter les résultats des Japonais dans leur ensemble dans le cadre de ce mémoire, des perspectives découlent de l'étude linguistique des sujets. En effet il nous faudra rester attentif aux réponses de certains sujets. C'est notamment le cas pour 02-F-JP qui pour nous a un profil très proche d'un francophone ainsi que 04-F-JP qui semble être relativement exposé à la langue française même si elle n'en a pas une bonne maîtrise. De même, nous avons les sujets 01-M-JP et 12-M-JP qui ont appris le français sans le maîtriser et sont fortement exposés à des stimuli français. Remarquons également 03-F-JP et 12-F-JP qui sont non apprenantes mais ont un contact assez important avec le milieu francophone, et enfin 22-F-JP qui est un cas assez ambigu et que nous devons aussi surveiller.

## **Partie 3**

### **Analyse des résultats expérimentaux**

## Chapitre 5 – Analyse des résultats japonais

A partir de la manipulation expérimentale du test de perception, nous avons réparti les résultats dans des matrices de confusion dont nous avons fait une analyse descriptive statistique, ce que nous présenterons ici. Nous avons également récolté les degrés de confiance des réponses de chaque locuteur qui seront illustrés dans une seconde section. Enfin, concernant les jugements des auditeurs (pour les stimuli en audio seul) sur l'origine japonaise ou étrangère des stimuli, nous évoquerons les résultats dans une dernière partie de ce chapitre.

### *Matrices de confusions et analyses statistiques*

Dans cette section, nous ferons tout d'abord le point sur la manière dont nous avons présenté nos données dans les matrices de confusion et les problèmes qu'elles imposent, puis nous parlerons des analyses statistiques qui ont été mises en œuvre afin de mettre en évidence la significativité de certains résultats et enfin nous présenterons une analyse descriptive des données obtenues dans une dernière partie.

### **Quelques précisions sur les matrices de confusion**

Les matrices de confusion sont des tableaux à double entrées qui s'appuient sur le recueil de données sur des questions à choix finis mono réponse. Elles mettent ainsi en corrélation des réponses attendues avec les réponses effectivement données en observant sa diagonale. De même, elles permettent de regarder la dispersion des réponses en dehors de la diagonale. Toutes les réponses ont donc toutes été résumées dans de telles matrices, où nous trouvons en horizontale les réponses données pour un stimulus (correspondant à une auto-annotation donnée) et à la verticale la dispersion d'une étiquette sur les différents stimuli d'un locuteur. Nous avons donc une matrice pour chacune des modalités (A-V et AV) associée à chaque locuteur soit un total de 18 matrices de confusions japonaises (cf. Annexe 7); à rajouter aux 18 matrices des Français (G. De Biasi, 2011) (cf. Annexe 8).

Pour nos matrices, nous avons dans chaque élément, à la fois le nombre brut des réponses et le pourcentage rapporté au nombre de réponses par stimuli. Le problème est de déterminer à partir de quel niveau une valeur devient significative. G. De Biasi avait réalisé des tests statistiques de type  $\chi^2$  et avait défini comme significatives, uniquement les valeurs deux fois au-dessus du hasard. Dans la comparaison entre sujet français et japonais nous

devrons prendre ce seuil pour référence, mais dans le cas de l'étude des données japonaises seules, il n'est pas pertinent d'utiliser ces valeurs car celles-ci sont très diffuses et moins élevées puisque les stimuli ne sont pas ceux de leur langue maternelle, ce qui leur rend la tâche plus difficile. Ce choix est purement arbitraire et n'est pas véritablement motivé, contrairement aux moyens statistiques que nous allons essayer de proposer pour l'analyse des données japonaises.

Le Tableau 4 est un exemple de matrice de confusion sur les réponses japonaises du sujet T. Les reports significatifs sont illustrés en bleu et les reconnaissances significatives sont en vert. La diagonale grisée correspond aux reconnaissances correctes de chaque type de stimuli.

réponse stimulus	Etonnée		Déception		(Assez) Calme		Angoissée		Hésitation	
Etonnée	3	7,14%	8	19,05%	6	14,29%	15	35,71%	10	23,81%
Déception	2	4,76%	29	69,05%	4	9,52%	5	11,90%	2	4,76%
(Assez) Calme	2	4,76%	9	21,43%	4	9,52%	12	28,57%	15	35,71%
Angoissée	14	33,33%	7	16,67%	8	19,05%	10	23,81%	3	7,14%
Hésitation	7	16,67%	16	38,10%	6	14,29%	6	14,29%	7	16,67%

**Tableau 4 : Exemple d'une matrice de confusion du sujet T (modalité A chez les auditeurs japonais)**

### Objectivation des données par des outils statistiques

Une première manipulation statistique fut la réalisation d'une analyse en composantes principales (PCA : Principal Component Analysis). Il s'agit d'une technique utilisant une table de valeurs corrélées à un ensemble de variables non corrélées, indépendantes du contexte pour retrouver ce que nous appelons les composantes principales, qui illustrent les plus fortes tendances (Abdi, 2004). Les résultats ne donnant pas de résultats très lisibles, nous avons en définitive décidé de travailler sur des analyses de correspondance (CA : Correspondance Analysis) qui est une forme généralisée de la PCA qui donne lieu à un tableau de contingence, soit une illustration de données à partir de comptage, ce qui est bien notre cas. Le but de cette technique est de récupérer deux facteurs de scores illustrant la colonne et la ligne de nos matrices, ce qui semble être le meilleur moyen de représenter leur corrélation. Par ailleurs la lecture de ces données peut se lire sur un graphique de type nuage de points qui rend compte de l'information

essentielle des matrices d'un point de vue statistique, les deux facteurs étant représentés dans le même espace car le calcul se base sur une variance identique (Abdi, 2010).

Pour réaliser ces calculs et ces figures nous avons utilisé le logiciel R<sup>17</sup> version 2.15.0, un logiciel libre sous licence GNU GPL développé par John Chambers et al. Au Bell Laboratories. Pour la manipulation des données via une interface simplifiée, nous avons également chargé le package FactoMineR<sup>18</sup> qui permet d'effectuer des analyses exploratoires de données multi variables, dont la CA.

Donc pour chacune des 18 matrices de confusion des sujets japonais, j'ai déterminé les CA correspondantes (sur la première et la seconde dimension car elles illustrent plus de 70% de la variance), accompagnées de leur dendogramme (qui est une représentation arborescente des regroupements hiérarchiques des données) ainsi que les regroupements (ou *clustering*) hiérarchiques, correspondants au dendogramme de la CA de chacune des matrices. La CA et le regroupement hiérarchique (qui reprend à la fois le *clustering* et le dendogramme), correspondant à chaque matrice, sont regroupés en Annexe 9.

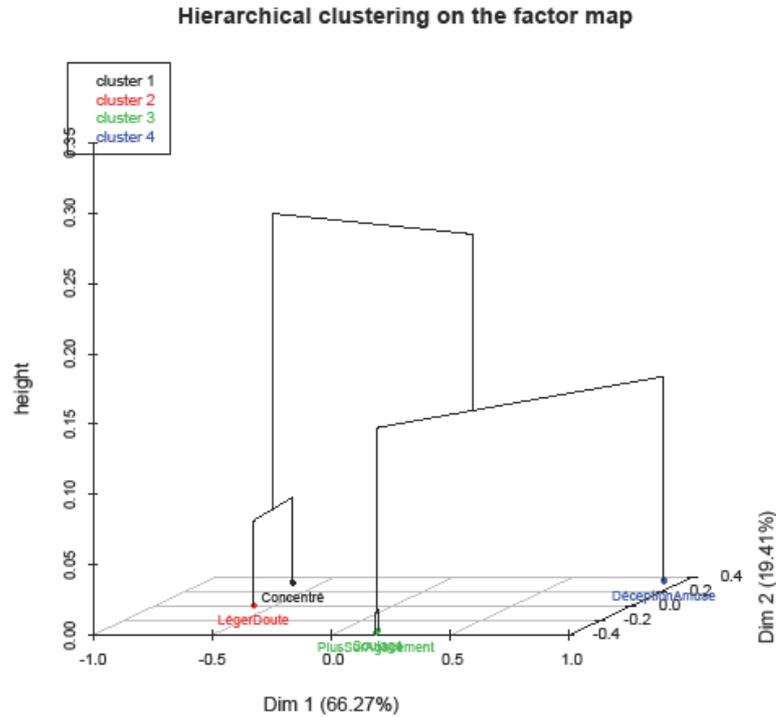
Un exemple du *clustering* hiérarchique des réponses japonaises du sujet T est donné en Figure 9. Cette figure montre par exemple qu'il existe une proximité de comportement entre « léger doute » et « concentré » sur des critères statistiques.

Dans ce regroupement hiérarchique, la couleur des *clusters* n'a aucune valeur car la hauteur du dendogramme a été fixée au hasard. D'ailleurs, nous ne souhaitons pas avoir de *clustering* automatique car, à terme, nous voulons réussir à regrouper des stimuli sur la base d'une observation empirique motivée de données et pas seulement sur une logique statistique. Le regroupement ne nous intéresse donc pas mais la proximité significative de ces données nous intéresse.

---

<sup>17</sup> Le site officiel de l'outil R se trouve à cette adresse : <http://www.r-project.org/index.html> (consulté en avril 2012)

<sup>18</sup> Le site officiel de FactoMineR se trouve à cette adresse : <http://factominer.free.fr/> (consulté en avril 2012) Ce complément est développé et mise à jour par F. Husson, J. Josse, S. Lê, de l'Agrocampus de Rennes, ainsi que par J. Mazet.



**Figure 9 : Exemple de regroupement hiérarchique de la matrice A du sujet T pour les réponses des auditeurs japonais**

Pour compléter ces données, nous avons réalisé des matrices de corrélation correspondant à chaque matrice de confusion. Les matrices de corrélation illustrent des demi-tableaux sous la diagonale, pour mettre en évidence la similitude de comportement des lignes et des colonnes, soit respectivement la dispersion des étiquettes sur un stimulus et la dispersion des stimuli pour une étiquette donnée dans notre cas, ce qui permet de faire le lien entre les différentes variables. Pour cela nous avons calculé les coefficients de corrélation<sup>19</sup> à partir du tableur Excel, la fonction existant dans le logiciel permettant de faire une corrélation par ligne d'une part et une corrélation par colonne d'autre part. Le coefficient de corrélation (dite de Pearson) se calcule par le quotient de la covariance de deux variables (ici une ligne par rapport aux lignes ou une colonne par rapport aux autres colonnes) sur le produit des écart-types de ces variables.

<sup>19</sup> Etant donné que nous ne présenterons pas d'analyses directes dans ce mémoire, les matrices de corrélations seront présentées dans l'annexe CD de ce mémoire.

## Lecture de données statistiques

Sur les CA, dont un exemple est donné en Figure 10, nous voyons l'illustration des lignes en rouge et celle des colonnes en bleu. Ainsi nous pouvons observer comment les stimuli (en rouge) sont reconnus via l'étiquette ou les étiquettes les plus saillantes (en bleu) puisque ces attributions apparaissent à proximité des stimuli correspondants.

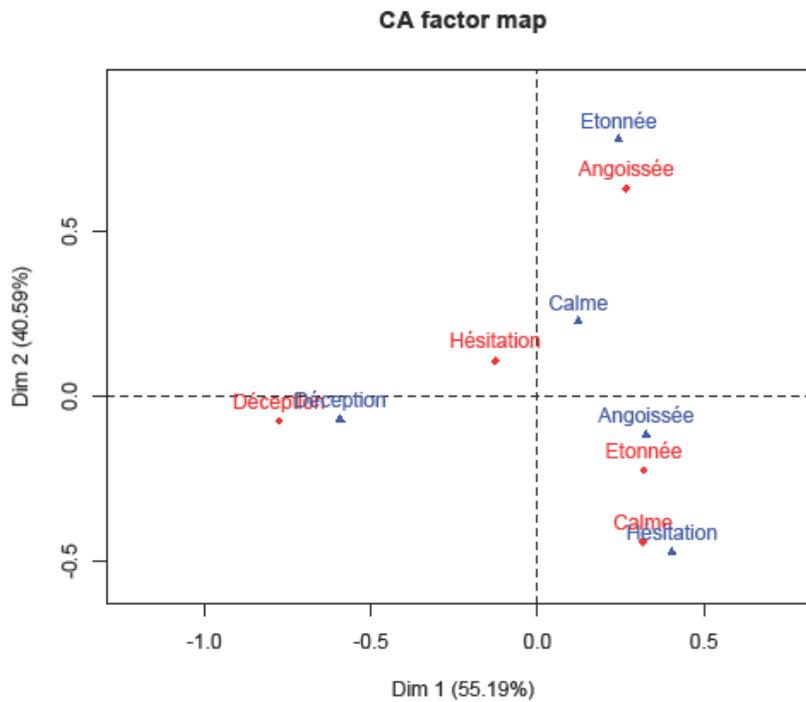


Figure 10 : Exemple de CA correspondant à la matrice A du sujet T des réponses japonaises

Sur les coefficients de corrélation qui varient entre -1 et 1, plus on s'approche des valeurs extrêmes, plus cela signifie qu'on a une similitude dans le profil de dispersion des étiquettes ou des stimuli. Le signe négatif permet de savoir si les deux variables varient ou non dans le même sens mais cela n'a aucune valeur significative pour nos matrices puisque les sens de lecture des lignes et des colonnes importent peu. Le Tableau 5 est un exemple de matrice de corrélation.

Coefficient de corrélation par ligne					
	Etonnée	Déception	Calme	Angoissée	Hésitation
Etonnée	1				
Déception	0,02	1			
Calme	0,81	0,07	1		
Angoissée	-0,37	-0,18	-0,74	1	
Hésitation	-0,1	0,97	0,07	-0,2	1

**Tableau 5 : Exemple de matrice de corrélation sur les lignes pour les résultats japonais de la matrice de confusion en A du sujet T**

### Les principales observations sur les différents sujets

Voici donc une description brève des observations que nous pouvons faire sur chacun de nos locuteurs à partir de l'analyse des différentes figures statistiques vues précédemment (cf. Annexe 7, 9 et Annexe CD pour les matrices de corrélation).

**Sujet T :** Pour ce sujet, le profil de reconnaissance est particulier en audio (A) mais plus proche pour le visuel (V) et l'audiovisuel (AV). Quelque soit la modalité, la « déception » est bien reconnue et les stimuli de « calme » sont confondus avec celle de « l'hésitation », et inversement. Par ailleurs, les stimuli de « angoissée » et « étonnée » sont mutuellement confondus dans les deux sens dans les trois modalités. Pour le visuel, « calme » est également confondue avec l'étiquette « angoissée ». En audiovisuel, la « déception » est également confondue avec de « l'hésitation ». Donc malgré quelques particularités de confusions en V et AV, nous avons des couples de confusion qui sont relativement constants pour le sujet T.

**Sujet S :** Le profil de dispersion de chacune des modalités est très différent. En A, la « surprise » est confondue avec « mission impossible ». Le « but » est bien reconnu mais il est également confondu avec la « concentration ». « Rigoler » est confondu avec « déçue » et « surprise » et les stimuli « concentration » et « mission impossible » sont assez isolés. Enfin « déçu » est un peu confondu avec « rigoler ». En V il existe beaucoup de confusions. Ce qui est clair c'est que « rigoler » est confondu avec « déçue ». La « concentration » est assez bien reconnue mais elle est surtout confondue avec « rigoler ». Sinon « surprise », « le but », « mission impossible » et « déçue » sont des attracteurs entre eux et « surprise » ainsi que « déçue » sont assez bien reconnus. Enfin en AV nous conservons la confusion de « rigoler » par « déçue ». La « surprise » reste bien reconnue et

garde sa confusion avec « mission impossible ». « Mission impossible » est lui confondu avec « le but » alors que la « concentration » l'est avec « rigoler ». Enfin « le but » et « déçue » sont confondues avec la « concentration ».

**Sujet M :** Les profils V et AV de la CA sont relativement semblables alors que celui de A est assez spécifique. Pour cette dernière modalité, « déçue » et « stress » ont des profils très similaires et sont tous les deux assez bien reconnus. « Sûr de moi », « peur » et « doute » sont tous trois confondus avec « doute » et « pas contente - déçue ». Enfin pas contente - déçue » est confondu avec « sûre de moi » mais à des taux assez bas. En V et en AV, « sûr de moi » et « peur » sont attracteurs des étiquettes « pas contente » et « stress ». En plus ces attractions concernent également l'étiquette « déçue » en V et celle du « doute » en AV. Sinon le « stress » est assez bien reconnu en visuel alors que c'est « déçue » qui l'est en AV. Enfin « pas contente - déçue » et « déçue » sont confondus avec « sûr de moi ».

**Sujet J :** Sur les 3 modalités, l'« inquiétude » est assez bien reconnue. En A et en AV, « irritation » et « On se paye ma tête » sont à l'origine de confusions et « ennui » reste un cas relativement isolé. Sinon les profils en A et en AV sont quasiment identiques. En V, l'« irritation » est un peu reconnue mais assez isolée. « On se paye ma tête » et « ennui » ont des profils similaires et « ennui » est légèrement reconnu.

**Sujet N :** Pour ce sujet, les confusions sont diffuses en audio mais plus précises en visuel et davantage en audiovisuel. En A, l'élément qui est clair est l'attractivité de « concentré » qui reçoit « léger doute » et « agacement ». « Soulagé » semble être confondu avec « concentré », « déception » avec « soulagé ». Il semble également y avoir un report de « concentré » sur « léger doute ». « Agacement » est lui confondu avec « déception - amusé ». En V, la dispersion est assez différente. La « déception » est bien reconnue mais ces stimuli sont aussi à l'origine d'un report sur « plus sûr - agacé ». « Léger doute » est confondu avec « concentré » et les stimuli « concentré » sont à l'origine des confusions avec « plus sûr - agacé » et « léger doute ». En AV, « déception - amusé » est confondu avec « soulagé » mais il est aussi un peu reconnu. « Concentré » est confondu avec « plus sûr - agacé » alors que c'est « léger doute » qui est confondu avec « concentré ». Enfin « léger doute » se reporte sur « soulagé » et « plus sûr - agacé ».

**Sujet R :** Les profils de dispersion des 3 modalités sont là encore très différents. Pourtant la « concentration » est bien reconnue dans les 3 modalités. En A, elle est aussi

confondu avec « ennui, concentration » et « surprise agacé » et « agacé » se confondent mutuellement ». L'« ennui » est également assez bien reconnu. Sinon « ennui concentration » est confondu avec « ennui incompréhension » qui est lui-même confondu avec « surprise, agacé ». En V, « ennui » et « ennui, incompréhension » sont répulsifs. « Surpris, agacé » attire « ennui incompréhension » et « ennui », « agacé » quant à lui attire « ennui, incompréhension et mon incompréhension est justifié ». Enfin « ennui concentration » est essentiellement confondu avec « agacé », et « mon incompréhension est justifié » attire « surprise, agacé ». En AV, quasiment tous les éléments sont assez bien reconnus sauf « mon incompréhension est justifiée » qui est au centre de quelques confusions ».

### **Synthèse des principales observations**

Concernant les bonnes réponses, la déception présente des stimuli qui sont bien reconnus chez les sujets, et cela dans les trois modalités. De même, la concentration et le doute sont également bien reconnus. Ensuite nous avons des spécificités liées à chaque sujet. Chez le sujet T ce sont les stimuli d'« angoissée » qui sont bien reconnus dans les trois modalités. Le sujet S présente une bonne reconnaissance de l'amusement en AV. Le stress quant à lui est bien reconnu chez le sujet M, en A et en V. Pour les locuteurs hommes, chez le sujet J, l'inquiétude est reconnue en AV et en V de même que l'irritation l'est en AV et en A. Le soulagement du sujet N est reconnu en V de même que l'ennui (plus ou moins la concentration) en A et V chez le sujet R. Donc certains stimuli sont bien porteurs d'indices, mais ces indices ne sont pas forcément additifs dans les différentes modalités.

Pour ce qui est des reports les plus saillants, nous en avons quelques uns qui systématiques, ce qui peut faire l'objet d'une éventuelle classification de stimuli. En effet pour les trois modalités, nous avons des reports symétriques pour 3 couples d'étiquettes : « étonnée »/« angoissée », « assez calme »/« hésitation », « pas contente »/« sûr de moi ». D'autres étiquettes sont également dans le même cas mais avec des reports moins marqués pour la modalité auditive. Il s'agit des couples « surprise »/« agacé », « irritation »/« On se paye ma tête ». De même nous avons des reports unidirectionnels, toujours moins marqués en acoustique mais que nous pouvons également noter tels que : la déception qui est reconnue comme de l'amusement ou la concentration qui est confondue par du doute.

### Remarques sur les faibles valeurs

Les matrices des sujets japonais montrent de manière générale que bien qu'il y ait des tendances remarquables sur certaines valeurs, celles correspondant aux bonnes réponses et qui sont donc sur la diagonale ne sont pas très élevées, tout comme cela a été le cas chez les Français. De plus, il existe beaucoup de dispersions dans les différents éléments de ces matrices, avec notamment des valeurs moyennes entre 15% et 20% qui insèrent du bruit et ne rendent pas les analyses statistiques efficaces.

La Figure 11 ci-dessous est un diagramme qui illustre sur l'exemple du sujet T, la dispersion et la tendance diffuse des réponses japonaises en reprenant les valeurs au-dessous du seuil de 1,5 fois le seuil du hasard.

Cela souligne surtout la nécessité d'améliorer le protocole expérimental afin d'avoir des résultats plus nets qui ne cacheraient pas les tendances émergentes qui semblent être du détail mais pourraient avoir une importance capitale dans la compréhension des tendances interculturelles.

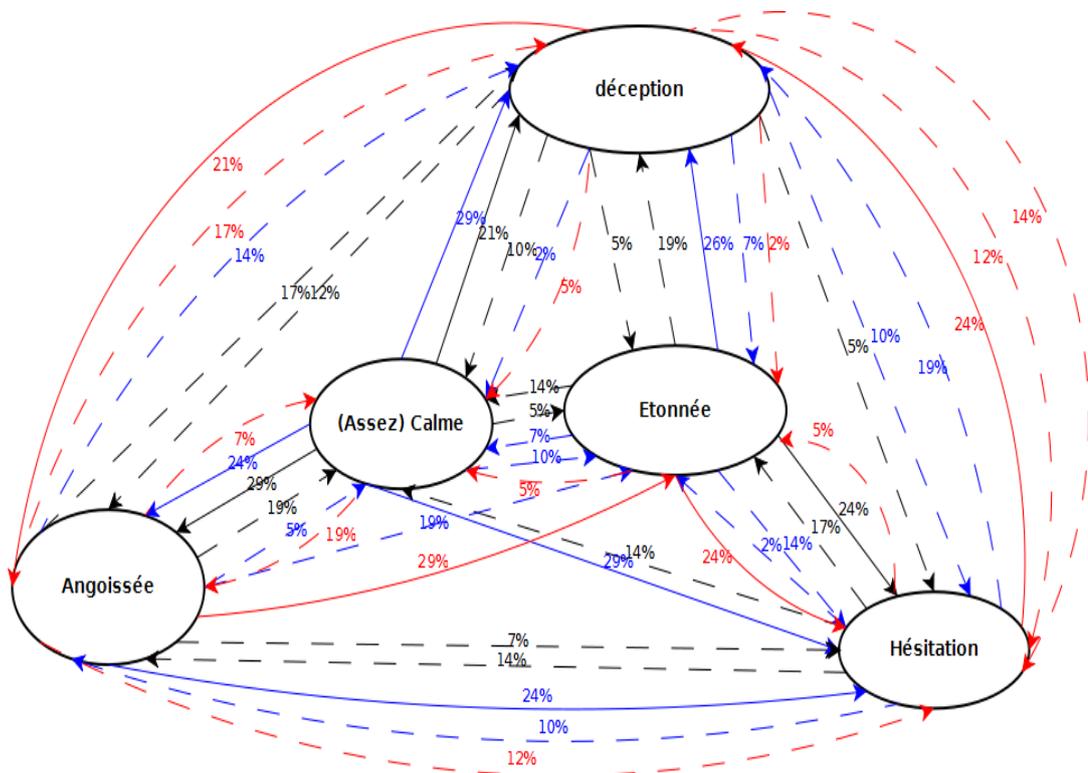


Figure 11 : Diagramme des réponses de faibles scores chez les japonais : exemple du sujet T

### ***Les degrés de confiance des réponses des sujets***

Le degré de confiance des sujets est un autre moyen de comprendre les données du test de perception. En effet, si le degré de confiance, défini par le sujet, a une forte valeur et que ce dernier répond correctement, nous pouvons penser que les stimuli concernés sont fortement informatifs en interculturel. Ainsi le choix des auditeurs serait motivé par certains indices portés par les stimuli. Si la réponse est fautive, cela pourrait être un signe d'élément intra-culturel à valeur informative spécifique. Dans l'idéal, nous devrions regarder ce degré de confiance stimulus par stimulus. Pour pallier ces analyses au cas par cas, nous avons d'une part utilisé une échelle de couleurs pour visualiser globalement l'ensemble de la répartition des degrés de confiance. D'autre part, nous avons calculé des moyennes et des écarts-types pour nous permettre d'avoir une vision d'ensemble de ces degrés de confiance corrélés aux scores de bonnes réponses afin de voir si des tendances remarquables pouvaient s'en dégager. Une fois encore nous n'avons pas pris en compte les sujets chinois ayant participé à notre expérience pour ne brüiter la dimension culturelle des réponses.

#### **Représentation colorimétrique des degrés de confiance**

Les degrés de confiance sont tous représentés dans trois tableaux pour chacune des modalités (cf. Annexe 11 pour l'illustration colorimétrique, l'Annexe CD pour accéder aux valeurs précises du fichier Excel). Nous avons volontairement séparé les hommes des femmes afin de vérifier qu'il n'y ait pas de différences apparentes sur ce critère. Ici nous présentons un extrait de ces données dans le Tableau 6 ci-dessous, qui récapitule uniquement quelques résultats des réponses japonaises chez les femmes, dans la modalité auditive.

Chaque colonne représente un sujet interrogé dans le test de perception et chaque ligne illustre le stimulus correspondant à chaque réponse. Nous avons attribué trois couleurs qui illustrent le niveau du degré de confiance, ce choix de regroupement étant totalement arbitraire. Ainsi nous avons du bleu pour des niveaux de faible confiance (allant de 0 à 3), de l'orange pour des niveaux intermédiaires (de 4 à 6) et les niveaux élevés de degré (de 7 à 10) sont en rose foncé. A priori G. De Biasi avait relevé que le degré de confiance chez les Français était surtout lié à une variation interpersonnelle et ne permettait pas véritablement de se prononcer sur des regroupements de stimuli sur ce critère.

Auto-annotations	Stimuli	02_F_JP	03_F_JP	04_F_JP	05_F_JP	06_F_JP	08_F_JP	10_F_JP	12_F_JP	22_F_JP	23_F_JP	MOY F	Ecart-type F
Inquiétude sur le but du jeu	JEA01.mov	7	8	1	10	10	3	4	0	0	4	5	4
	JEA02.mov	8	7	0	5	4	4	6	3	0	0	4	3
Léger doute	NIC01.mov	8	9	3	5	2	4	5	1	0	0	4	3
	NIC02.mov	7	7	7	10	6	3	3	1	0	0	4	3
Surprise, nervosité / Surpris, agacé	ROB09.mov	8	8	7	5	10	4	4	2	1	0	5	3
	ROB10.mov	8	6	1	5	8	4	5	1	0	0	4	3
Au pif, une envie de rigoler	SAB09.mov	7	8	9	10	8	3	4	1	1	0	5	4
	SAB10.mov	6	5	2	5	10	6	4	1	0	0	4	3
Hésitation	TIP09.mov	10	8	10	0	10	6	7	4	0	2	6	4
	TIP10.mov	9	7	9	10	10	4	5	0	0	5	6	4
MOYENNE des degrés de confiance par locuteur		8	7	5	7	8	4	5	2	0	1		
Ecart-type des degrés de confiance par locuteur		1	1	4	3	3	1	1	1	0	2		
Scores bonnes réponses (%)		19	10	15	14	14	13	10	16	18	15		

**Tableau 6 : Extrait d'un tableau récapitulatif des degrés de confiance: exemple de la modalité auditive des réponses des femmes japonaises**

Ce qui est sûr, c'est qu'il y a effectivement d'importantes variations interpersonnelles chez les Japonais tout comme chez les Français. Sur un aperçu très général des couleurs (cf. Annexe 11), nous voyons tout d'abord qu'en audio, les femmes sont beaucoup plus confiantes que les hommes, bien qu'il y ait des femmes qui sont très peu confiantes par rapport aux autres. En visuel, les hommes semblent avoir un plus haut degré de confiance, ce qui semble également vrai en audio-visuel. Il nous faut donc faire une analyse plus précise en corrélant ces degrés de confiance avec le score de bonnes réponses de ces participants.

D'autre part, nous voulions souligner le cas de 23\_F\_JP, qui est à l'origine d'une hypothèse possible car elle a un comportement similaire à celui des hommes (cf. Annexe CD, le fichier Excel degreConfiance.xls). En effet, cette femme, vis-à-vis de la culture japonaise, est souvent qualifiée de masculine, dans la manière de se comporter, de parler ou dans le choix de ses loisirs. Donc la personnalité d'une personne pourrait éventuellement illustrer les réponses données dans le test. Sachant également que la langue japonaise est une langue très marquée par le genre, nous pouvons poser comme hypothèse que la reconnaissance des micro-expressions est elle-même influencée par le sexe, tout au moins pour les Japonais, ce que nous essaierons également de vérifier dans l'analyse des corrélations entre les moyennes/écart-types et les scores de bonnes réponses.

### **Corrélation entre degré de confiance et score de bonnes réponses**

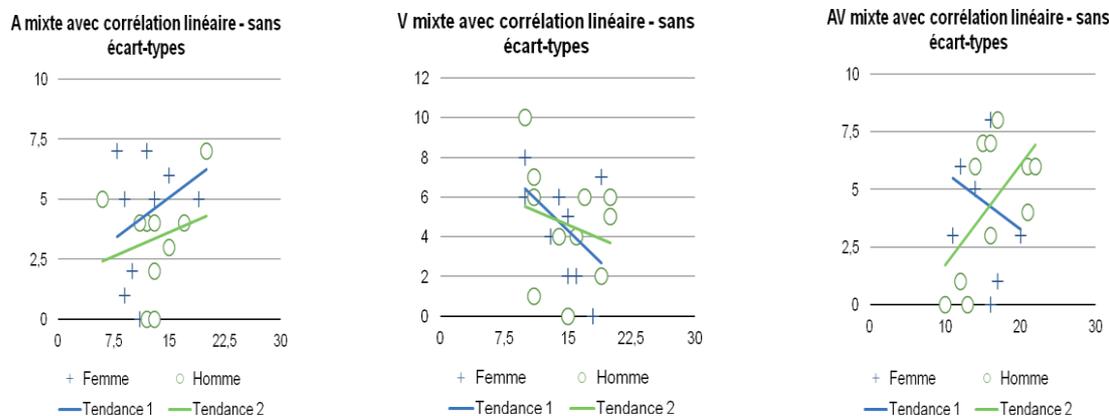
Les écart-types et les moyennes de chaque stimulus sont calculés pour l'ensemble des sujets interrogés, hommes et femmes. C'est ce que nous voyons dans les colonnes grisées. Les moyennes et les écart-types des degrés de confiance sont également calculés pour chaque locuteur et sont représentées dans les lignes grisées en bas de tableau. Enfin nous avons également calculé le score moyen de bonnes réponses de chaque sujet auditeur pour chacune des modalités. Pour avoir une corrélation entre ces moyennes des et ces scores, nous avons réalisé des graphiques pour chacune des modalités (A-V-AV) avec en abscisse le score de bonnes réponses en pourcentage et en abscisse le degré de confiance qui est donnée sur une échelle de 0 à 10. Nous avons, pour la lisibilité, 6 graphiques en Annexe 11, qui différencient les résultats donnés par les hommes et femmes, pour chacune des modalités, avec les écart-types des moyennes. Nous avons également 3 graphiques mixtes pour chaque modalité qui représente les tendances des hommes et des femmes sous forme de corrélations linéaires afin de pouvoir effectuer une comparaison entre les deux sexes (cf. Figure 12).

### **Observations des scores et écarts-types les graphiques de corrélation**

Concernant les graphiques illustrant les écart-types (cf. Annexe 11), les scores entre hommes et femmes sont semblables, voir meilleurs chez les hommes que chez les femmes pour les modalités visuelles et audiovisuelles. Pour l'audio, nous avons des scores qui peuvent être plus bas que ceux du visuel ou de l'audiovisuel. Concernant la modalité auditive, les femmes ont un degré de confiance assez élevé (pour la grande majorité au-dessus de 5), alors que celui-ci est plus bas pour les hommes (au-dessous de 5). Cela rejoint nos observations sur la colorimétrie. Pour le visuel nous avons des figures semblables entre homme et femmes avec des scores rassemblés de part et d'autre de 15%. Cependant nous remarquons que les écart-types des femmes sont beaucoup plus importants que ceux des hommes, donc les réponses des femmes sont beaucoup plus variables que celles des hommes. Donc les indices visibles au sein des stimuli semblent être sujets à confusion chez les femmes alors que les hommes sont assez consensuels et plus confiants que pour la modalité auditive. Enfin en audiovisuel, les écart-types des femmes restent assez importants et les degrés de confiance restent relativement les identiques que ceux de la modalité visuelle. Chez les hommes nous avons à nouveau des degrés de confiance plus élevés et une augmentation des scores de bonnes réponses en audiovisuel.

## La différence des tendances de corrélation entre hommes et femmes

Sur les matrices de corrélation linéaire, présentées ci-dessous dans la Figure 12, ces différentes tendances sont bien mises en évidence.



**Figure 12 : Tendances de corrélation entre scores de bonnes réponses et degré de confiance**

En audio, les femmes proposent un bien meilleur degré de confiance que les hommes. Mais aussi bien chez les hommes que chez les femmes, plus leurs degrés de confiance sont élevés, meilleurs sont les scores.

En visuel nous observons la tendance inverse. En effet, plus les résultats sont meilleurs chez les sujets, plus leurs degrés de confiance diminuent. Donc leurs choix semblent moins motivés et moins réflexifs, ce que nous pouvons considérer comme l'entrée dans une boucle involontaire pour le traitement de ces données dans la modalité visuelle. Mais la perception non réflexive donne tout de même de bons résultats. Cette tendance est d'ailleurs plus accentuée chez les femmes car leurs degrés de confiance qui étaient plus importants que ceux des hommes deviennent à l'inverse plus faibles. Ce changement a lieu aux alentours des scores de bonnes réponses de 14-15%.

Enfin les tendances des hommes et des femmes sont complètement inversées pour la modalité audiovisuelle. Chez les hommes la tendance reste la même (voire plus marquée) que celle observée en audio, avec une augmentation des scores de bonnes réponses corrélée à de meilleurs degrés de confiance. Pour les femmes, celles-ci gardent la même tendance que pour la modalité visuelle, à savoir une diminution des degrés de confiance avec l'augmentation des scores. Donc cela signifierait qu'en audiovisuel, les hommes restent sur des choix plus volontaires et réflexifs que les femmes. Les femmes, elles, resteraient plus intuitives quand le visuel intervient dans les modalités de perception.

### *La perception culturelle des micro-expressions*

A partir des réponses sur la perception japonaise du caractère natif et non natif des stimuli, nous avons réalisé des regroupements pour la modalité auditive. Il s'agit d'une première classification des micro-expressions sur un critère perceptif « ethnologique ». Le tableau des réponses correspondant à cette perception se trouve en Annexe 13. A partir de la colonne « non japonais » de ce tableau, nous obtenons une échelle graduée en pourcentage illustrant les réponses des japonais quant à leur impression sur l'origine des stimuli. A 0% nous sommes à l'extrême « perception de stimuli natif » et à 100% nous sommes sur des stimuli perçus comme « non japonais » (cf. Figure 13).

+ japonais							- japonais								
0	10	20	30	40	50	60	70	80	90	100					
			38%	43%	48%	52%	57%	62%	67%	71%	76%	81%	86%	90%	95%
			JEA0	TIPO	SAB	SAB	SAB	TIPO	TIPO	TIPO	SAB	MAR	SAB	TIPO	TIPO
			3	2	01	06	07	5	1	4	03	06	04	3	6
				SAB	SAB	SAB	ROB	TIPO	TIPO	MAR	NICO	MAR	SAB		NICO
				11	02	09	08	7	8	10	7	07	12		2
					JEA0	MAR	ROB	SAB	TIPO	MAR	NIC1	NICO	NICO		
					4	04	05	05	9	12	0	4	6		
					JEA0	JEA0		SAB	TIP1	MAR					
					5	1		08	0	03					
					JEA0	JEA0		MAR	SAB	MAR					
					6	7		02	10	08					
					JEA0	ROB		ROB	MAR	MAR					
					8	01		02	05	09					
					NICO	ROB		ROB	MAR	JEA0					
					8	03		07	11	2					
					NICO			ROB	MAR	NICO					
					9			14	01	1					
					ROB			ROB	NICO	NICO					
					04			06	5	3					
									ROB						
									10						
									ROB						
									09						
									ROB						
									11						
									ROB						
									12						
									ROB						
									13						

Figure 13 : Echelle de perception des stimuli les plus japonais aux moins japonais en modalité

A

D'après cette échelle, nous pouvons voir que la grande majorité des stimuli sont perçus comme non natifs (55 stimuli sont définis à plus de 52% comme des stimuli non

natifs). Parmi ceux-là, un nombre important de micro-expressions s'accroissent autour d'une perception de 62% à 71% « non japonais ». Puis à partir de 76% le nombre de stimuli diminue. Cependant il existe des stimuli qui sont considérés comme des stimuli non natifs jusqu'à 95%. Cela forme donc un second groupe de stimuli avec des caractéristiques très culturelles. Enfin nous avons des stimuli perçus de 38% à 57% des « non japonais ». Ces objets sont donc assez ambigus et ont certainement des caractéristiques acoustiques relativement universelles. En effet, si nous avons des productions très proches de sons japonais nous aurions eu des résultats à des taux bien plus faibles, or le fait qu'il y ait un nombre assez important de stimuli dans la zone du hasard signifie que certains stimuli sont plus universels que d'autres. D'ailleurs, tous les stimuli de JEA soit le sujet J se trouvent dans la partie « + japonais » de l'échelle. Donc nous pouvons supposer que les stimuli de ce locuteur sont relativement universels par rapport à celui des autres locuteurs.

Nous n'irons pas plus loin dans ce mémoire concernant cet aspect de la perception, cependant de nombreuses perspectives sont possibles pour étudier plus en détail cette classification de stimuli, notamment en analysant les caractéristiques acoustiques de ces groupes. Cependant nous pouvons d'ores et déjà voir que ces micro-expressions sonores sont assez fortement marquées culturellement, ce qui permet de renforcer notre hypothèse selon laquelle les micro-expressions seraient des objets langagiers.

### ***Conclusion du chapitre 5***

Nous avons vu dans ce chapitre qu'il a fallu de nombreux tests statistiques pour mettre en évidence les réponses pertinentes car nous avons beaucoup de bruits dans nos matrices. En fonction des sujets, les caractéristiques de dispersion sont complètement différentes mais nous notons tout de même que les reports étaient surtout réalisés entre des étiquettes et des stimuli sémantiquement assez proches. Il existe également des reconnaissances de stimuli et des reports qui apparaissent de façon systématique.

Par ailleurs, les degrés de confiance ont montré des tendances qui nous laissent à penser qu'il existerait peut-être une voie d'observation sur les différences entre hommes et femmes (pour les Japonais), dans l'hypothèse que, tout comme la langue japonaise, la perception est elle aussi sensible au genre des sujets observés.

Enfin, la perception du caractère japonais ou étranger des objets acoustiques montre que les Japonais perçoivent des indices culturels dans les micro-expressions mais qu'il

existe aussi une gradation de la discrimination culturelle des stimuli qui est plus ou moins évidente en fonction des objets observés. Ce critère peut donc être intéressant pour effectuer une classification des stimuli si nous le couplons à une analyse psychoacoustique des données, ce qui est une perspective directe de cette étude.

## Chapitre 6 – Analyse comparative franco-japonaise

Nous avons précédemment vu l'analyse des données pour les sujets japonais. Dans le chapitre 6, nous allons nous intéresser à la comparaison entre les réponses des Français et celles des Japonais. Notons que cette analyse est très succincte et ne s'appuie que sur les tendances les plus importantes (prises en compte des pourcentages deux fois au dessus du seuil du hasard) pour avoir un premier aperçu général des similitudes et des différences. En revanche, les petites variations ont autant d'importance que les grandes.

### *Analyse comparative sur les taux de bonnes réponses générales*

Quand nous regardons les tableaux en Annexe 10 illustrant le taux de bonnes réponses françaises et japonaises, nous nous apercevons que les Japonais ne sont pas toujours plus mauvais que les Français dans la reconnaissance des stimuli (comme nous l'avions pensé), même si d'une manière générale, le taux de reconnaissance global des Français est meilleur que celui des Japonais.

En effet, pour la modalité auditive les stimuli des sujets T et N sont mieux reconnus par les Japonais que par les Français. De même, nous faisons la même observation sur les stimuli des sujets M et J en visuel. Donc il existe très certainement des indices subtils et une analyse approfondie des caractéristiques acoustiques et visuelles devra être faite permettrait de donner des pistes d'investigation en faveur d'une classification de ces stimuli. Cette observation est encore plus spécifique en audiovisuel, car les Japonais reconnaissent mieux les stimuli des hommes que ceux des femmes comme nous l'illustre le Tableau 7. En termes de perspectives, nous pouvons alors nous demander si le sexe des locuteurs n'aurait pas une influence sur la perception, en tout cas des Japonais.

Sujets	Nb bonnes réponses des Japonais	Taux de bonnes des Japonais	Nb bonnes réponses des Français	Taux de bonnes des Français
T	44	20,95%	40	26,67%
S	53	21,03%	50	27,78%
S	47	18,65%	42	23,33%
J	58	34,52%	33	27,50%
N	47	22,38%	31	20,67%
R	79	26,87%	53	25,24%
<b>Total</b>	<b>328</b>	<b>23,67%</b>	<b>249</b>	<b>25,20%</b>

Tableau 7 : Scores de bonnes réponses par sujets chez les Japonais et les Français pour la modalité audiovisuelle

## Explication de l'analyse comparative étiquette par étiquette

La comparaison s'appuiera sur les tendances les plus importantes, chez les auditeurs français et chez les auditeurs japonais. L'analyse se fera étiquette par étiquette pour chacun des sujets en utilisant une nomenclature spécifique.

### Méthodologie

Pour la comparaison, nous ne prendrons en compte que les valeurs qui sont supérieures à deux fois le seuil du hasard puisqu'il s'agit de la limite de significativité qui a été choisie pour l'étude des réponses françaises (De Biasi, 2011). Ces réponses, nous les avons résumées sous forme de diagrammes représentant soit les réponses des sujets français, soit les réponses des sujets japonais. Toutes les figures correspondantes se trouvent en Annexe 14.

Voici un exemple de ce diagramme ci-dessous en Figure 14. Les reports et les reconnaissances des stimuli sont illustrées sous forme de flèches. En noir, nous avons les résultats de la modalité auditive seule, en bleu ceux de la modalité visuelle seule, et enfin en rouge les réponses de la modalité audiovisuelle. Sur les diagrammes des Japonais les taux compris entre 1,5 fois et 2 fois le hasard sont également affichés (flèches en pointillé). Le sens des flèches indiquent le sens des reports.

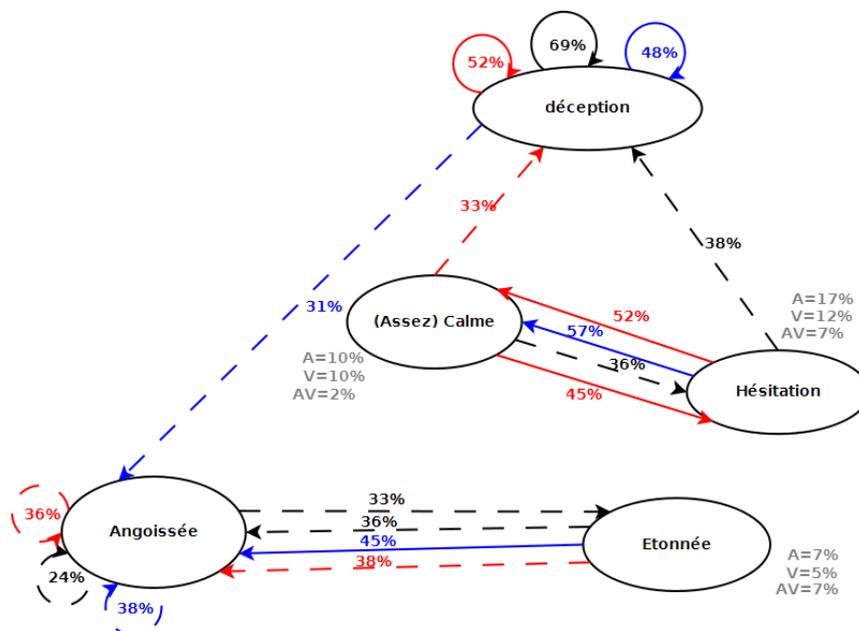


Figure 14 : Diagramme de reports et de reconnaissances élevés : exemple des réponses japonaises sur les stimuli du sujet T

## Nomenclature d'analyse

Pour chaque étiquette nous utiliserons un formalisme de description pour résumer les relations observées entre les étiquettes dans les différentes modalités puis nous commenterons brièvement nos observations sur chaque sujet.

Pour chaque étiquette, les nationalités des auditeurs sont marquées par FR (française) ou JP (japonaise). Puis pour chaque type d'auditeur nous donnons les modalités pour lesquels il y a eu un bon taux de reconnaissance. Cela est noté : Reco= A (pour l'audio), V (pour le visuel), AV (pour l'audiovisuel) en enfin 0 si aucune modalité n'est reconnue. Mais dans ce dernier cas, nous indiquons entre parenthèse l'ordre d'influence des modalités. Ces modalités sont donc rangées par ordre croissant de taux de reconnaissance et nous utiliserons des connecteurs de comparaison (<, > et  $\approx$  quand les valeurs sont similaires). Ces connecteurs logiques sont également utilisés dans le cas des reports. Pour les reports, nous utilisons des flèches => illustrant le choix d'étiquettes des auditeurs pour les stimuli correspondants à l'auto-annotation donnée les locuteurs. Cette notation est accompagnée des modalités concernées par la confusion<sup>20</sup>.

### *Observations sur les stimuli du sujet T*

#### **La déception**

FR /// Reco = AV<V  
Déception => hésitation (AV<A)

JP /// Reco = V<AV<A  
Déception => stress (V)

Ici le « stress » désigne l'étiquette « angoissée ».

#### **Hésitation**

FR /// Reco = 0 (AV<A<V)  
Hésitation => stress (AV)

JP /// Reco = 0 (AV<V<A)  
Hésitation => déception (A) + calme (AV>A)

---

<sup>20</sup> Par exemple, *Déception* => *hésitation* (AV<A) signifie que les stimuli de la « déception » sont reconnus comme de « l'hésitation », surtout en AV mais également en A.

### **(Assez) Calme**

FR /// Reco = 0 ( $A \approx AV < V$ )  
Calme => hésitation ( $V < A \approx AV$ )

JP /// Reco = 0 ( $AV < A \approx V$ )  
Calme => déception (AV) + hésitation ( $A < AV$ )

### **Etonnée**

FR /// Reco = 0 ( $A < AV < V$ )  
Surprise => calme (A)

JP /// Reco = 0 ( $V < A \approx AV$ )  
Surprise => stress ( $A < AV < V$ )

### **Angoissée**

FR /// Reco =  $V < AV$   
Stress => surprise (A)

JP /// Reco =  $A < V < AV$   
Stress => surprise (A)

### **Synthèse des observations du sujet T**

Les stimuli sont peu reconnus, sauf « déception » et « angoissée » qui marchent bien pour les deux nationalités. Dans ces cas-là, les modalités de reconnaissance sont souvent inversées entre Français et Japonais (les sujets français reconnaissent en général plus en V qu'en A contrairement aux sujets japonais). Les reports ne sont pas sur les mêmes éléments, ni sur la même modalité sauf pour le « stress » qui est dans tous les cas confondu avec la « surprise ».

### ***Observations sur les stimuli du sujet S***

#### **Mission Impossible**

FR /// Reco = 0  
Doute => déception (A)

JP /// Reco = 0  
Doute => amusement (A) + concentration (AV)

L'étiquette « mission impossible » est représentative du « doute ».

### **Le but est de trouver des solutions**

FR /// Reco = AV ( $V < A$ )  
FoT => concentration (A)

JP /// Reco = 0 ( $A < V$ )  
FoT => doute (A) + concentration ( $V \approx AV$ )

L'étiquette « le but est de trouver des solutions » est représentative de la notion du « Feeling of Thinking » (ou FoT<sup>21</sup>) décrit dans l'état de l'art de ce mémoire.

### **Surprise par les résultats**

FR /// Reco = 0 ( $V < A \approx AV$ )  
Surprise => déception (A)

JP /// Reco = 0 ( $A < V < AV$ )  
Surprise => doute ( $A \approx V$ ) + concentration (A)

### **Au pif, une envie de rigoler**

FR /// Reco =  $V \approx AV$   
Amusement => surprise ( $V \approx AV$ )

JP /// Reco = AV ( $V < A$ )  
Amusement => déception ( $V \approx AV$ )

L'étiquette « au pif, une envie de rigoler » est représentative de l'« amusement ».

### **Toujours la concentration**

FR /// Reco = 0 ( $V \approx AV < A$ )  
Concentration => déception ( $V < AV$ )

JP /// Reco =  $V < AV$   
Concentration => amusement ( $V < A < AV$ )

### **Déçue par les résultats**

FR /// Reco = 0 ( $V < A \approx AV$ )  
Déception => 0

---

<sup>21</sup> cf. Chapitre 3 § *Les motivations d'études sur les micro-expressions et terminologie*, Section : *Feeling of Thinking* pour retrouver la définition du concept

JP /// Reco = 0 (A<AV<V)  
Déception => doute (A) + amusement (A) + concentration (AV)

### **Synthèse des observations du sujet S**

Aussi bien pour les Japonais que pour les Français, le « doute », la « surprise » et la « déception » ne sont pas reconnus. Le « FoT » et la « concentration » ne sont reconnus que partiellement, surtout en AV, le premier par les Français, le second par les Japonais. Enfin « l'amusement » est reconnu par tous en AV.

Pour les reports, ceux-ci peuvent être différents entre Français et Japonais comme c'est le cas pour le « doute », la « surprise », l'« amusement » ou la « concentration ». Ces reports peuvent également être similaires mais sur des modalités différentes comme pour le « FoT ». Enfin parfois il y a des cas très ambigus comme la « déception » qui ne présente aucun report pour les Français et beaucoup de confusions pour les Japonais.

### **Observations sur les stimuli du sujet M**

#### **Peur**

FR /// Reco = 0 (V<AV)  
Peur => doute (V<AV<A)

JP /// Reco = 0 (V=AV)  
Peur => doute (A) + DecIrr (AV<V)

#### **Doute**

FR /// Reco = 0 (A<AV)  
Doute => 0

JP /// Reco = 0 (A<AV)  
Doute => DecIrr (V<A)

Les auto-annotations désignées par « DecIrr » représentent un amalgame entre la « déception » et l'« irritation ».

#### **Pas contente-déçue**

FR /// Reco = A (V<AV)  
DecIrr => stress (AV)

JP /// Reco = V (A<AV)

DecIrr => contentement ( $V < AV$ )

### **Sûr de moi / Bien à l'aise**

FR /// Reco = 0 ( $A \approx V \approx AV$ )  
Contentement => doute ( $A \approx AV$ )

JP /// Reco = 0 ( $A < V < AV$ )  
Contentement => DecIrr ( $AV < A \approx V$ ) + doute ( $AV < A$ ) + stress (V)

L'étiquette « sûre de moi / bien à l'aise » est représentative du « contentement ».

### **Stress**

FR /// Reco = 0 ( $V < AV < A$ )  
Stress => doute ( $AV < V < A$ ) + DecIrr (V)

JP /// Reco =  $A < V$   
Stress => doute ( $AV < V$ ) + déception (AV)

### **Déception**

FR /// Reco = 0 ( $AV < A \approx V$ )  
Déception => doute (AV)

JP /// Reco = AV ( $V < A$ )  
Déception => DecIrr ( $V \approx AV$ ) + contentement (V)

### **Synthèse des observations du sujet M**

Le « doute », la « peur » et le « contentement » sont très peu reconnus quelque soit la population observée. Le « stress » et la « déception » ne sont reconnus que par les Japonais, en A mais surtout en V pour l'un, en AV pour le second. Enfin l'amalgame « déception + irritation » est reconnu en A par les Français, en V par les Japonais. Le « stress », le « contentement » et la « peur » présentent des reports similaires pour les deux populations mais sont très différentes en étiquettes et en modalités pour les autres stimuli.

### ***Observations sur les stimuli du sujet J***

#### **Petite lassitude / l'indifférence s'installe**

FR /// Reco = 0 ( $V < AV < A$ )  
Ennui => 0

JP /// Reco = AV (A<V)  
Ennui => irritation (AV)  
L'étiquette « petite lassitude / l'indifférence s'installe » est représentative de l'  
« ennui ».

#### **Petite irritation**

FR /// Reco = V (AV<A)  
Irritation => stress (AV)

JP /// Reco = A <AV  
Irritation => 0

#### **Inquiétude sur le but du jeu**

FR /// Reco = V (A<AV)  
Stress => 0

JP /// Reco = V < AV  
Stress => ennui (A)

L'étiquette « inquiétude sur le but du jeu » est représentative du « stress ».

#### **Je commence à comprendre qu'on se paye ma tête**

FR /// Reco = 0 (A ≈ AV < V)  
FoT => irritation (AV)

JP /// Reco = 0 (A<AV<V)  
FoT => irritation (A ≈ AV)

L'étiquette « je commence à comprendre qu'on se paye ma tête » est représentative  
du « FoT ».

#### **Synthèse des observations du sujet J**

Le « FoT » n'est reconnu par personne alors que le « stress » est reconnu par les  
deux populations en V. En revanche, les reconnaissances de l'« irritation » et l'« ennui »  
sont asymétriques. Soit ces derniers sont reconnus pour des modalités différentes (A et V),  
soit l'un est reconnu par les Japonais alors qu'il ne l'est pour les Français. Par ailleurs les  
reports sont complètement différents chez les différents sujets sauf pour le « FoT » qui  
renvoie vers une « irritation », surtout pour ce qui concerne la modalité AV.

## *Observations sur les stimuli du sujet N*

### **Déception mais m'en m'amuse**

FR /// Reco = AV<V  
DecAmu => DouIrr (A)

JP /// Reco = V<AV  
DecAmu => soulagement (AV)

« DecAmu » qui est le diminutif choisi pour l'étiquette « déception mais m'en amuse », représente un amalgame de « déception » et d'« amusement ».

### **Léger doute**

FR /// Reco = 0 (AV<A<V)  
Doute => concentration (A<V<AV)

JP /// Reco = V (A ≈ AV)  
Doute => concentration (V<AV) + DouIrr (A)

L'étiquette « DouIrr » représente un amalgame entre le « doute » et l'« irritation ».

### **Soulagé**

FR /// Reco = 0 (A<V ≈ AV)  
Soulagement => doute (A)

JP /// Reco = V (A<AV)  
Soulagement => doute (AV<A)

### **Essaie de rester concentré**

FR /// Reco = V (A ≈ AV)  
Concentration => doute (A) + DouIrr (AV)

JP /// Reco = 0 (A<AV<V)  
Concentration => DouIrr (A<V<AV)

### **Ne suis plus sûr de moi, léger agacement**

FR /// Reco = 0 (AV<V<A)

DouIrr => doute (AV<A)

JP /// Reco = A (AV<V)  
DouIrr => doute (A<AV)

« DouIrr » qui est le diminutif choisi pour l'étiquette « ne suis plus sûr de moi, léger agacement », représente un amalgame de « doute » et d'« irritation ».

### **Synthèse des observations du sujet N**

Les reports sur les stimuli du sujet sont relativement proches entre Français et Japonais, sauf pour l'amalgame entre la « déception » et l'« amusement ». En termes de reconnaissance, les Japonais sont généralement meilleurs que les Français et ces derniers mettent surtout en jeu la modalité visuelle. Cependant l'amalgame entre la « déception » et l'« amusement » est bien reconnu par les deux nationalités, et cela particulièrement pour la modalité audiovisuelle.

### ***Observations sur les stimuli du sujet R***

#### **Concentration**

FR /// Reco = A<AV<V  
Concentration => ConEnn (AV<A) + EnnDou (AV)

JP /// Reco = A<AV<V  
Concentration => ConEnn (AV)

« ConEnn » désigne un amalgame entre « concentration » et « ennui ».

#### **Ennui**

FR /// Reco = 0 (V ≈ AV<A)  
Ennui => irritation (V)

JP /// Reco = 0 (AV<A)  
Ennui => concentration (V) + FoT (A<AV)

#### **Ennui-concentration**

FR /// Reco = 0 (A<AV<V)  
ConEnn => concentration (AV<A) + EnnDou (AV)

JP ///  $\text{Reco} = AV (A=V)$   
 $\text{ConEnn} \Rightarrow \text{EnnDou} (V<AV) + \text{FoT} (A)$

« EnnDou » désigne un amalgame entre « ennui » et « doute ».

### **Ennui, incompréhension**

FR ///  $\text{Reco} = V - AV$   
 $\text{EnnDou} \Rightarrow 0$

JP ///  $\text{Reco} = AV (A \approx V)$   
 $\text{EnnDou} \Rightarrow \text{FoT} (A) + \text{ConEnn} (AV)$

### **Mon incompréhension est justifiée**

FR ///  $\text{Reco} = 0 (A<V<AV)$   
 $\text{FoT} \Rightarrow \text{EnnDou} (V) + \text{SurIrr} (AV)$

JP ///  $\text{Reco} = 0 (V<A \approx AV)$   
 $\text{FoT} \Rightarrow \text{EnnDou} (A \approx AV<V)$

« SurIrr » désigne un amalgame entre « surprise » et « irritation ». Par ailleurs l'étiquette « mon incompréhension est justifiée » désigne la notion de « FoT ».

### **Agacé**

FR ///  $\text{Reco} = AV<A$   
 $\text{Irritation} \Rightarrow 0$

JP ///  $\text{Reco} = 0 (V<A<AV)$   
 $\text{Irritation} \Rightarrow \text{FoT} (AV<A \approx V) + \text{EnnDou} (AV) + \text{ConEnn} (V)$

### **Surprise, nervosité / Surprise, agacé**

FR ///  $\text{Reco} = A (V \approx AV)$   
 $\text{SurIrr} \Rightarrow \text{FoT} (V<AV) + \text{irritation} (A)$

JP ///  $\text{Reco} = AV (V<A)$   
 $\text{SurIrr} \Rightarrow \text{FoT} (AV<A)$

### **Synthèse des observations du sujet R**

Le profil de reconnaissance pour les stimuli de la « concentration » est semblable pour les deux populations. Par ailleurs, l'« ennui » et le « FoT » ne sont pas reconnus. Sinon nous voyons aussi bien des dissymétries et des similitudes pour les autres reconnaissances, sans que les tendances soient identiques entre sujets japonais et français.

Pour les reports, nous trouvons des similitudes pour les stimuli visuels de « FoT » confondus par l'amalgame « ennui + doute » pour les deux nationalités, de même que les stimuli de « surprise+irritation » qui sont confondus par du « FoT » en audiovisuels. Sinon pour tous les autres reports, nous avons de grandes variations aussi bien dans le choix des étiquettes et des modalités.

## Chapitre 7 – Les observations de synergologie

La principale expérience de ce mémoire ainsi que ses résultats, ont été présentés dans les trois chapitres précédents. C'est cette expérience qui sous-tend toutes les observations que nous envisageons. Il faut rappeler que même si nous adoptons une méthodologie expérimentale scientifiquement vérifiable pour traiter ces objets que sont les micro-expressions, nous nous basons sur des auto-annotations empiriques, qui sont volontairement non reliées à des hypothèses théoriques a priori. Ce qui signifie que l'objectivation globale de nos observations ne pourra que découler à terme d'une boucle expérimentale qui sera à même de relier, par ce processus méthodologique quasi inductif (comme on l'applique en éthologie ou en ethno-méthodologie (F Loyau, 2007) ou A. (Vanpé, 2012)), nos observations à des constructions théoriques.

Dans les études portant sur les émotions, nous avons pu voir que de nombreuses recherches portaient sur des indices visuels d'expressions faciales. De nouvelles disciplines s'intéressent également sur ces aspects visuels mais plus globalement sur le langage corporel. C'est le cas d'un mouvement appelé la synergologie qui a attiré notre attention pour ses particularités que nous présenterons dans ce chapitre 7.

### *Aperçu de la synergologie*

#### **Contexte et intérêts pour la synergologie**

Aux Etats-Unis, une série s'appelant « Lie To Me » a mis au grand jour une discipline de psychologie s'appuyant sur l'observation d'expressions faciales, directement inspirées du modèle du FACS (Facial Coding System Investigation) qu'a développé P. (Ekman, 1975), scientifique très renommé, dont tous les travaux, avant ces dernières années, entraînent largement dans les processus méthodologiques de falsification/vérification propre à toute démarche scientifique. Ainsi le modèle d'Ekman développé avec un souci de rigueur scientifique est souvent pris comme référence dans les études de l'analyse expressive faciale. En revanche, ce modèle pose un problème car bien qu'il soit actuellement à la base d'une grande majorité d'études, il ne fait désormais plus l'objet d'une remise en cause qui passe par une confrontation aux évaluations qui s'appuient sur des protocoles scientifiques.

D'autre part, il y a apparition d'un grand nombre de méthodes qui s'intéressent au décryptage du langage mais qui malheureusement ne s'appuient sur aucune base

scientifique, car il existe un véritable engouement autour des séries, des émissions ou des thématiques sur les « mentalistes », la détection de mensonge, « les profilers », « le coaching » ou encore des réflexions qui tendent à savoir ce que l'autrui ressent et pense ou plus globalement sur sa personnalité.

Dans ce contexte, un mouvement récent fait son apparition en France sous le nom de « synergologie » (marque déposée par Philippe Turchet, son fondateur) en étudiant le langage corporel des micro-gestes. Cette synergologie est basée sur la proposition d'une grammaire universelle de micro-gestes, qui a l'avantage de vouloir se confronter aux méthodes scientifiques afin de confirmer ses méthodes par une validation objective de ses observations. Nous nous proposons dans notre travail de faire un premier pas dans cette observation de la synergologie comme hypothèses scientifiques. Pour cela nous donnons à des synergologues les mêmes stimuli qu'aux sujets naïfs (rappelons qu'une précédente expérience a mesuré les performances perceptives de Français naïfs), et nous noterons comment et sur quels critères certains de nos stimuli sont identifiables en synergologie.

### **Présentation générale de la discipline<sup>22</sup>**

Philippe Turchet est considéré comme le père de la synergologie dès lors qu'il entamait des études doctorales sur la rationalité du comportement dans les années 1980. L'ouvrage le plus abouti à l'heure actuelle serait « Le langage universel du corps » dont il en est l'auteur. L'approche est officiellement fondée en 1996.

Etymologiquement, la synergologie est composée de « sun » (être ensemble), « ergo » (actif) et « logos » (discours) qui signifierait donc « être actifs en situation de production de discours ». La synergologie est une approche qui fait des classifications d'informations non verbales gestuelles. Elle part du pré-requis qu'il existe une universalité du langage corporel au-delà des particularités culturelles qu'elle n'estimerait qu'aux environs de 1% parmi toute la gestualité corporelle de l'homme.

Pour elle, le corps et l'esprit ne forme qu'une seule et même entité et s'intéresse au mécanisme cognitif que l'on retrouve dans l'interaction humaine. Pour cela elle s'appuie sur une nomenclature spécifique en découpant l'information non verbale qu'elle corrèle avec des grilles de catégories émotionnelles qu'elle aurait elle-même établie. Elle fait par

---

<sup>22</sup> Cette discipline étant assez récente donc très peu de manuels ou de références scientifiques existent pour la décrire. En France, un site internet sert de référence à tous les synergologues et présente la nomenclature ainsi que la discipline : <http://www.synergologie.org/> . La présentation qui en est faite s'en inspire très largement.

ailleurs la distinction entre ce qu'elle appelle les « émotions sous-jacentes » et « sur-jacentes » pour s'intéresser aux phénomènes du « non dit ». Enfin cette démarche affirme valider ses connaissances en utilisant le principe de réfutabilité (affirmer la réalité d'un fait par l'impossibilité d'en trouver des contre-exemples, soit un raisonnement par l'absurde) en se confrontant à des protocoles scientifiques. Ce critère de réfutabilité (Karl Popper, 1972) est décrit grâce au principe d'induction. Ainsi elle s'appuierait sur des données neurobiologiques pour maintenir la cohérence de ses observations. Elle ne recherche pas à valider mais à réfuter ses observations pour pouvoir les confirmer. Cette volonté de rigueur scientifique est également à l'origine de la mise en place d'un code éthique en 25 articles.

Enfin, la synergologie est une approche qui s'apprend à travers une formation de 300 heures soit environ trois ans pour maîtriser l'observation d'éléments pertinents de la scène visuelle selon une nomenclature précise.

### **Les méthodes d'observations de la synergologie**

La synergologie se déclare scientifique pour trois de ces aspects : le découpage des informations qu'elle observe, les propositions qui sont émises ainsi que le mode de validation des propositions.

Elle utilise donc un corpus de propositions (recueillies auprès de la communauté des synergologues), sur lequel elle s'appuie pour sa propre description. A l'heure actuelle, la discipline semble en accord avec 344 propositions et celles-ci sont regroupées dans 6 classes appelées respectivement : « Conditions d'apparition de la synergologie », « Principes de classification de la connaissance », « Détermination d'un champ heuristique », « Le paradigme synergologique », « Des stratégies universelles de communication » et « Théorisation de la relation ». Ces propositions permettent aux synergologues de préciser leurs méthodologies, leurs terminologies et les approches sur lesquelles ils s'appuient entre autres.

Enfin concernant les méthodes d'observations, la synergologie utilise un éthogramme composé de 12 catégories d'objets nommées: « les gestes d'autocontact sur le visage », « les microréactions sur le visage », « les axes de la tête », « les gestes d'autocontact sur le cou », « les geste d'autocontact sur le torse », « les microattitudes ou gestes d'autocontact sur le ventre », « les microattitude ou gestes d'autocontact sur le bassin », « les gestes d'autocontact sur les membres supérieurs », « les geste d'autocontact ou microcontact sur les membres inférieures », « les gestes dans l'espace », « les boucles

de rétroaction », « les boucles de rétroactions secondaires », « les gestes de préhension », « les positions assises sur la chaise ». Chacune de ces catégories est composée de nombreuses étiquettes codées comme par exemple : *A\_0\_D\_N\_30\_56* : « *Autocontact sur le visage microdémangeaison dans la zone du nez sous le nez en partant vers l'extérieur avec la main droite* ».

Une catégorie pouvant contenir jusqu'à 852 étiquettes, nous ne les détaillerons pas ici, mais tous ces codes et leur description sont disponibles sur le site officielle de la synergologie<sup>23</sup>.

### **Les approches sur lesquelles se basent la synergologie**

La synergologie donne accès à des manuels de lecture mais ne présente pas ou peu de fondements théoriques sur ces observations car il s'agit d'une science empirique. Elle se nourrit de différentes études issues des sciences affectives et tout particulièrement de celles qui s'intéressent au « non verbal ». Ces études sur le non verbal sont extrêmement nombreuses et les synergologues dressent une bibliographie entière de lectures dont ils s'inspirent et qui permettent de comprendre leur démarche<sup>24</sup>. Nous retrouvons parmi ses sources les grands noms des théories émotionnelles tels que Darwin, Ekman ou Damasio que nous avons vu dans le Chapitre 2. La synergologie s'appuie également sur des études de neurobiologies axées sur des caractéristiques asymétriques du fonctionnement du visage et du corps (Kimura et Humphrys, 1981 ; Skinner, 1991 ; Borod, 1992 ; Buser, 1998). Elle s'intéresse également aux observations utilisant l'éthologie comme les études de Tinbergen sur des enfants autistes (Tinbergen, 1972), les dimensions sociales de la gestualité (Calbris, 2003) ou sur des sujets variés faisant intervenir la gestualité tel que la communication, la kinesthésie, la mémoire et bien d'autres phénomènes encore.

### ***Recueil d'analyse de synergologie***

Comme nous l'avons précisé précédemment, nous ne cherchons pas à émettre un jugement sur la synergologie mais nous voulons en faire un objet d'observation scientifique, et cela de manière systématique y compris dans le cadre de réflexions ultérieures. Ainsi il nous faut pouvoir appréhender cette nouvelle démarche sous un angle

---

<sup>23</sup> Nomenclature de l'éthogramme disponible à l'adresse : <http://www.synergologie.org/ethogramme> (consulté en mars 2012)

<sup>24</sup> Voici une liste d'ouvrages scientifiques en lien indirect avec la synergologie : <http://www.synergologie.org/regards-croises/ouvrages-interessants-sur-la-communication-non-verbale> (consulté en avril 2012)

d'observation scientifique et envisager des manipulations expérimentales adéquates pour l'avenir. Nous avons donc mis en place un premier protocole d'étude préliminaire complémentaire aux observations de nos tests de perception. Celle-ci est certainement imparfaite car il s'agit d'un nouvel essai en la matière, mais elle nous permettra d'avoir un premier aperçu de cette discipline et d'effectuer les rectifications nécessaires pour de prochaines études.

### **Contexte de recueil et présentation des sujets synergologues**

Le nombre de synergologues étant très restreints en France (moins d'une dizaine) et même dans le monde (une petite centaine), une démarche quantitative était d'ores et déjà exclue dans la récolte de notre corpus d'analyses synergologiques. Sur Grenoble, nous avons pu prendre contact avec un synergologue<sup>25</sup> par le biais de P. Turchet. Ce synergologue a accepté d'analyser les mêmes stimuli que ceux que nous présentons à nos sujets, à savoir ceux du test de perception présenté aux Français (De Biasi, 2011) ainsi que ceux de mon propre protocole expérimental avec les Japonais puisqu'il s'agit des mêmes stimuli. En revanche, comme ces spécialistes affirment n'utiliser que le visuel, nous avons supprimé les sons des vidéos pour être sûr que l'analyse soit entièrement basée sur cette unique modalité. Par ailleurs les fichiers ont été entièrement renommés aléatoirement afin que les synergologues ne puissent retrouver les associations étiquette/stimulus. Par ailleurs, pour que le choix soit purement effectué en fonction de la valeur sémantique des étiquettes (et non sur des indices grammaticaux de genre), les étiquettes des locutrices avec des genres saillants ont été neutralisées.

Etant donné que l'analyse en soi ne nécessite pas de cadre expérimental physique spécifique, nous avons une fois encore réalisé le recueil des analyses à distance. Cependant pour pouvoir guider et contrôler la direction des analyses pour la récolte d'un matériel exploitable, nous avons mis en place un protocole basé sur deux questionnaires. Ceci a été également précédé d'une entrevue pour expliquer plus précisément la tâche au synergologue mais aussi pour avoir directement certaines informations concernant la synergologie.

---

<sup>25</sup> En réalité il ne s'agit pas du premier contact mais un collègue du premier synergologue qui a effectué en définitive l'analyse.

### **Protocole d'analyse imposé aux synergologues**

Pour recueillir les analyses de synergologie nous avons estimé, peut-être à tort, que le protocole des naïfs serait trop facile pour les synergologues. En effet, le choix d'une étiquette d'auto-annotation parmi un nombre très restreint d'étiquettes d'un seul locuteur nous semblait réduire les champs d'observation et l'étendue de ce que la synergologie pouvait donner en termes d'analyse. Ainsi, bien qu'il nous soit dès lors impossible de faire des comparaisons directes entre les réponses des sujets naïfs et celles des synergologues, nous pourrions tout de même mettre en évidence une tendance sur le choix des étiquettes par le biais de notre protocole.

Dans un premier temps, le synergologue a été amené à choisir une étiquette d'auto-annotations parmi toutes les étiquettes associées aux 66 stimuli qui ont été sélectionnés. Ces étiquettes ont cependant étaient regroupées par locuteur pour pouvoir faciliter à minimum la présentation des auto-annotations et donner une cohérence qui est fonction du locuteur qui les a fournies. Nous avons ainsi laissé la liberté de choisir les étiquettes parmi ceux de n'importe quels locuteurs en restreignant tout de même ce choix à une étiquette parmi les choix que propose un locuteur afin de pouvoir garder une base de comparaison avec les réponses des naïfs. De plus, c'est une manière de contraindre les synergologues sur une étiquette globale, la plus saillante car pour ces spécialistes, nos stimuli qui nous paraissent pourtant court en tant que naïfs, sont parfois décomposables en plus d'unités qui sont chacune susceptible d'être annotée par des étiquettes différentes. Néanmoins, pour ne pas perdre ces analyses détaillées, nous avons laissé un espace de commentaire libre où le synergologue était libre de faire son analyse avec ses propres critères. Cette mise en place à fait l'objet d'un questionnaire (cf. Annexe 15). Ainsi pour chaque stimulus nous avons demandé de remplir un questionnaire, afin de pouvoir récupérer les résultats des analyses au fur et à mesure<sup>26</sup>.

Puis un second questionnaire consistait à faire l'association d'un ensemble d'étiquettes (correspondant à celle donné à un locuteur) avec un ensemble de stimuli représentants produits chaque locuteur. En gros la tâche consistait à voir si les synergologues pouvaient retrouver l'ensemble d'étiquettes associées à un locuteur en s'appuyant sur un éventuel état global qui qualifierait le locuteur à partir de ses auto-

---

<sup>26</sup> Finalement cette procédure n'a pas servie car les résultats nous ont été communiqués en un seul envoi

annotations. Ce deuxième questionnaire (cf. annexe 16) était l'objet d'une réponse suite à l'analyse de tous les stimuli d'un locuteur.

### ***Résultats des questionnaires***

Nous avons rassemblé toutes les réponses du synergologue dans un tableau (cf. Annexe 17). Dans cette section nous présenterons donc les résultats de ce protocole en rappelant en parallèle les réponses données par les sujets naïfs français, sur la modalité visuelle de leur test de perception, en nous appuyant sur les matrices de confusion des Français naïfs (De Biasi, 2011) (cf. Annexe 8). La présentation de ces résultats se fera étiquette par étiquette pour chaque locuteur sujet.

#### **Les observations sur le sujet T**

Le sujet T présente 10 stimuli soit 5 étiquettes à retrouver dont :

**Etonnée** : Les naïfs voient principalement de la « déception » dans les stimuli TIP01 et TIP 02 bien qu'il y ait des confusions notables entre « étonnée », « angoissée » et « hésitation ». Chez les synergologues, TIP 02 est reconnu comme stimuli du sujet T et lui associe de « l'hésitation » voire même de la « peur » d'après leurs critères. TIP01 est lui du « doute » (étiquette de M). Ils soulignent en revanche une similitude entre les deux stimuli en notant des indices de la « peur » qui pour eux se voit à l'aide de plissement du front. Donc aussi bien les naïfs que les synergologues ne reconnaissent pas l'étonnement décrits par T.

**Déception** : Les naïfs reconnaissent bien la déception dans les signaux TIP03 et TIP04. En revanche les synergologues ne la reconnaissent pas comme étiquette saillante. En effet, ils ont préféré choisir « l'hésitation » pour le stimulus TIP04 et ne se prononcent pas sur une étiquette précise pour TIP03 mais notent un certain « stress » illustré par un faux sourire. Notons tout de même que les synergologues conservent une certaine cohérence puisque quand ils choisissent l'étiquette « hésitation » ils remarquent à chaque fois les indices de leur « peur ».

**(Assez) calme** : Pour les naïfs, les stimuli TIP05 et TIP06 représentent très largement de « l'hésitation ». Les synergologues ont choisi le « doute » (du sujet M) pour le stimulus TIP05 et ne se sont pas prononcés pour TIP06 car d'après leurs critères la séquence est très courte. En tout cas, nous pouvons noter que le choix des synergologues est proche de celui des naïfs via le « doute ». De plus, nous retrouvons l'indice des rides

sur le front désignant la « peur » dans TIP05, or nous avons vu qu'à chaque fois qu'ils avaient précédemment évoqué cet indice, ils avaient choisi l'étiquette « hésitation », ce qui confirme la proximité de choix avec les sujets naïfs.

**Angoissée** : les naïfs reconnaissent bien cette étiquette en visuel sur les stimuli TIP07 et TIP08. Les synergologues eux font à nouveau le choix de « l'hésitation » sur TIP08 et ne proposent pas d'étiquette sur TIP07. Mais comme dans tous les cas où l'étiquette « hésitation » a été choisie par les synergologues, ils notent des indices de la « peur ». Or cette « peur » est proche de l'étiquette « angoissée » donc on peut supposer que la reconnaissance est également réalisée par les synergologues. D'ailleurs, l'indice de la peur est également présent dans le stimulus TIP07 d'après ces derniers.

**Hésitation** : chez les naïfs, les stimuli TIP09 et TIP10 ne permettent qu'un choix d'étiquettes au hasard. Pour le synergologue, nous avons également deux étiquettes différentes pour ces stimuli. TIP09 est défini comme de « l'inquiétude » (du sujet J) et TIP10 comme de la « déception ».

Au final, le synergologue a réussi à retrouver le locuteur T à partir des stimuli et des étiquettes qui lui ont été fournis.

### **Les observations sur le sujet S**

Le sujet S présente 12 stimuli soit 6 étiquettes à retrouver dont :

**Surprise par les résultats car un doute persistait sur la prononciation. Néanmoins les résultats donnent une certaine satisfaction personnelle** : SAB01 et SAB02 sont très mal reconnus par les naïfs pour qui les choix se font au hasard. Cela est également vrai pour le synergologue interrogé. Cependant, ce dernier trouve des indices d'une véritable assertion dans SAB01 ce qui rejoint l'étiquette au niveau de la satisfaction personnelle. Il a donc choisi l'annotation « Sûr de moi / (bien) à l'aise (du sujet M). En revanche, SAB02 lui semble trop court et il y perçoit du « stress » sans confirmer d'étiquette parmi celles que nous avons proposées.

**Toujours la concentration, l'écoute attentive, un peu déçue du résultat et de la difficulté. Néanmoins le résultat n'est pas pris comme un échec cuisant** : Les naïfs considèrent SAB03 et SAB04 comme de la « déception ». Pour le synergologue, SAB03 représenterait un « léger doute » (du sujet N) et détecte un vrai sourire, ce qui semble avoir un lien avec le « néanmoins (...) » de l'étiquette. Pour SAB04, ce dernier ne choisit aucune

étiquette car la séquence semble couper un mouvement qui serait utilisé dans leur nomenclature. Il détecte cependant un « problème lié à soi » qui est peut-être un rappel de la « déception » mais ici nous n'avons pas vraiment d'observations concluantes.

**Mission impossible- concentration mais réponse au pif :** Chez les naïfs il n'y a pas de réponses saillantes sur SAB05 et SAB06, seulement une légère tendance vers la « surprise » qui n'est pas confirmée au niveau statistique. Le synergologue lui a choisi la « peur » (du sujet M) pour SAB05 et « ne suis plus sûr de moi » (« doute » de N) pour SAB06. Dans SAB05 il détecte également un « vrai sourire » qui suggère la proximité du « au pif » et SAB06 est finalement assez proche de l'étiquette, sachant que la synergologie remarque également des indices de « dégoût/colère » et encore une fois un haussement de sourcil gauche (assez fréquent chez le sujet puisqu'il apparaît 4 fois), ce qui suggérerait un « problème lié à soi ». Nous pouvons donc voir qu'ici la synergologie a de meilleurs résultats que chez les naïfs.

**Déçu par les résultats, j'essaie de trouver des solutions :** Les naïfs répondent au hasard sur les stimuli SAB07 et SAB08 malgré une légère tendance non confirmée statistiquement pour un choix tourné vers la « concentration ». Le synergologue a lui assez bien reconnu la « déception » des stimuli. Pour SAB07, il a choisi « étonné » (du sujet T) ce qui ne semble pas réellement correspondre alors que dans ses analyses, il utilise le mot « dépit » qui est beaucoup plus proche de la « déception ». Pour SAB08 il choisit bien l'étiquette d'une déception (celle du sujet N) qu'il associe également à du « mépris ».

**Au pif, une envie de rigoler :** Les naïfs reconnaissent bien cette étiquette mais font une confusion assez marquée avec la « surprise » (associée à une « satisfaction ») pour les stimuli SAB09 et SAB10. Les analyses du synergologue pour ces deux stimuli sont relativement semblables noté d'un « clignement » avec la présence d'un vrai « sourire » dans les deux cas. Donc « l'envie de rigoler » s'illustre par ces indices. Concernant les étiquettes choisies, SAB09 se voit attribuer « soulagé » (de N) qu'on ne peut pas vraiment comprendre, mais SAB10 est associé à « ne suis plus sûr de moi » (de N) qui est proche de l'étiquette « au pif ». Donc ces stimuli marchent bien pour les naïfs et le synergologue.

**Le but est de trouver des solutions à ces résultats catastrophiques : tous les moyens sont bons et en plus ça rassure :** Les sujets français ne reconnaissent pas cette étiquette, ils répondent au hasard avec une légère préférence pour « surprise » et « concentration » sans confirmation statistique. Le synergologue choisit l'étiquette

« étonné » (de T) pour SAB11, ce qui rejoint la « surprise » des naïfs. De même il associe « ne suis plus sûr de moi » au stimulus SAB12, ce qui n'est pas incompatible avec la « concentration ». Nous avons donc des similitudes entre les naïfs et le synergologue.

Le synergologue a associé l'ensemble d'étiquettes de N aux stimuli du sujet S, donc les étiquettes de S ne lui ont pas permis de la reconnaître, et l'état global qu'il en retire est plus proche de celui de N.

### **Les observations sur le sujet M**

Le sujet M présente 12 stimuli soit 6 étiquettes à retrouver dont :

**Peur** : Notons que ce sont les stimuli de MAR01 et MAR02 qui sont dupliqués car l'annotation ne correspondait qu'un à un seul stimulus. Chez les naïfs, l'étiquette attribuée au stimulus est le « doute ». Le synergologue a bien remarqué que les deux stimuli MAR01 et MAR02 étaient identiques et n'a fait qu'une analyse. Il associe cette vidéo à « Mission impossible – concentration mais réponse au pif » (du sujet S).

**Doute** : Les stimuli MAR08 et MAR09 ne sont pas significativement reconnus mais sont marqués par une tendance vers le « doute » chez les naïfs. Pour le synergologue, MAR07 est associé à « Mission impossible » (de S), un choix d'étiquette qui était associé au « doute » naïf pour le stimulus de la « peur » du sujet M. En revanche, il semble insatisfait de cette étiquette car il pense qu'elle est fautive. MAR08 est quant à lui associé à « Mon incompréhension est justifiée ! » (de R) qui est une étiquette relativement proche du « doute ». Donc les réponses des naïfs et du synergologue sont assez semblables.

**Stress** : Chez les naïfs, les stimuli MAR03 et MAR04 sont significativement associés au « doute » et à « pas contente-déçue ». Le synergologue semble avoir des avis assez semblables car il choisit la « concentration » (du sujet R) pour MAR03, ce qui le rapproche du « doute » des naïfs, et « agacé » (de R) pour MAR04, ce qui rappelle le « stress » de ces stimuli mais également « pas contente-déçue » qu'avait choisi les sujets naïfs. Pour cette dernière étiquette, le synergologue semble être assez sûr de lui. Donc à nouveau nous avons une similitude entre naïfs et synergologue.

**Sûre de moi, (bien) à l'aise** : Les naïfs répondent au hasard pour les stimuli MAR05 et MAR06 mais nous pouvons relever une petite préférence non significative pour l'étiquette « stress ». De même, le synergologue associe sur ces deux stimuli l'étiquette « agacé » (du sujet R) mais note également à chaque fois le « mépris » dans ses analyses

détaillées. Ce « mépris » est assez proche de l'assurance que le sujet M se dit avoir. Donc certains naïfs pressentent peut-être une attitude négative qu'ils définissent comme du « stress » que la synergologie elle qualifie systématiquement de « mépris ».

**Pas contente-déçue** : chez les naïfs le choix des étiquettes de MAR09 et MAR10 est complètement aléatoire. Pour les synergologues, MAR09 est associé à l'étiquette « au pif, une envie de rigoler » (de S) et MAR10 à une « petite irritation » (de J). MAR10 est donc bien reconnu par le synergologue alors que le choix de MAR09 est moins saisissant. D'ailleurs les indices des vidéos sont assez différents d'après la nomenclature de synergologie. Notons également qu'à nouveau, sur MAR10, le synergologue note du « mépris ».

**Déçue** : Les sujets naïfs font là encore un choix relativement aléatoire pour MAR11 et MAR12 mais notons tout de même qu'ils hésitent essentiellement entre des étiquettes antagonistes opposées « Sûre de moi » et « doute » bien que ce choix ne soit pas statistiquement significatif. Le synergologue attribue quant à lui, « au pif, une envie de rigoler » (de S) ou du « dépit » à MAR11 et « surprise » (de S) ou du « mépris » à MAR12. Les choix « dépit » et « mépris » que le synergologue a sélectionné de sa nomenclature semblent relativement proches de la notion de « déception » contrairement aux étiquettes choisies par les naïfs. Notons également que le sujet M semble avoir des stratégies d'expressions de la déception assez différente entre les deux stimuli, l'un avec un « sourire prononcé » qui expliquerait le choix des sujets vers l'étiquette « Sûre de moi » et le second avec beaucoup d'indices visuels d'après la synergologie, qui ferait transparaître du « doute » pour les naïfs.

Au final le sujet M est associé aux étiquettes du sujet S par le synergologue. En terme d'attitude général, nous avons pu voir que le synergologue proposait souvent le « mépris » pour décrire les attitudes de M. Or ce trait avait également été noté par les expérimentateurs qui ont créé le corpus. En effet, ces derniers qualifiaient M comme « méprisable vis-à-vis de l'outil qu'elle était en train de “tester” ».

### **Les observations sur le sujet J**

Le sujet J présente 8 stimuli soit 4 étiquettes à retrouver dont :

**Inquiétude sur le but du jeu** : Chez les sujets naïfs, JEA01 et JEA02 sont bien reconnus comme de « l'inquiétude ». Pour le synergologue JEA02 serait du « stress » (de M) aussi bien comme auto-annotation que dans sa propre terminologie. JEA01 est associé

à « déçue » (de M). D'après sa nomenclature, il se rapproche de cette inquiétude car il note pour ce stimulus « de l'incompréhension ou de la concentration » en s'appuyant sur des « froncements de sourcils » et éventuellement un mouvement incomplet de recul qui serait une attitude de « fuite ». Donc là où nous avons une assez bonne reconnaissance chez les naïfs, la synergologie relève plutôt des marqueurs de cet état émotionnel.

**Petite irritation** : Les stimuli JEA03 et JEA04 sont également bien reconnus par les sujets naïfs. Le synergologue associe JEA03 à du « doute » (de M) et JEA04 avec « sûre de moi, / bien à l'aise » (de M), ce qui n'est pas réellement très proche de « l'irritation ». En termes de nomenclature synergologique, JEA03 serait illustré par un mouvement de retrait et des indices de « ne sais pas ». JEA04 porterait des marques de « délectation » à cause des caractéristiques de la « langue ». Nous notons donc une vraie différence entre les sujets naïfs et les synergologues pour ces stimuli.

**Petite lassitude - l'indifférence s'installe** : Les naïfs ne reconnaissent pas l'auto-annotation de JEA05 et JEA06. Le synergologue lui associe JEA05 à « agacé » (de R) ce qui se rapproche de l'étiquette et JEA06 à « étonnée » (de T), une attribution assez différente de l'auto-annotation. Dans JEA05 nous soulignons à nouveau les marques de la délectation, mais également du « mépris » et du « dépit ». JEA06 est lui synonyme de « problème lié à soi » et de « retrait ou fuite ». La nomenclature synergologique s'approche donc plus de l'auto-annotation qui a été proposée par le sujet J.

**Je commence à comprendre qu'on se paye ma tête** : Les naïfs choisissent au hasard les étiquettes pour les stimuli JEA07 et JEA08. Le synergologue attribue à JEA07 l'annotation « ne suis plus sûr de moi, léger agacement » (de N) et ne choisit pas d'étiquette pour JEA08 mais désigne ce stimuli par les termes « empathie » et « dépit ». De même, JEA est associé à des termes du vocabulaire synergologique tels que « retrait », « délectation » et « dépit ». Donc aussi bien chez les naïfs que chez le synergologue, cette étiquette est mal reconnue.

Le synergologue attribue les étiquettes du sujet M au sujet J. De manière générale, nous retrouvons souvent les indications synergologiques « retrait, fuite » ou « dépit » pour ce qui concerne le locuteur J. Or il est à noter également que durant le recueil de corpus, les expérimentateurs ont eu un sentiment d'incertitude de J face aux tâches linguistiques qui lui étaient présentées, du fait que J soit non spécialiste. Peut-être trouvons-nous donc

cet état global dans les indices de la synergologie, comme nous avons pu le voir précédemment pour M.

### **Les observations sur le sujet N**

Le sujet N présente 10 stimuli soit 5 étiquettes à retrouver dont :

**Léger doute** : Les stimuli NIC01 et NIC02 sont reconnus comme « essaie de rester concentré » par les sujets naïfs français des tests de perception. Le synergologue lui a choisi les étiquettes « léger doute » pour NIC01 soit la bonne étiquette du bon locuteur. Il attribue d'autre part l'étiquette « étonnée » (de T) au stimulus NIC02, ce qui est moins évident à comprendre. Cependant, dans ses critères synergologiques, nous voyons qu'il trouve des indices d'« étonnement » et de « concentration » pour NIC01, deux descriptions qui rejoignent à la fois la réponse des naïfs et sa réponse sur NIC02.

**Ne suis plus sûr de moi, léger agacement** : les naïfs ne répondent pas de manière significative sur les stimuli NIC03 et NIC04, mais présentent leurs plus fortes confusions pour « déception » et « soulagé ». En synergologie, NIC03 se voit attribuer « sûr de moi / (bien) à l'aise » (de M) et « essaie de rester concentré » (de N). Les nomenclatures synergologiques remarquent 2 signaux : « sourire » et « clignement ». Bien qu'on puisse trouver une certaine proximité des réponses naïves et synergologiques, aucun indice n'est véritablement explicite sur ces stimuli.

**Essaie de rester concentré** : Pour NIC05 et NIC06, les français naïfs reconnaissent bien la « concentration ». C'est également le cas du synergologue qui associe une étiquette « concentration » (de R) pour le stimulus NIC06. En revanche, NIC05 se voit attribuer l'étiquette « déception » (de T), un choix qui ne tente pas les naïfs bien qu'il existe une étiquette « déception » chez N puisque très peu de confusion sont faites sur cette étiquette.

**Déception mais m'en amuse** : la « déception » est bien l'étiquette largement choisie par les sujets naïfs pour les stimuli NIC07 et NIC08. Pour les synergologues, cette réponse n'est pas évidente puisque NIC07 illustrerait du « doute » (de T) et NIC08 n'est associée à aucune étiquette car la séquence serait trop courte. Donc la naïveté des sujets semblent ici plus efficace que le filtre synergologique. En revanche nous savons que la synergologie s'intéresse autant aux émotions sous-jacentes et sur-jacentes, ce qui pourrait avoir un lien avec la divergence observée.

**Soulagé :** Les stimuli NIC09 et NIC10 ne sont pas très reconnus par les naïfs et aucune confusion n'est véritablement saillante. Pour les synergologues, NIC09 souligne bien le « soulagement » et d'ailleurs c'est l'étiquette de N qui a été utilisé pour désigner le stimulus. En revanche il associe NIC10 à la « surprise » de S, une étiquette qui désigne aussi « une satisfaction personnelle » que l'on peut comprendre comme un certain « soulagement ». Donc les résultats de la synergologie sont meilleurs que ceux des naïfs pour ces stimuli-ci.

Malgré les bonnes réponses trouvées sur les étiquettes de N, le synergologue n'a su reconnaître les étiquettes de N comme lui appartenant. Il ne fait d'ailleurs aucune association entre groupe d'étiquettes et locuteur pour les stimuli de N. Deux d'entre elles ont bien été utilisées mais il existe aussi de grandes divergences entre les réponses des sujets naïfs et les auto-annotations choisies par le synergologue.

#### **Les observations sur le sujet R**

Le sujet R présente 14 stimuli soit 7 étiquettes à retrouver dont :

**Concentration :** Les stimuli ROB01 et ROB02 sont très bien reconnus chez les sujets naïfs. ROB02 l'est aussi par le synergologue qui a su choisir la bonne étiquette du bon locuteur. En revanche, ROB01 est étiqueté comme « agacé » (de R) où une « déglutition » semble faire penser à du « stress ». Donc ROB02 est clair aussi bien pour les naïfs et les synergologue alors que le stimulus ROB01 laisse à réfléchir.

**Ennui-concentration :** Il n'y a pas de réponses claires chez les naïfs mais nous observons une tendance logique vers les étiquettes « concentration » et « ennui-concentration », donc une assez bonne reconnaissance des stimuli ROB03 et ROB04. Pour le synergologue, ROB04 est lui aussi reconnu comme de la « concentration » notamment à partir de ce qu'il appelle « moue sur droite » puisqu'il en définit des dénominations telles que « réflexion » et « hésitation » qui sont proches de la « concentration ». Pour le signal ROB03 il note du « stress » (de N) mais nous n'en voyons pas véritablement les raisons. Nous retrouvons tout de même le marqueur ASI qu'il avait utilisé pour le stimulus de la « concentration », ROB01, pour lequel il avait choisi l'étiquette « agacé ».

**Ennui :** ROB05 et ROB06 se voient attribuer les étiquettes d'« agacé » chez les naïfs. Pour le synergologue nous avons une certaine similitude puisque ROB06 est associé à « stress ». Cependant d'après des observations sur la « lèvre inférieure », ROB05 mais également ROB06 serait illustratif d'une certaine « peur » (que l'on ne retrouve d'ailleurs

que chez M). Donc aussi bien chez les naïfs que chez les synergologues, nous avons la reconnaissance d'une certaine forme de « stress ».

**Ennui, incompréhension** : Chez les naïfs, nous avons une bonne reconnaissance des stimuli ROB07 et ROB08. Pour les synergologues ROB08 est associé au « doute » (de M) et ROB07 n'a pas d'étiquette car il semblerait que la séquence soit trop courte pour pouvoir déterminer de quoi il s'agit. Le « doute » s'apparente d'une certaine manière à « l'incompréhension » donc le choix du synergologue n'est pas mauvais. Pour ROB07 il est vrai que la séquence est coupée avant la fin d'un geste expiratoire qui semble assez clair, mais c'est certainement la méthodologie de la synergologie qui impose l'observation complète d'un signe avant de le définir.

**Surprise, nervosité / surpris, agacé** : ROB09 et ROB10 sont reconnus comme « Mon incompréhension est justifiée ! » chez les sujets naïfs, ce qui reste assez proche de l'auto-annotation donnée par R. Pour le synergologue ROB09 est associé à la « peur » (de M) et ROB10 est sans étiquette encore une fois car la séquence semble trop courte pour conclure. En revanche, la « peur » peut aussi illustrer la partie « surprise/nervosité » de l'étiquette. Donc Les naïfs et le synergologue sont relativement complémentaires vis-à-vis de l'étiquette même si le focus émotionnel ne semble par être sur la même cible, certainement parce que les indices sur lesquels ils s'appuient (ou considèrent comme plus importants) sont différents.

**Agacé** : Les stimuli ROB11 et ROB 12 ne trouvent pas d'étiquettes adéquates selon le choix des sujets naïfs. Pour le synergologue, ROB12 est bien reconnu avec la bonne étiquette du bon locuteur et nous voyons en plus les termes « dégoût » et « colère » qui leur sont associés. En revanche, il attribue à ROB11 l'étiquette « hésitation » (de R), ce qui ne correspond pas véritablement à l'auto-annotation.

**Mon incompréhension est justifiée !** : Pour les naïfs c'est surtout l'étiquette « ennui-incompréhension » qui a été attribuée aux stimuli ROB13 et ROB14, ce qui reste proche de l'auto-annotation. Pour le synergologue, ROB13 s'est vu attribué la bonne annotation mais ROB14 reste sans annotation toujours pour la même raison de durée trop courte. Donc globalement nous pouvons dire que les naïfs et le synergologue ont d'assez bons résultats sur ces deux stimuli.

Le synergologue n'a pas attribué de locuteur pour les stimuli de R. Pour les stimuli de ce locuteur, il note principalement des indices de la « peur » et « l'incompréhension »

par les termes « dubitatif », « ne sais pas ». Par ailleurs, ce groupe de stimuli a eu de nombreuses remarques sur la durée, jugée trop courte pour certains signaux pour permettre de réaliser une annotation, tout au moins sur des critères synergologiques.

### ***Conclusion du Chapitre 7***

Le choix des étiquettes des naïfs et du synergologue sont relativement semblables de manière générale. En revanche nous avons également vu que parfois il y avait des divergences qui ne sont pas réellement motivées chez les naïfs mais souvent en faveur des synergologues. Il peut parfois tout de même s'avérer que la recherche d'indices fins peut parfois être nuisibles à la reconnaissance d'évidences. Donc nous devons rester vigilant avec la manipulation de ces données et ne pas prendre tout pour acquis. La première observation de synergologie nous laisse à penser que cette approche nous sera également utile à l'avenir, notamment quand nous aborderons les thèmes de l'inhibition, puisque cette discipline s'intéresse au non-dit. Mais de toute évidence elle nous est utile pour observer les événements visuels rien qu'en terme d'analyse descriptive grâce au niveau de détail des observations faites par ces spécialistes.

## **Partie 4**

### **Perspectives et enjeux applicatifs**

## **Chapitre 8 – Bilan de l'étude et perspectives**

A partir des éléments recueillis nous avons réalisé une infime partie des possibles en termes d'analyses. Le protocole perceptif que nous avons utilisé a montré ses limites : la tâche en choix fermé est vraiment complexe, compte tenu du nombre et de la complexité des étiquettes, et de la variabilité du contenu informatif des stimuli observés – variabilité que nous souhaitons précisément mesurer. L'étude contrastive entre la perception par des français et japonais nous renseigne : (1) par ses similarités dans la non reconnaissance de stimuli => sur la valeur non informationnelle de ces stimuli (2) par ses similarités sur la reconnaissance comparable de stimuli => sur la valeur informationnelle non langagière (innée) ou bien universellement langagière de ces stimuli (3) par ses différences dans la reconnaissance => sur la valeur informative langagière de ces stimuli .

Ce dernier chapitre tentera donc de présenter les points qui n'ont pas pu être abordés, voire les améliorations que nous pouvons apporter sur les différents aspects de notre étude.

### ***Les problèmes et difficultés mis en évidence***

Etant donné que nous avons trois parties qui précèdent ce chapitre, nous mettrons en évidence les difficultés et les problèmes liés à ces parties.

#### **Les difficultés que souligne l'état de l'art**

L'état de l'art de ce mémoire a un gros défaut : celui d'être incomplet. Bien entendu nous ne pouvons parler de tout ce qui existe dans un état de l'art mais pour ce cas précis, il me paraît manquer une grosse partie sur l'apport des disciplines des sciences du langage pour expliquer en partie la forme (et pas seulement le contenu) des micro-expressions. Nous avons également une grosse lacune concernant les apports de la pragmatique sur ce sujet. Malheureusement je n'ai pas eu suffisamment de temps pour développer cette partie.

Par ailleurs, les difficultés de l'état de l'art sur un tel sujet est la sélection de la quantité d'informations. En effet, étant donné que le sujet a parfois une longue histoire et surtout que les disciplines qui s'intéressent à ce mémoire sont extrêmement nombreuses, résumer les informations a été une grande difficulté.

### **Limites et problèmes des protocoles expérimentaux**

Le premier défaut, qui est aussi un atout des protocoles, est comme nous l'avons vu la modalité à distance sur laquelle nous ne reviendrons pas car ce sujet a largement été développé auparavant. De plus, nous avons également le fait que l'expérience n'est pas entièrement bilingue, il aurait été intéressant d'avoir aussi bien des productions de Japonais que de Français. Un troisième point concernait la robustesse des programmes. Par ailleurs, au niveau du contenu de l'expérience, il existe un petit défaut sur la question du degré de confiance. En effet, l'ajout de la question sur le caractère japonais et étranger des micro-expressions pour la modalité auditive induit une ambiguïté. Nous ne savons pas si les sujets associent le degré de confiance à la réponse portant sur ce caractère japonais/étranger, à l'ensemble des réponses ou uniquement sur le choix de l'étiquette.

D'une manière générale, dans les matrices de confusions, les taux de reconnaissance des stimuli restent globalement faibles, y compris pour les meilleurs, par rapport aux taux de reports. Plus précisément, la répartition de ces confusions est très diffuse. Ceci est important de noter car un report massif, nous l'avons vu, est très informatif sur une potentielle catégorisation plus large. Autrement dit les résultats sont statistiquement très bruités. Ainsi, associer cette expérience à d'autres protocoles à plus faible charge cognitive, comme par exemple des présentations par paires, serait sans doute très intéressant pour mieux comprendre et interpréter globalement nos résultats.

Pour l'évaluation par les synergologues, nous aurions finalement pu peut-être garder la sélection de la modalité de questionnement comme pour les sujets naïfs. Nous avons fait ce choix de modalité strictement visuelle, car les synergologues travaillent avec une grammaire orientée strictement sur le visuel, mais il aurait été intéressant de vérifier malgré tout que leurs compétences implicites n'utilisent jamais le canal oral.

### **Les non-dits de l'analyse**

L'une des perspectives prévues par les études antérieures, et à laquelle nous n'avons pas eu le temps de répondre, est l'analyse psychoacoustique des stimuli utilisés. En effet, cette analyse aurait complété les observations du visuel et la classification des micro-expressions. Malheureusement le logiciel Praat ne permet pas de faire des analyses assez fines sur les données du fait des caractéristiques acoustiques très particulières de nos données. De plus, le temps nous a tout simplement manqué.

Nous n'avons pas non plus pu analyser les réponses de nos sujets chinois.

L'analyse de la perception par les Japonais que j'ai présentée reste malgré tout superficielle et descriptive, l'observation plus poussée des données est donc absolument nécessaire.

### ***Les perspectives d'amélioration des problèmes soulignés***

Voici quelques pistes d'améliorations possibles pour nos expériences et les solutions apportées aux difficultés rencontrées.

#### **Les choix et une pluridisciplinarité ouverte mais contrôlée**

Les expériences que j'ai conduites sont dans la continuité des travaux du GIPSA qui ont été présentés dans l'état de l'art de ce mémoire. Il m'a en effet paru important de comprendre de manière détaillée ce sujet de recherche puis son origine. D'ailleurs c'est parce que je n'ai pas eu le temps de traiter la forme même des micro-expressions ni les considérations pragmatiques qui leurs sont associées, que j'ai privilégié les travaux antérieurs de ce laboratoire puisqu'ils abordaient ces aspects.

L'ouverture du domaine des sciences cognitives et affectives est un besoin réel mais il convient de rester vigilant et de ne pas dériver des objets qui nous intéressent.

#### **Les améliorations des protocoles de perception et de synergologie**

L'amélioration des protocoles passe par les révisions algorithmiques du programme du test de perception pour éviter les non réponses des sujets interrogés.

#### **L'analyse acoustique des données**

Philippe Martin a développé et commercialisé un outil nommé Winpitch qui permet à la fois de faire des analyses précises, en particulier de la prosodie. Cet outil permet également de réaliser des resynthèses de certains objets acoustiques et il sera alors intéressant d'analyser nos stimuli en fonction de la reconnaissance par les sujets japonais versus français et de transplanter certains paramètres en resynthèses pour en mesurer les incidences perceptives.

## Conclusion

Nous avons essayé de mesurer ici conjointement la valeur culturelle et la portée informative de micro-événements audio-visuels, en donnant à percevoir des micro-expressions de sujets français à des sujets japonais, dans chaque modalité séparée puis dans la bimodalité. Nous avons comparé nos résultats avec ceux obtenus dans un travail précédent pour la perception par des sujets français. Ce que nous pouvons très globalement retirer de cette étude, c'est que ces stimuli, dans leur modalité acoustique en particulier, sont étonnamment bien reconnus par les Japonais en comparaison avec les Français. Nous savons que la langue et la culture japonaise utilisent beaucoup la qualité de voix et possède un nombre varié d'interjections et d'onomatopées, très fréquemment utilisées (Matsumoto, 1988 ; Ishi, 2012 ; Maekawa 2004). On peut peut-être trouver là une explication de cette performance perceptive. Ce qui est le plus intéressant, ce sont les différences dans le comportement perceptif entre les français et les japonais, par exemple sur les étiquettes d'hésitation, qui pourrait être un indice supplémentaire à celui apporté par l'étude de Signorello (2010) dans son étude contrastive italo-française, qui cependant ne mesurait la qualité de la reconnaissance de ces stimuli, mais uniquement leur perception d'appartenance à une langue. Il faudrait maintenant entrer beaucoup plus finement dans cette analyse en ciblant sur chaque sujet plus précisément (le sujet R a par exemple été reconnu de façon assez semblable par les sujets français et japonais, ce qui signifierait que ses productions sont informatives mais non empruntées par la langue), et également sur des « catégories » ou des axes extraits des étiquettes afin de rassembler les productions de tous les sujets autour d'étiquettes empiriques analogues.

La première approche du travail des synergologues a été en partie concluante : il serait maintenant intéressant de reprendre en particulier, pour les stimuli bien reconnus, les indices de leur grammaire pour vérifier visuellement dans un premier temps si ces indices permettent d'extraire d'autres stimuli étiquetés de façon analogue. En parallèle, nous avons donné aux synergologues les résultats de leurs analyses et la comparaison avec les scores des naïfs. En particulier, l'informativité pour les naïfs de la modalité acoustique sera certainement pour les synergologues un retour utile.

Ce présent travail, aussi bien l'état de l'art et les résultats des études expérimentales, a ainsi essayé d'illustrer la portée des micro-expressions dans de multiples domaines ainsi que leurs caractéristiques dans le but de mettre à jour une nouvelle vision

de futurs enjeux applicatifs. En effet, un véritable engouement refait surface autour des sciences de l'affect et des émotions longtemps délaissées, car de véritables besoins émergent des sociétés au contact des évolutions technologiques. De plus, alors que nous étions dans l'impossibilité d'observer certaines données par manque de moyens ou d'outils, les avancées telles que l'imagerie médicale, la puissance des machines actuelles ou encore l'apparition de nouvelles ressources, nous permettent enfin de nous atteler aux gros problèmes que souligne le TAL qui est la compréhension et la modélisation du caractère naturel et humain des données qui nous permettraient de désambiguïser la complexité du langage. Les besoins sont partout et dans des domaines inimaginables : les nouvelles technologies de la communication, les agents conversationnels animés, la télécommunication et les sociétés sans écriture, le multilinguisme, les technologies d'aide à la personne, la compréhension des diverses pathologies, la domotique et l'aide aux personnes âgées, l'apprentissage des langues, l'apprentissage des langues par leur culture, les formations à distance, l'oralité et la rétroaction, l'apprentissage spécifique (par exemple la surdité)...

Pour le japonais en particulier, se développent actuellement de nombreuses nouvelles applications mettant en œuvre le type d'événements que nous avons étudiés (Hayashi, 1999 ; Nakagawa et al, 1995 ; Sagisaka et al, 1997 ; Minato, 2007 ; Gilleade et al, 2005).

## Bibliographie

Abdi, H.; (2004). Multivariate Analysis. In M. Lewis, A. Bryman, & T. Futing (Ed.), *Encyclopedia of social science research methods* (pp. 699-702.). Thousand Oaks: Sage.

Abdi, H., Williams, L.J., (2010), Correspondence Analysis. In N.J. Salkind, D.M. Dougherty & B. Frey (Ed.), *Encyclopedia of Research Design* (pp. 267-278). Thousand Oaks: Sage.

Aristote (2001). *Les politiques*. Paris :Flammarion.

Aristote (2007). *Rhétorique*, Paris: Pocket.

Arnold, M.B. (1960), Psychological Aspects. In Arnold, M.B. (Ed.), *Emotion and Personality* (Vol 1) (pp. 296). New York : Columbia University Press.

Aubergé, V., Rilliard, A., & Audibert, N. (2005) De E-Wiz à E-Clone : méthodologie expérimentale pour la modélisation des émotions et affects authentiques. In *Workshop Francophone sur les Agents Conversationnels Animés, (Grenoble, France, 13-14 Juin 2005)* (pp. 125-134)

Aubergé, V., Audibert, N. & Rilliard, A. (2006). De E-Wiz à E-Clone : recueil, modélisation et synthèse d'expressions authentiques. In *Revue d'Intelligence Artificielle – Numéro spécial « Interactions émotionnelles »*, (Vol. 20) (4-5), (pp. 499-528).

Audibert, N., Rilliard, A. & Aubergé, V. (2005). La plateforme E-Wiz (Expressive-Wizard of Oz) : capture d'expressions authentiques en Interaction Homme-Machine. In *1er Workshop francophone sur les Agents Conversationnels Animés (WACA'01) (Grenoble, France)* (pp. 161-164).

Averill, J.R. (1975), A semantic atlas of emotional concepts. In *JSAS, Catalog of Selected Documents in Psychology (Vol 5)*, (N°330) Amherst, USA: University of Massachusetts.

Borod, J.C. (1992) Interhemispheric and intrahemispheric control of emotion: a focus on unilateral brain damage, In *Journal of Consulting and Clinical psychology* (Vol 60) (pp 339-348).

Burckhardt, R.W. (2005) *Patterns of behavior, Konrad Lorenz, Niko Tinbergen and the founding of ethology*. USA: University of Chicago Press. (pp. 1-30)

Buser, P. (1998) *Cerveau de soi, cerveau de l'autre*. Paris : Odile Jacob. (pp. 291).

Calbris, G. (2003) *L'expression gestuelle de la pensée d'un homme politique*, Paris : CNRS Éditions (Communication) (pp. 205).

Campbell, N. (2000): Databases of emotional speech, In *SpeechEmotion-2000*. (pp. 34-38).

Campbell, N. (2007a) Differences in the speaking styles of a Japanese male according to interlocutor; showing the effects of affect in conversational speech, In *International Journal of Computational Linguistics and Chinese Language Processing* (3 mai 2007) (Vol. 11, no. 1) (pp.116)

Campbell N. (2007b) The Role and Use of Speech Gestures in Discourse, In *Archives of Acoustics (Warsawa)* (Vol 32) (N° 4, 803814)

Cannon, W.B. (1927) The James-Lange theory of emotions. In *American Journal of Psychology*, (Vol. 39) (pp. 115-124).

Colletta, J., Kunene, R., Venouil, A., Kaufmann, V., & Simon, J. (2009) Multi-track Annotation of Child Language and Gestures. In Kipp M., Martin J., Paggio P., & Heylen D. (Ed.) *Multimodal Corpora, Lecture Notes in Computer Science* (Vol. 5509, pp. 54-72). Berlin, Heidelberg: Springer.

Darwin, C. R. (1872). *The expression of the emotions in man and animals*. London: John Murray. 1st edition. Accès : <http://darwin-online.org.uk/content/frameset?itemID=F1142&viewtype=text&pageseq=1> (consulté en novembre 2011)

Damasio, A. (1995) *L'erreur de Descartes*, Paris : O. Jacob.

Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (2003) *Handbook of Affective Sciences*, New York: Oxford University Press.

De Biasi, G. (2011) *Etude de la perception des émotions dans les microévénements non verbaux chez l'être humain*. Mémoire de master Sciences du langage. Université Stendhal, Grenoble.

De Bonis, M. (1996) *Connaître les émotions humaines*, Liège : Mardaga.

Douglas-Cowie, E., Cowie, R., Schröder, M. (2000) A new emotion database: considerations, sources and scope. In *Proceedings of the ISCA Workshop on Speech and Emotion: A Conceptual Framework for Research, Northern Ireland, Belfast: NI: Textflow*.

Ekman, P.; Sorenson, E. R.; Friesen W.V. (1969) *Pan-Cultural Elements in Facial Displays of Emotion Science*, New Series, (Vol. 164) (No. 3875) (pp. 86-88).

Frijda, N. (1986) *The emotions*. Angleterre : Cambridge University Press.

Gabrielson, A., Juslin, P.N. (2003) Emotional expression in music. In Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (Ed.) *Handbook of Affective Sciences*, New York : Oxford University Press (pp. 503).

Gerrards-Hesse, A., Spies, K., Hesse, F.W. (1994) Experimental inductions of emotional states and their affectiveness: A review. In *British Journal of Psychology* (Vol. 85) (pp. 55-78).

Giese-David, J., Spiegel, D. (2003) Emotional expression and cancer expression. In Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (Ed.) *Handbook of Affective Sciences*, New York : Oxford University Press (pp.1053).

Gilleade, K., Dix, A., Allanson, J. (2005) Affective Videogames and Modes of Affective Gaming: Assist Me, Challenge Me, Emote Me. In: *Proceedings of DIGRA'2005*. Accès: <http://www.hcibook.com/alan/papers/DIGRA2005/>

Goldsmith, H., (2003), Genetics of emotional development, In Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (Ed.) *Handbook of Affective Sciences*, New York: Oxford University Press. (pp. 300).

Guaïtella, I. et al. Groupe de recherche Geste et Voix (1998) Les relations voco-gestuelles dans la communication interpersonnelle : émergence d'une problématique et carrefour interdisciplinaire. In Santi, S., Guaïtella, I., Cavé C. et Konopczynski, G. (Ed.) *Oralité et gestualité, communication multimodale, interaction : actes du colloque ORAGE'98*, Montréal : L'Harmattan.

Hayashi, Y. (1999) Recognition of vocal expression of mental attitudes in Japanese: Using the interjection "eh". In *International Congress of Phonetic Sciences (San Francisco)* (pp. 135-139)

- Insel, T.R. (2003) The neurobiology of affiliation: implications for autism. In Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (Ed.) *Handbook of Affective Sciences*, New York: Oxford University Press. (pp. 1010)
- Ishi, C.T., Hatano, H., & Hagitari, N. (2012) Extraction of paralinguistic information carried by mono-syllabic interjections in Japanese, In *Proceedings of Speech Prosody*, à paraître.
- Izard, C. E. (1980) Cross-cultural perspectives on emotion and emotion communication. In Triandis H. & Lonner W. (Ed.), *Handbook of cross-cultural psychology: Basic processes* (Vol. 3) Boston: Allyn & Bacon. (pp. 185-222).
- Jonstone, T., Scherer, K. (1999) The effects of emotions on voice quality, In *Proceedings of the XIVth International Congress of Phonetic Sciences (San Francisco)* (pp. 2029-2032).
- Kaiser, S., Wherle, T. (1994) Emotion research and AI: some theoretical and technical issues. In *Geneva studies in emotion and communication* (Vol. 8) (pp. 1-16).
- Kimura D et Humphrys C.A. (1981) A comparison of left and right arm movement during speaking, In *Neuropsychologia*, (Vol. 19) (pp. 807-812).
- Kleinginna, P.R., Kleinginna, A.M., (1981) A categorized list of emotion definitions with suggestions for a consensual definition, In *Motivation and Emotion*. (chap. 5, pp 345-346).
- Lazarus, R. S. (1966) *Psychological stress and the coping process*. New York: McGraw-Hill.
- Lewis, M., Haviland-Jone, J.M. (2000) *Handbook of emotions, Second edition*, New York: The Guilford Press.
- Loyau, F. (2007) *Expressions des états mentaux et émotionnels de l'humain en interaction : ébauches du « Feeling of Thinking »*. Thèse en sciences cognitives. Institut National Polytechnique de Grenoble.
- Loyau, F., & Aubergé, V (2006) *Expressions outside the talk turn: ethograms of the feeling of thinking*. In *5th LREC* (pp. 47-50).
- Magno Caldognetto, E., & Poggi, I. (1997) Micro- and Macro-Bimodality. In Benoit C. and Campbell R. (Ed). *Proceedings of the ESCA Workshop on Audio-Visual Speech Processing (AVSP'97), (Rhodes, Greece, September 1997)*.
- Maekawa, K. (2004) Production and perception of 'Paralinguistic' information. In *Proceedings of Speech Prosody 2004 (Nara, Japan)* (pp. 367-374).
- Matsumoto, D., Ekman, P. (1989) American-Japanese cultural differences in ratings of facial expressions of emotion, In *Motivation and emotion*, (chap 13, pp. 143-157).
- Matsumoto, D., Yoo Hee, S., & Fontaine, J. (2008) Mapping Expressive Differences Around the World: The Relationship Between Emotional Display Rules and Individualism Versus Collectivism. In *Journal of Cross-Cultural Psychology*, (Vol. 39-1) (pp.55-74).
- Matsumoto, Y. (1988) From Bound Grammatical Markers to Free Discourse Markers: History of some Japanese Connectives. In *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 340-351).
- Mandler, G. (1984) *Mind and body: Psychology of emotion and stress*, New York: Norton.

- Minato, J. (2007) *Corpus-based Analysis of Japanese-English Emotional Expressions*, Tokushima: Tokushima Univ. (pp. 413-441)
- Naito, S. (2002) Difficultés linguistiques et culturelles dans l'enseignement du français à un public japonais, In *Ela. Études de linguistique appliquée*, (Vol. 2002/2 n° 126) (pp.145-155).
- Nakagawa, S., Kobayashi, S. (1995) Phenomena and acoustic variation on interjections, pauses and repairs in spontaneous speech. In *J. Acoustical Society of Japan* (Vol. 51, No. 3) (pp. 202-210).
- Plutchik, R. (1977) Cognitions in the Service of Emotions : An Evolutionary Perspective. In Candland D.K., Fell J.P., Keen E., Leshner A.I., Plutchik R., Tarpay R.M. (Ed.) *Emotions*. Monterey: Brooks/Cole (pp. 529-553).
- Pettorino, M., Giannini, A., Chiari, I., Dovetto F.M. (2010) *Spoken Communication*. Cambridge : Cambridge Scholar Publishing.
- Popper, K. (1972) *La logique de la découverte scientifique*, Paris : Payot.
- Reder, L.M., Ritter, F.E. (1992) What determines initial feeling of knowing? Familiarity with question terms, not with the answer, In *Journal of Experimental Psychology: Learning Memory, and Cognition* (Vol 18-3) (pp. 435-451).
- Sagisaka, Y., Campbell, N., Higuchi, N. (1997) *Computing prosody: computational models for processing spontaneous speech*. New York: Springer.
- Saussure, F. (1995) *Cours de linguistique générale*. Paris : Payot.
- Schröder, M. (1998) *L'expression vocale de l'amusement : premières expériences audiovisuelles*. TER. Maîtrise Sciences du Langage, Université Stendhal Grenoble 3, ICP (institut de la communication parlée)
- Shanahan, J.G., Qu, Y., Wiebe, J. (2006) *Computing attitude and affect in text : Theory and applications*, Dordrecht : Springer.
- Scherer, K.R. (2001) Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Ed.). *Appraisal processes in emotion: Theory, Methods, Research*. New York and Oxford: Oxford University Press. (pp. 92-120).
- Signorello, R. (2010) *La parole ne s'arrête pas aux mots : à la recherche d'indices de culture dans le backchannel audio-visuel*. Mémoire de master en sciences du langage. Université Stendhal, Grenoble.
- Skinner M et Mullen B (1991) Facial assymetry in emotional expression : A meta-analysis of research. In *British Journal of psychology* (Vol.30) (pp. 113-124).
- Solomon, R.C., (2000), The philosophy of emotions, In Lewis M., Haviland-Jone J.M. (Ed.) *Handbook of emotions, Second edition*. New York: The Guilford Press. (pp. 3-15).
- Thivel, A. (1997) Hippocrate et la théorie des humeurs. In *Noesis* (en ligne, N°1 | 1997). Mis en ligne le 02 mars 2009. Accès: <http://noesis.revues.org/index1419.html> (consulté le 2 mai 2012)
- Tinbergen, J. (1972) *Early childhood autism - an ethological approach*, Berlin & Hamburg: Parey.

Tomkins, S.S. (1980) Affect as amplification: Some modifications in theory. In R. Plutchik & H. Kellerman (Ed.) *Emotion: Theory, research and experience* (Vol. 1). San Diego: Academic Press. (pp. 141-187).

Ishi, C.T., Hatano, H., Hagita, H. (2012) Extraction of paralinguistic information carried by monosyllabic interjections in Japanese, In *Proceedings of Speech Prosody*, à paraître.

Turchet, P. (2000) *La synergologie*. Canada : les éditions de l'Homme.

Vanpé, A. (2011) *Expressions et micro-expressions spontanées de la face et de la voix en Interaction Homme-Machine: esquisse d'un modèle du "Feeling of Thinking"*. Thèse dans le domaine de Sciences de l'Homme et Société, Informatique. Université Stendhal, Grenoble.

Vanpé, A., Aubergé, V. (2012) Early meaning before the phonemes concatenation? Prosodic cues for Feeling of Thinking. In *VII GSCP international conference: speech and corpora*. Accès : [http://www.lettras.ufmg.br/gscp2012-eng/data1/arquivos/gscp2012\\_submission\\_100.pdf](http://www.lettras.ufmg.br/gscp2012-eng/data1/arquivos/gscp2012_submission_100.pdf) (consulté en 13 mai 2012)

Wierzbicka, A. (1999) *Emotions Across Languages and Cultures: Diversity and Universals*. Cambridge: Cambridge University Press.

Williams, C.E., Stevens, K.N. (1972) Emotions and speech: some acoustical correlates. In *Journal of the Acoustical Society of America*, (Vol. 52-4) (pp. 1238-1250).

# Table des matières

<b>AVANT-PROPOS</b> .....	<b>4</b>
<b>REMERCIEMENTS</b> .....	<b>7</b>
<b>INTRODUCTION</b> .....	<b>9</b>
<b>PARTIE 1 ETAT DE L'ART</b> .....	<b>11</b>
<i>Chapitre 1 - Présentation du laboratoire</i> .....	12
<i>Chapitre 2 – Survol de la communication expressive</i> .....	14
<i>Réflexions philosophiques sur les émotions</i> .....	14
L'évolution de la grande opposition entre raison et émotion .....	14
Les premières définitions et notion d'éthique .....	15
Les émotions et le corps .....	15
La source des émotions passant du corps vers l'esprit .....	16
Les considérations philosophiques cognitives des émotions .....	16
Une description favorable et précise des émotions sans théorisation .....	17
La philosophie moderne et les émotions .....	17
<i>Les principales théories des émotions dans le domaine de la psychologie</i> .....	18
Avant de rentrer dans les grands courants des émotions .....	18
Le courant évolutionniste .....	19
Le courant physiologique .....	21
Les théories cognitivistes .....	21
Le constructivisme social des émotions .....	22
Interdisciplinarité, sciences affectives et nouvelles réflexions .....	23
<i>Conclusion du Chapitre 2</i> .....	24
<i>Chapitre 3 - Les Travaux précédents à l'origine du mémoire</i> .....	25
Les différentes méthodes de recueil des micro-expressions .....	25
Recueil de corpus E-Wiz avec le protocole Sound Teacher .....	27
<i>Les motivations d'études sur les micro-expressions et terminologie</i> .....	29
Le Feeling of Thinking .....	29
Quelques dichotomies à noter sur les micro-expressions .....	29
<i>Premières observations et annotations de corpus</i> .....	30
Des méthodes d'observations basées sur l'éthologie .....	30
Les différentes annotations du corpus .....	31
Les perspectives d'études mis en évidence par A. Vanpé .....	32
Une étude perceptive culturelle des micro-expressions .....	32
La dimension socio-affective dans un test de perception .....	33
<i>Conclusion du Chapitre 3</i> .....	34
<b>PARTIE 2 EXPERIENCES PERCEPTIVES</b> .....	<b>35</b>
<i>Chapitre 4 – Test de perception de micro-expressions par des sujets japonais</i> .....	36
<i>Une manipulation répondant aux perspectives d'études antérieures</i> .....	36
<i>Protocole et outil expérimental</i> .....	37
Les outils utilisés .....	37
Description générale .....	38
Interfaces du test de perception .....	40
Adaptation du test en Japonais et modification algorithmique .....	43
<i>Contexte de recueil de données</i> .....	45
Les contraintes de distance avec les sujets .....	45
Les solutions apportées aux contraintes d'étude à distance .....	45
<i>Profils des sujets japonais</i> .....	48
Questionnaire .....	48
Description générale des sujets .....	49
Connaissances et expériences linguistiques .....	50
Séjours en France .....	52
La proximité avec des Français de pratique linguistique variée .....	54
La langue française dans les loisirs des sujets japonais .....	56

<i>Conclusion sur le chapitre 4</i> .....	57
PARTIE 3 ANALYSE DES RESULTATS EXPERIMENTAUX .....	58
<i>Chapitre 5 – Analyse des résultats japonais</i> .....	59
<i>Matrices de confusions et analyses statistiques</i> .....	59
Quelques précisions sur les matrices de confusion .....	59
Objectivation des données par des outils statistiques .....	60
Lecture de données statistiques .....	63
Les principales observations sur les différents sujets .....	64
Synthèse des principales observations .....	66
Remarques sur les faibles valeurs .....	67
<i>Les degrés de confiance des réponses des sujets</i> .....	68
Représentation colorimétrique des degrés de confiance.....	68
Corrélation entre degré de confiance et score de bonnes réponses.....	70
Observations des scores et écarts-types les graphiques de corrélation .....	70
La différence des tendances de corrélation entre hommes et femmes.....	71
<i>La perception culturelle des micro-expressions</i> .....	72
<i>Conclusion du chapitre 5</i> .....	73
<i>Chapitre 6 – Analyse comparative franco-japonaise</i> .....	75
<i>Analyse comparative sur les taux de bonnes réponses générales</i> .....	75
<i>Explication de l’analyse comparative étiquette par étiquette</i> .....	76
<i>Méthodologie</i> .....	76
<i>Nomenclature d’analyse</i> .....	77
<i>Observations sur les stimuli du sujet T</i> .....	77
La déception.....	77
Hésitation.....	77
(Assez) Calme.....	78
Etonnée .....	78
Angoissée.....	78
Synthèse des observations du sujet T .....	78
<i>Observations sur les stimuli du sujet S</i> .....	78
Mission Impossible .....	78
Le but est de trouver des solutions .....	79
Surprise par les résultats .....	79
Au pif, une envie de rigoler.....	79
Toujours la concentration.....	79
Déçue par les résultats.....	79
Synthèse des observations du sujet S .....	80
<i>Observations sur les stimuli du sujet M</i> .....	80
Peur .....	80
Doute.....	80
Pas contente-déçue.....	80
Sûr de moi / Bien à l’aise .....	81
Stress.....	81
Déception.....	81
Synthèse des observations du sujet M.....	81
<i>Observations sur les stimuli du sujet J</i> .....	81
Petite lassitude / l’indifférence s’installe.....	81
Petite irritation .....	82
Inquiétude sur le but du jeu.....	82
Je commence à comprendre qu’on se paye ma tête .....	82
Synthèse des observations du sujet J.....	82
<i>Observations sur les stimuli du sujet N</i> .....	83
Déception mais m’en m’amuse .....	83
Léger doute .....	83
Soulagé .....	83
Essaie de rester concentré .....	83
Ne suis plus sûr de moi, léger agacement .....	83
Synthèse des observations du sujet N .....	84
<i>Observations sur les stimuli du sujet R</i> .....	84
Concentration.....	84

Ennui.....	84
Ennui-concentration.....	84
Ennui, incompréhension.....	85
Mon incompréhension est justifiée.....	85
Agacé.....	85
Surprise, nervosité / Surprise, agacé.....	85
Synthèse des observations du sujet R.....	86
<i>Chapitre 7 – Les observations de synergologie.....</i>	<i>87</i>
<i>Aperçu de la synergologie.....</i>	<i>87</i>
Contexte et intérêts pour la synergologie.....	87
Présentation générale de la discipline.....	88
Les méthodes d’observations de la synergologie.....	89
Les approches sur lesquelles se basent la synergologie.....	90
<i>Recueil d’analyse de synergologie.....</i>	<i>90</i>
Contexte de recueil et présentation des sujets synergologues.....	91
Protocole d’analyse imposé aux synergologues.....	92
<i>Résultats des questionnaires.....</i>	<i>93</i>
Les observations sur le sujet T.....	93
Les observations sur le sujet S.....	94
Les observations sur le sujet M.....	96
Les observations sur le sujet J.....	97
Les observations sur le sujet N.....	99
Les observations sur le sujet R.....	100
<i>Conclusion du Chapitre 7.....</i>	<i>102</i>
<b>PARTIE 4 PERSPECTIVES ET ENJEUX APPLICATIFS.....</b>	<b>103</b>
<i>Chapitre 8 – Bilan de l’étude et perspectives.....</i>	<i>104</i>
<i>Les problèmes et difficultés mis en évidence.....</i>	<i>104</i>
Les difficultés que souligne l’état de l’art.....	104
Limites et problèmes des protocoles expérimentaux.....	105
Les non-dits de l’analyse.....	105
<i>Les perspectives d’amélioration des problèmes soulignés.....</i>	<i>106</i>
Les choix et une pluridisciplinarité ouverte mais contrôlée.....	106
Les améliorations des protocoles de perception et de synergologie.....	106
L’analyse acoustique des données.....	106
<b>CONCLUSION.....</b>	<b>107</b>

**MOTS-CLÉS** : micro-expressions de la parole, synergologie, perception, attitudes sociaux-affectifs, TAL (Traitement Automatique de la Langue), enjeux applicatifs

## **RÉSUMÉ**

L'évolution des habitudes sociales crée chaque jour de nouveaux besoins dans le domaine du Traitement Automatique de la Langue. Plus particulièrement dans le cas des technologies de la parole, l'un des enjeux majeurs est l'analyse des bruits non lexicalisés, accompagnés de gestes spécifiques, prémisses du langage humain. Ces objets sont désignés sous le terme de micro-expressions de la parole dans le cadre de cette étude. En reconnaissance, ces objets permettent de comprendre les attitudes socio-affectives de l'humain qui sont au cœur des nouveaux enjeux applicatifs actuels. Ainsi, si nous avons d'ores et déjà des pistes d'investigation sur les aspects acoustiques de ces micro-expressions, l'analyse visuelle en est une toute autre affaire. Il est alors intéressant de voir comment les nouveaux courants à la mode tels que la synergologie peut servir à une recherche scientifique. L'étude en cours vise à comprendre tous ces mécanismes et ces enjeux à travers une expérience de perception des micro-expressions de la parole en contraste français-japonais.

**KEYWORDS** : speech micro-expressions, synergologie, perception, social affective attitudes, NLP (Natural Language Processing), applications strategies

## **ABSTRACT**

The changing of social habits creates every day new needs in the area of Natural Language Processing. Especially in the case of speech technologies, one of the major challenges is the analysis of non-lexicalized sounds, accompanied by specific gestures, premises of human language. These objects are referred to as speech micro-expressions in this study. In recognition, these objects provide an understanding of the socio-emotional attitudes of humans which are the heart of the new challenges of modern applications. If we already have some ideas to investigate acoustic aspects of these micro-expressions, visual analysis is another matter. It is interesting to see how new movements in fashion as synergologie can be used in a scientific research. The current study aims to understand these mechanisms and issues through an experiment of speech micro-expressions perception in a French-Japanese contrast.

DECLARATION

1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

Signature :

SASA Yako  
Le 14/05/2012



Mémoire de Master 2 Sciences du Langage – Industries de la Langue. Université Stendhal Grenoble 3.  
Soutenu le 16/05/2012